SYSTEMATIC CHARACTERIZATION AND MODELING OF SMALL AND LARGE SIGNAL PERFORMANCE OF 50 - 200 GHZ SIGE HBTS

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.

_____

Jun Pan

Certificate of Approval:

_____	_____
Stuart M. Wentworth	Guofu Niu, Chair
Associate Professor	Professor
Electrical and Computer Engineering	Electrical and Computer Engineering
_____	_____
Foster Dai	Stephen L. McFarland
Associate Professor	Acting Dean
Electrical and Computer Engineering	Graduate School

SYSTEMATIC CHARACTERIZATION AND MODELING OF SMALL AND LARGE SIGNAL PERFORMANCE OF 50 - 200 GHZ SIGE HBTS

Jun Pan

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 8, 2005

VITA

Jun Pan, son of Cunguan Pan and Lanfang Lu, was born on 7 October, 1977, in DongTai, JiangSu Province, P. R. China. He received his BS degree from Jilin University in 1997, majoring in Physics. He received his MS degree from Nanjing University in 2000, majoring in Electrical Engineering. In Spring 2001, he entered the Electrical and Computer Engineering department of Auburn University, Auburn, Alabama, where he has been pursuing his Ph.D degree.

DISSERTATION ABSTRACT

SYSTEMATIC CHARACTERIZATION AND MODELING OF SMALL AND LARGE SIGNAL PERFORMANCE OF 50 - 200 GHZ SIGE HBTS

Jun Pan

Doctor of Philosophy, August 8, 2005
(M.S., Nanjing University, 2000)
(B.S., Jiling University, 1997)

134 Typed Pages

Directed by Dr. Guofu Niu

Over the last ten years, SiGe BiCMOS technology has become the leading technology in analog circuit design for both wired and wireless telecommunication applications. However, the endless pursuit for high performance is fraught with difficulties in characterizing and modeling the SiGe HBTs. In order to obtain a high cut off frequency f_T, devices are scaled to extremes and collector doping is increased to allow more current flow before the onset of the Kirk effect, allowing the high speed benefit of a smaller base transit time to be realized. As a result, self-heating plays an important role in the already complicated HBT characteristics.

A good example of this is the characterization of avalanche multiplication. As an inevitable result of pursuing high f_T, the breakdown voltage is decreased, which makes the characterization of avalanche multiplication more important. However, with severe self-heating, conventional methods fail at a practical bias (below peak f_T current). Chapter 2 gives a review of the measurement methods currently available for characterizing avalanche multiplication in SiGe HBTs. With the scaling of devices, conventional methods can no longer be applied. New methods are proposed to accurately measure avalanche multiplication factor (M-1) even in the severe

self-heating region. The current dependence of $M-1$ is demonstrated. The results show that the CB breakdown voltage at the J_E of peak f_T is higher than that at either low J_E or in the off state by a significant 1 V in a 120 GHz peak f_T device. Also in Chapter 2, the current dependence of $M-1$ is found to be considerably smaller by taking into account the extrinsic collector resistance. Later, in Chapter 5, a simplified model for the current dependent $M-1$ is proposed.

In Chapter 3, RF characterization methods are discussed, including S-parameters characterization, large signal power characterization and third order intermodulation characterization. The S-parameters are measured using a Vector Network Analyzer and a DC power supply. In general, S-parameters characterize small signal parameters and can be used to extract the base resistance R_B, the cut-off frequency f_T and the maximum oscillation frequency f_{max}. The large signal system built utilizes the same equipment setup used for the S-parameter system, while being able to measure device performance at large input power. By monitoring the DC voltage and current, the power added efficiency can be calculated. The third order intermodulation system built is more complicated than the previous two systems in the sense that the system distortion level can significantly affect the accuracy of the measurement results. With care, accurate intermodulation measurements are realized for HBTs and MOSFETs. All of the three systems are controlled by in-house programs written in VEE. The program for the S-parameter system was written by William E. Ansley. Modification of the program is made in this work to increase the stability of operation and the application range of the program. Programs for the large signal system and the IIP3 system are written during the construction of both systems. All of the programs are now capable of measuring the characteristics of both HBT and MOSFET devices with high accuracy while requiring little attention from operators.

The large signal performance characterization of SiGe HBTs is of great importance to both circuit design and process design. Chapter 4 experimentally investigates SiGe profile and collector profile optimization from a large signal performance standpoint, as well as the impact of technology scaling. The results show that device and circuit designs that only consider optimum small signal performance could inadvertently degrade large signal performance. The tradeoffs in SiGe profile design between small signal and large signal performance, as well as the impact of speed-breakdown tradeoff on large signal performance, are experimentally examined. The SiGe HBTs from a 200 GHz technology show impressive small and large signal performance at 20 GHz, demonstrating the benefits of technology scaling, despite decreased breakdown voltage.

Intermodulation linearity is an another important figure-of-merit for SiGe HBTs, as it relates to the selectivity of an RF receiver and the spectral purity of an RF transmitter. Chapter 5 presents a systematic characterization of the intermodulation linearity for SiGe HBTs in order to gain insight into the device physics underlying linearity behavior, and to construct guidelines to optimal sizing, biasing, and device selection (e.g. high breakdown versus low breakdown versions). The input 3rd order intercept point, IIP3, is measured on $I_C - V_{CE}$ plane for devices of various size, breakdown voltage, Ge profile, and technology generation. Later in this chapter, problems of VBIC model for simulating IIP3 are presented. Improvements for base collector capacitance and avalanche modelling in the VBIC model are suggested and implemented in Verilog-A to give a much better fit to the measurement results.

ACKNOWLEDGMENTS

I would like to thank my advisor, Dr. Guofu Niu, for his guidance, advice, and support throughout the completion of this dissertation, without which this work would not have materialized. His constant encouragement and interest strengthened me during my toughest times. His dedication to excellence will always be my example. I would like to thank Dr. Stuart M. Wentworth for introducing me to RF measurements and to express my appreciation to his encouragement and technical assistance. I am also grateful to Dr. Foster Dai for his valuable technical insights and suggestions. Special thanks are due to Dr. John D. Cressler for helping me enroll at Auburn University.

Several people deserve special recognition for their contributions to this work. I would like to thank Jin Tang, Shiming Zhang and Yuan Lu for helping me with the S-parameter measurement and Zhiming Feng for helping me with Verilog-A modules for simulation. I would also like to thank all the members of my research group for their friendship and their contributions to this work.

Thanks are also due for the financial support provided by the National Science Foundation under ECS-0119623 and ECS-0112923, the Semiconductor Research Corporation under SRC #2001-NJ-937, and an IBM Faculty Partner Award. I would like to thank IBM's SiGe team for providing wafers from various technology generations.

Finally, I sincerely thank Zhiyun Luo for her support and encouragement throughout the course of this work and also for her emotional support and the patience she has shown over the past few years.

Style manual or journal used IEEE Transactions on Electron Devices (together with the style known as "aums"). 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 `aums.sty`. The plots were generated using VossPlot®, MATLAB®and Microsoft Visio®.

TABLE OF CONTENTS

List of Figures

LIST OF TABLES

CHAPTER 1

INTRODUCTION

Over the last ten years, SiGe BiCMOS technology has become the leading technology in analog and RF circuit design for both wired and wireless telecommunication applications. Since the first generation technology of 50 GHz peak cut off frequency (f_T), SiGe heterojunction bipolar junction (HBT) devices have faced challenges on both speed and power from other technologies [1]. The newly announced 350 GHz peak f_T SiGe technology has answered the speed challenge [2], while higher power devices incorporating SiGe HBTs are also being fabricated, as reported by recent papers [3]. At present, it appears that SiGe BiCMOS technology is the current force to be reckoned with and is also the trend for the future.

With the endless pursuit of high performance comes difficulties in characterizing and modeling the SiGe HBTs. In order to obtain high f_T, devices are scaled to extremes and collector doping is increased to allow more current to flow before the onset of the Kirk effect to allow the high speed benefit of smaller base transit time to be realized. This makes self-heating an important concern. A good example of this is the characterization of avalanche multiplication. As an inevitable result of pursuing high f_T, the breakdown voltage is decreased, which makes characterizing avalanche multiplication more important. However, due to severe self-heating, conventional methods fail at a practical bias (below peak f_T current). In chapter 2, we will present two new methods to solve this problem.

Large signal performance characterization of SiGe HBTs is of great importance for both circuit design and process design. In this work, the tradeoffs between large signal performance and small signal performance when designing SiGe HBTs will be investigated.

Intermodulation linearity is another important figure-of-merit for SiGe HBTs, as it relates to the selectivity of an RF receiver and the spectral purity of an RF transmitter. In Chapter 5, a systematic characterization of the intermodulation linearity for SiGe HBTs will be presented. Problems with the VBIC (Vertical Bipolar Inter-Company) model for simulating IIP3 are also described in this chapter and suggestions given for using Verilog-A to implement a current dependent collector base charge model.

1.1 SiGe HBT Basics

An interesting thing in the history of bipolar junction transistors is that the first BJT is actually fabricated from a bar of germanium with two closely spaced alloyed contacts. In December of 1987, after many years of using Si to build BJTs, the first functional SiGe HBT was demonstrated [4]. In June of 1990, a non-self-aligned SiGe HBT grown by ultra-high vacuum/chemical vapor deposition (UHV/CVD) was demonstrated with a peak cutoff frequency of 75 GHz, which is about twice the performance of state-of-art Si BJTs [5]. Since then, the bright future of SiGe HBT has been embraced by the laboratories all around the world.

Adding Ge to Si BJT introduced a number of exciting performance improvements. The base region of SiGe HBTs is typically the region where SiGe alloy is used instead of Si. The basic operational principle of SiGe HBT can be best understood by considering the band diagram shown in Figure 1.1. The Ge mole fraction is graded from the emitter towards the collector, creating an accelerating electric field in the neutral base.

The important *dc* consequence of adding Ge into the base lies with the collector current density (J_C). The Ge-induced band offset exponentially increases the intrinsic carrier density in the base, which in turn decreases the effective Gummel number and hence increases J_C [6]. Because the emitter region is the same, the base current density J_B is roughly the same for SiGe

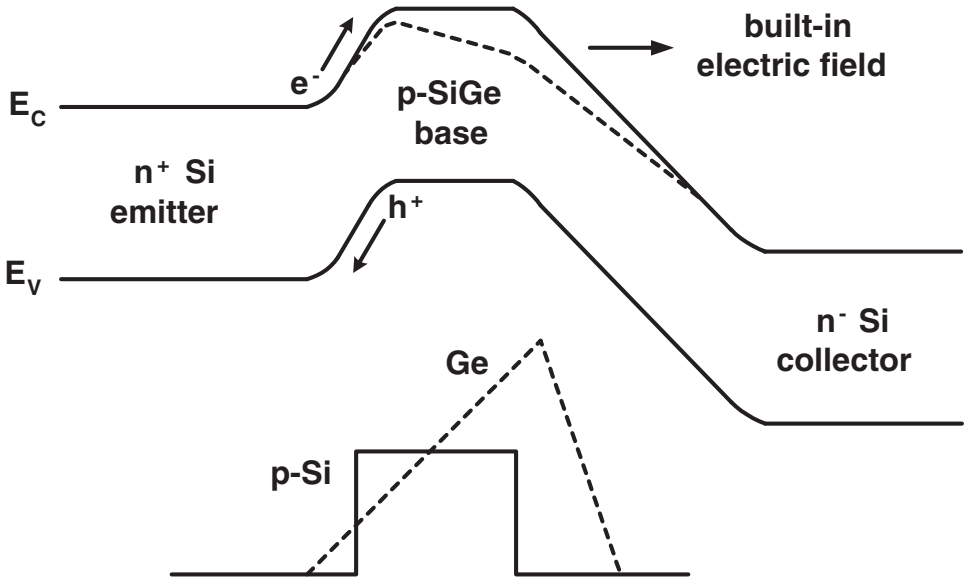


Figure 1.1: Energy band diagram of a graded-base SiGe HBT compared to a Si BJT.

HBT and Si BJT. Figure 1.2 compares the Gummel characteristics for the SiGe HBT and the comparably constructed Si BJT used in this work. The J_C of the SiGe HBT is much higher than that of the Si BJT, while the J_B of these two devices are similar. As a result, current gain β, defined as J_C/J_B, is higher in the SiGe HBT.

In most RF and microwave circuit applications, it is the transistor frequency response that limits system performance. An important frequency response figure-of-merit is the unity-gain cutoff frequency (f_T), which is given by

$$\frac{1}{2\pi f_T} = \tau_b + \tau_e + \tau_c + \frac{1}{g_m}(C_{te} + C_{tc}), \tag{1.1}$$

where τ_b, τ_c and τ_e are base, collector and emitter transit time, respectively, g_m is the transconductance, C_{te} and C_{tc} are EB and CB junction depletion capacitances. In conventional Si BJT's,

τ_b typically limits the maximum f_T. The built-in electric field induced by the Ge grading across the neutral base aids the transport of minority carriers (electrons) from emitter to collector, leading to faster base transit and thus higher f_T. Figure 1.3 compares the f_T characteristics for a SiGe HBT and a comparably constructed Si BJT. The f_T for the SiGe HBT is indeed much higher than the Si BJT. Another figure-of-merit for RF applications is the maximum oscillation frequency, f_{max}:

$$f_{max} = \sqrt{\frac{1}{8\pi C_{tc}} \frac{f_T}{r_b}}, \tag{1.2}$$

which indicates that the f_T/r_b ratio must be increased to improve f_{max} or transistor power gain.

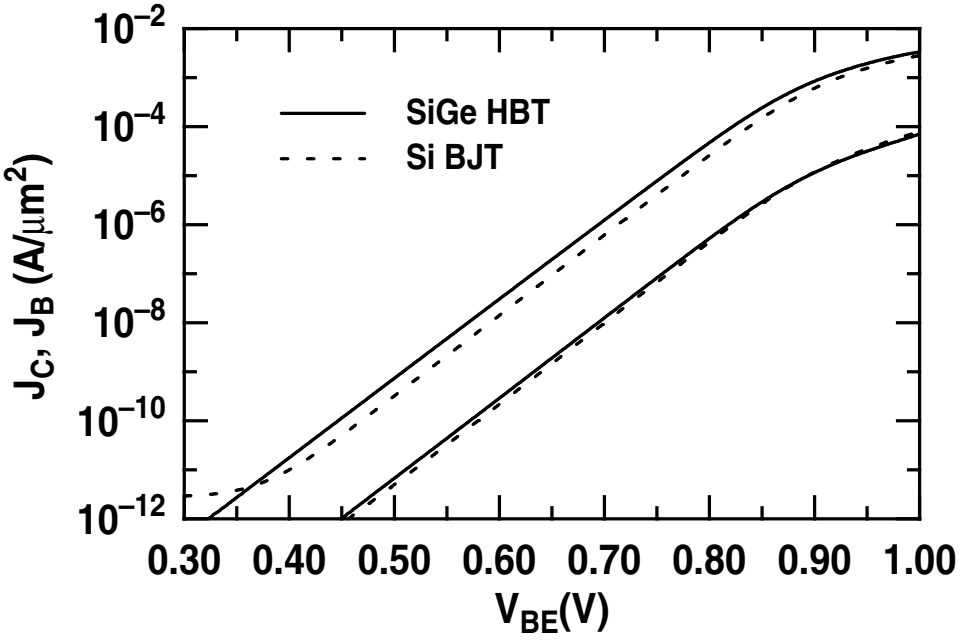


Figure 1.2: Comparison of gummel characteristics for a SiGe HBT and a Si BJT.

For double base contact, base resistance r_b is related to the product of base doping (N_B) and base width (W_B) by:

$$r_b = \frac{1}{12}\frac{1}{q\mu_p}\frac{X_E}{L_E}\frac{1}{N_B W_B}, \tag{1.3}$$

where X_E is the lateral emitter width, L_E is the lateral emitter length, and μ_p is the lateral mobility of holes in the base. With bandgap engineering, r_b can be reduced by increasing base doping N_B without compromising β, because β can be increased by ΔE_g exponentially.

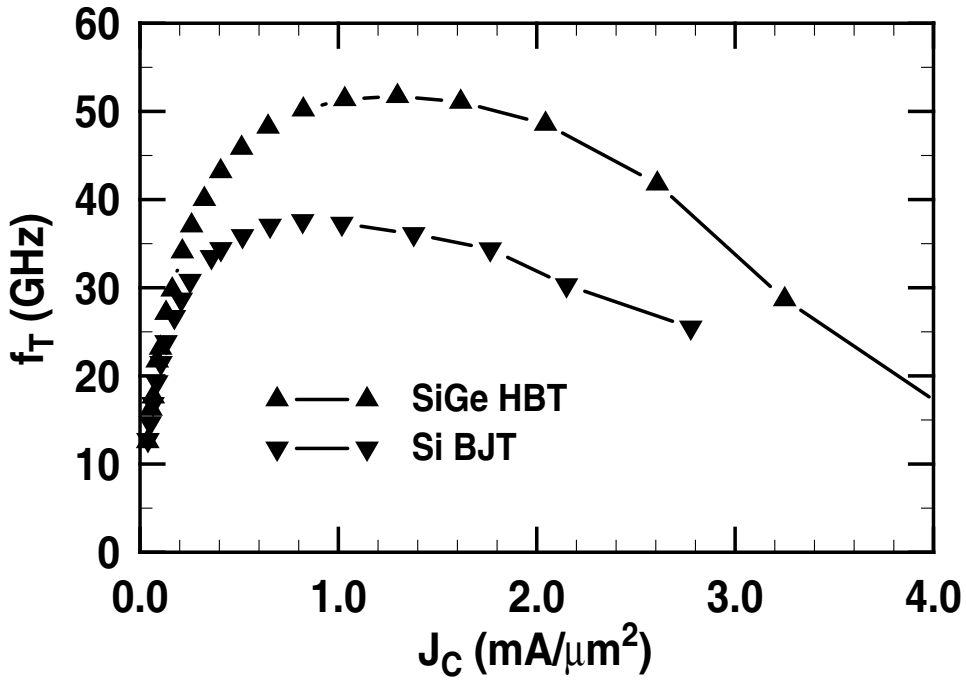


Figure 1.3: Comparison of f_T characteristics for a SiGe HBT and a Si BJT.

Figure 1.4 shows a schematic cross-section of a SiGe HBT used in this work. The SiGe HBT has a planar, self-aligned structure with a conventional poly emitter contact, silicided extrinsic base, and deep- and shallow-trench isolation. The extrinsic resistive and capacitive parasitics are intentionally minimized to improve the maximum oscillation frequency (f_{max}) of the

transistor. Details of the fabrication process can be found in [1]. The SiGe base was grown using the ultra high vacuum/chemical vapor deposition (UHV/CVD) technique [7]. Devices of different breakdown voltages were obtained on the same chip in the same fabrication flow by selective implantation during collector formation.

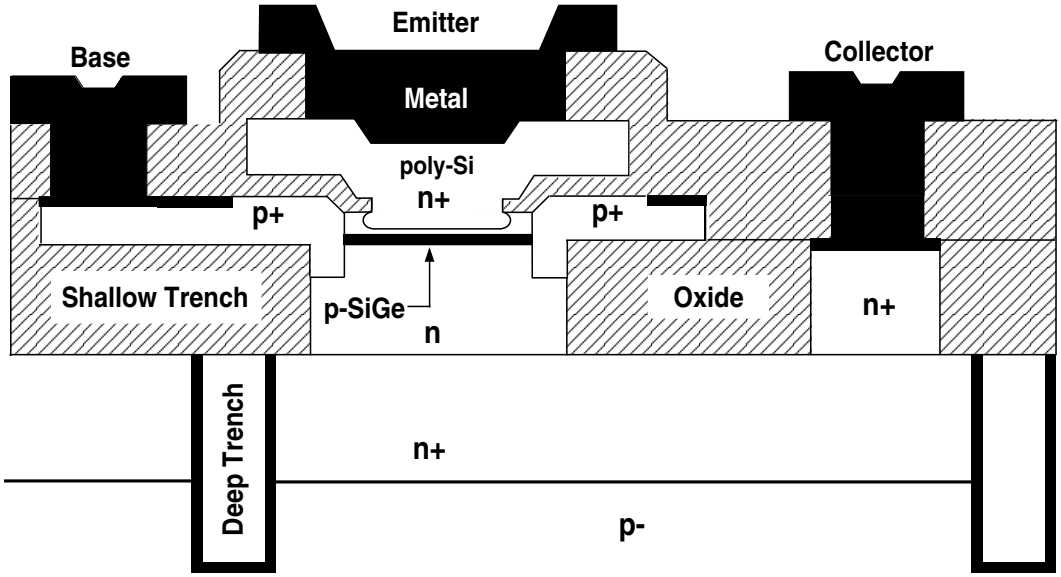


Figure 1.4: Schematic cross-section of a SiGe HBT used in this work.

1.2 SiGe Applications and Leverage of SiGe HBTs for RFIC

The SiGe HBT marketplace covers a wide range of product requirements. SiGe products are now appearing in virtually all analog and high-frequency market segments. SiGe is used in wireless cellular CDMA and GSM standards at 900 MHz and 2.4 GHz, both in handsets and base stations [8]- [9]. Wireless local area network (WLAN) chipsets at 2.4 GHz have been announced where the use of SiGe reduced the IC chip count and power consumption by 50%. SiGe is also widely used in high-speed/high-capacity network applications including a 10 Gbps synchronous

optical network (SONET) transmit-and-receive module [10], trans-impedance amplifiers, and 1-2.5 Gb/s Ethernet applications. 40 Gbps applications are now being produced utilizing SiGe bipolar and BiCMOS production technologies with f_T of 120 GHz. Ring-oscillators with 4.2-ps ECL delays have been demonstrated using 200 GHz f_T SiGe technology [11]. Other applications include products with lower level of integrations such as LNAs, VCOs, mixers, power amplifiers (PA), and GPS receivers.

There is a wide variety of wireless communication systems today: from pagers, cordless to analog and digital cellular telephones, satellite television services which utilize RF integrated circuits. The explosion for wireless communication has created a rapidly expanding market for RFIC's. The design requirements imposed on these transceiver components are truly challenging and multi-dimensional, including power dissipation, speed, noise (both broadband and near carrier), linearity, gain. As a result, the optimum IC technology choice for RF transceivers, in terms of optimum devices and levels of integration, is still evolving.

SiGe provides an exceptional integrated solution to the demanding wireless communication market when combined with RF elements made possible by a good BiCMOS process. High-isolation structures can be realized by combing deep trench, metal layers, and the active layers. Si has high thermal conductivity compared to GaAs, which allows denser integration, power devices and higher reliability. Such a process can satisfy the requirements of RF circuits, mixed-signal circuits, and precision analog circuits.

The heart of the SiGe BiCMOS process, SiGe HBT, has the desired performance in the areas of gain, noise, V_{BE} matching, and linearity [12]. The good gain and linearity offer current savings — an important aspect of portable electronics that are powered by batteries. The HBT needs to have a high degree of V_{BE} matching to enable, for example, precision current mirrors used in biasing, the reduction of offset voltage in operational amplifiers, and in general, to increase

product yields. The process of the SiGe HBT ensures consistent V_{BE} to meet these circuit demands [13].

1.3 Review of This Dissertation

This dissertation focuses on the characterization and modeling of some important SiGe HBT parameters with the discussion of the implications to analog circuit design.

Chapter 2 gives a review of measurement methods for characterizing avalanche multiplication in SiGe HBTs. With the scaling of device, conventional methods fail at practical bias. New methods are proposed to accurately measure avalanche multiplication factor (M-1) even in the severe self-heating region. Current dependence of $M-1$ is demonstrated. The results show that the CB breakdown voltage at the J_E of peak f_T is higher than that at either low J_E or off state by a significant 1 V in a 120 GHz peak f_T device. Also in Chapter 2, current dependence of $M-1$ is found to be considerably smaller by taking into account the extrinsic collector resistance. Later in chapter 5, a simplified model for the current dependent $M-1$ is proposed.

In chapter 3, RF characterization methods are discussed, including S-parameters, large signal power characterization and third order intermodulation. In general, S-parameters characterize small signal parameters and can be used to extract base resistance R_B, cut-off frequency f_T and maximum oscillation frequency f_{max}. The large signal system built utilizes the same equipment setup of the S-parameter system, while being able to measure device performance at large input power. With monitoring the DC voltage and current, power added efficiency can be calculated. The third order intermodulation system built is more complicated than the previous two in the sense that the system distortion level can largely affect the accuracy of the measurement result. With careful setting, accurate IIP3 measurements are done for HBTs and MOSFETs. All of the three systems are controlled by in-house programs written in VEE. The program of the

S-parameter system was written by William E. Ansley. Modification of the program is made in this work to increase the stability of operation and the application range of the program. Programs for the large signal system and the IIP3 system are written during the construction of both systems. All of the programs are now capable of measuring both HBT and MOSFET devices with high accuracy while requiring little care from operators.

Chapter 4 experimentally investigates SiGe profile and collector profile optimization from a large signal performance standpoint, as well as the impact of technology scaling. The results show that device and circuit designs that only consider optimum small signal performance could inadvertently degrade large signal performance. The tradeoffs in SiGe profile design between small signal and large signal performance, as well as the impact of speed-breakdown tradeoff on large signal performance are experimentally examined. The SiGe HBTs from a 200 GHz technology show impressive small and large signal performance at 20 GHz, demonstrating the benefits of technology scaling, despite decreased breakdown voltage.

Chapter 5 presents systematic characterization of the intermodulation linearity for SiGe HBTs to gain insight into the device physics underlying linearity behavior, and to provide guidelines to optimal sizing, biasing, and device selection (e.g. high breakdown versus low breakdown versions). The input 3rd order intercept point, IIP3, is measured on $I_C - V_{CE}$ plane for devices of various size, breakdown voltage, Ge profile, and technology generation. Later in this chapter, problems of VBIC model for simulating IIP3 are presented. Improvements for base collector capacitance and avalanche modelling in the VBIC model are suggested and implemented in Verilog-A to give a much better fitting to the measurement results.

Chapter 6 presents the major conclusions of this work.

Reliable characterization of S-parameter, large signal performance and IIP3 involves time-consuming automation programing and large effort to co-ordinate equipments. Appendix A to C give measurement guides and computer programs used in this measurements.

CHAPTER 2

AVALANCHE MULTIPLICATION

Accurate measurement of the avalanche multiplication factor ($M-1$) is critical to the design of high-speed circuits that need to operate well above the transistor open base breakdown voltage (BV_{CEO}), including power amplifiers in wireless systems, and modulator drivers in light wave communication systems [14]. Conventional $M-1$ measurement techniques, however, are only applicable to relatively low collector current densitie (J_C), for which self-heating is negligible. A fundamental approach to improving speed is to reduce the transit times through vertical scaling and increase the operating J_C. The J_C of peak f_T has increased from 1 mA/μm^2 in a 50 GHz SiGe HBT technology to 7 mA/μm^2 in a 120 GHz SiGe HBT technology [15] [16]. In order to suppress high injection effects, the collector doping must be increased, thus increasing the electric field in the collector-base (CB) junction. The combination of increasing J_C and increasing CB electric field worsens the self-heating effect ($\vec{J}\cdot\vec{E}$), making conventional $M-1$ measurement techniques invalid in newer generations of SiGe HBTs, as detailed below.

This chapter presents two new techniques for $M-1$ measurement at the higher J_C's found in newer SiGe HBTs. One of the methods is based on photo current generated during the avalanche process [17]- [19], while the other uses a pulsed IV measurement system to eliminate the effect of self-heating. The techniques are demonstrated in a 120 GHz peak f_T SiGe HBT technology.

2.1 Avalanche Physics

As bipolar devices are scaled down, the collector doping concentration must be increased in order to prevent base push out and the Kirk effect, which cause performance degradation

at high current density. The latter increases the electric field in the CB junction. In normal

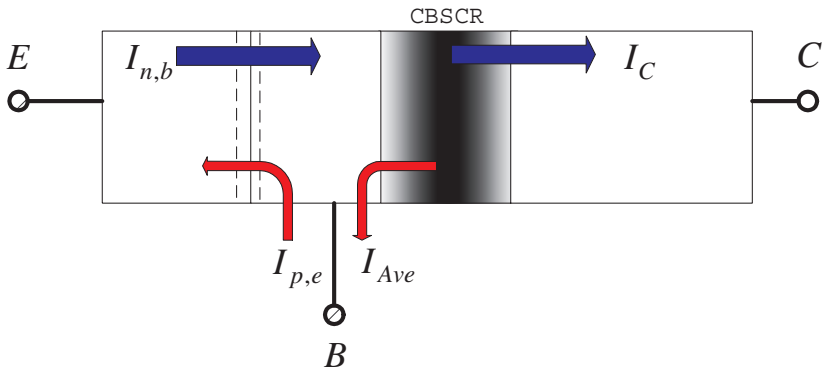


Figure 2.1: Illustration for the avalanche process in HBTs.

transistor operation, the CB junction is reverse biased. Depending on the doping profile and circuit configuration, the imposed reverse bias may cause the carriers injected from the emitter to generate electron-hole pairs via impact ionization as they traverse the CB depletion region. The CB reverse bias voltage will cause the generated electrons to be swept into the collector electrode, contributing a positive term to terminal collector current, while the generated holes will back flow into the neutral base region [20], as shown in Fig. 2.1. This carrier generation process is known as “impact ionization”. The electrons and holes generated by impact ionization can subsequently acquire energy from the strong electric field, and create additional electron-hole pairs by further impact ionization. This process of multiplicative impact ionization is known as "avalanche multiplication." The net effect is that the electron current leaving the CB space-charge region is larger than the I_C that would be observed without avalanche multiplication. The ratio of the two currents is known as the avalanche multiplication factor M

$$M = \frac{I_C}{I_{n,b}}. \tag{2.1}$$

where $I_{n,b}$ represents the electron current entering the CB space-charge region. In practice, $M-1$ is often used instead of M, which is simply:

$$M - 1 = \frac{I_C - I_{n,b}}{I_{n,b}}. \tag{2.2}$$

2.2 Conventional Measurement Method

2.2.1 Forced V_{BE} Measurement Method

The forced V_{BE} measurement setup is shown in Fig. 2.2. At $V_{CB} = 0$, I_B is dominated by the hole injection into the emitter ($I_{p,e}$). At higher V_{CB} and the same V_{BE}, without self-heating, $I_{p,e}$ should remain the same while the holes generated in the CB space-charge region flow into the base. The total base current is thus decreased by the hole current. Since the hole current equals the electron current avalanche generated, the avalanche current can be obtained by monitoring the I_B difference at different V_{CB}:

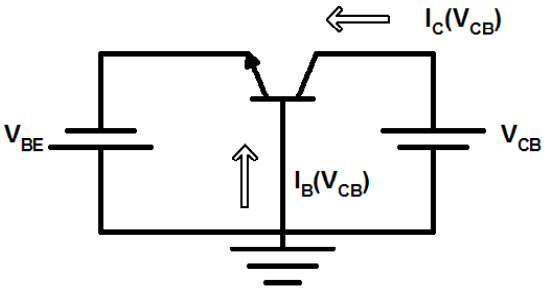
$$I_{Ave} = \Delta I_B = I_B(V_{CB} = 0) - I_B(V_{CB}) \tag{2.3}$$

The measured collector current is thus the avalanche current plus the electron current injected into the CB space-charge region. The electron current injected into the CB space-charge region is:

$$I_{n,b} = I_C(V_{CB}) - I_{Ave} \tag{2.4}$$

The $M-1$ factor can then be expressed as:

$$M-1=\frac{\Delta I_B}{I_C(V_{CB})-I_{Ave}} \tag{2.5}$$



$$M-1=\frac{\Delta I_B}{I_C\left(V_{CB}\right)-\Delta I_B}$$

$$\Delta I_B=I_B\left(V_{CB}=0\right)-I_B\left(V_{CB}\right)$$

Figure 2.2: Forced V_{BE} setup for $M-1$ measurement.

Problems with the forced V_{BE} measurement method

1. Thermal runaway

For a forced V_{BE} measurement setup, when V_{CB} increases more power is consumed by the device, heat is generated and the temperature is increased. At a fixed V_{BE}, this temperature increase means a higher current which heats up the device even further. The whole process forms a positive feedback loop which easily kills the device at high V_{BE}.

2. Self-heating

As the temperature is different at $V_{CB}=0$ and the specified V_{CB}, the difference in the base current includes not only the avalanche current but also the self-heating-generated base current

difference, whose value depends on the temperature difference. The accuracy of the avalanche current estimated by this method is thus largely significantly compromised.

2.2.2 Forced I_E Measurement Method

As shown in Fig. 2.3, to measure $M-1$, the collector-base voltage (V_{CB}) is swept at a fixed emitter current (I_E), and the emitter-base voltage (V_{BE}) is recorded. For modern transistors, hole injection into the emitter is far more significant than neutral base recombination, and therefore we have:

$$I_{n,b} = I_E - I_B(V_{BE})|_{V_{CB}=0} \tag{2.6}$$

The $M-1$ factor can then be expressed as [21]:

$$M-1 = \frac{I_C}{I_E - I_B(V_{BE})|_{V_{CB}=0}} - 1 \tag{2.7}$$

where V_{BE} is recorded during the V_{CB} sweep, and $I_B(V_{BE})|_{V_{CB}=0}$ is found from the $I_B - V_{BE}$ curve obtained with $V_{CB} = 0$ V.

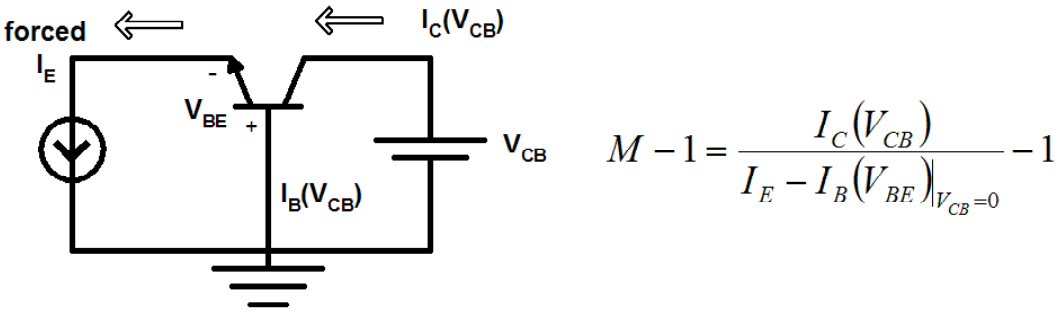


Figure 2.3: Forced I_E setup for $M-1$ measurement.

The forced-I_E method makes the measurement of $M-1$ safer because the total amount of current injected into the CB space-charge region is always limited by I_E. The feedback mechanism for avalanche multiplication is thus effectively limited in the presence of self-heating. This is a significant improvement over the forced-V_{BE} measurement method. The problem of the poor accuracy due to self-heating, again, still exists in the forced-I_E method. $I_B(V_{BE})|_{V_{CB}=0}$ is measured at $V_{CB} = 0$ V, where the device operates at a lower temperature than it does at higher V_{CB} values.

2.3 Photo Current Based Measurement Method

2.3.1 Measurement Technique

The proposed technique is based on photo carrier generation by hot carrier induced light which produces electron-hole pairs in the collector-substrate junction, as shown in Fig. 2.4 (a). Collection of these electron-hole pairs leads to a substrate current (I_{Sub}), which can be used to monitor the occurrence of avalanche multiplication [19].

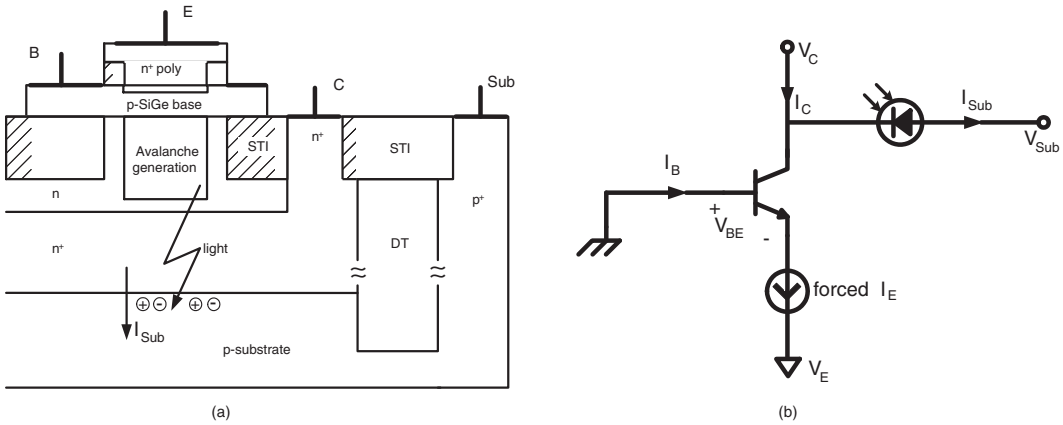


Figure 2.4: (a) Illustration of photo carrier generation in the SiGe HBTs used. (b) Experimental setup of the substrate current based $M-1$ measurement technique.

Fig. 2.4 (b) shows a schematic of the measurement setup. The base is grounded. The collector voltage V_C is set to the desired V_{CB}. The substrate voltage is chosen such that the collector-substrate bias $V_{CS} \geq 0$. An emitter current I_E is forced, and the value of I_E is swept. V_{BE}, I_B, I_C and I_{Sub} are recorded during the I_E sweep. In the absence of self-heating, the avalanche current can be obtained as the difference in I_B between high V_{CB} and $V_{CB} = 0$ V, denoted as I'_{Ave} [37]:

$$I'_{Ave} = I_B(V_{BE}, V_{CB} = 0) - I_B(V_{BE}, V_{CB}), \tag{2.8}$$

where $I_B(V_{BE}, V_{CB} = 0)$ is the I_B at the V_{BE} values recorded during the I_E sweep for a 0 V V_{CB}, and can be determined using a separate measurement. $I_B(V_{BE}, V_{CB} = 0)$ represents the hole current injected into the emitter. At high I_E, self-heating becomes severe, and the junction temperature increases with V_{CB} significantly. Therefore, $I_B(V_{BE}, V_{CB} = 0)$ gives the hole current injected into the emitter at a junction temperature lower than at the desired V_{CB}. The hole current injected into the emitter at the desired V_{CB} is thus *underestimated* by $I_B(V_{BE}, V_{CB} = 0)$. Consequently, the avalanche current is underestimated by I'_{Ave}. Negative I'_{Ave} can be obtained, which is clearly unphysical for avalanche current.

Fig. 2.5 (a) shows the measured I'_{Ave}, I_{Sub} and the I_{Sub}/I'_{Ave} ratio as a function of I_E. I'_{Ave} and I_{Sub} are shown on the left axis, and I_{Sub}/I'_{Ave} is shown on the right axis. V_{CB}=1.2 V. The SiGe HBT used has a 120 GHz peak f_T [16], and an emitter area $A_E = 0.2 \times 6.4$ μm^2. I'_{Ave} first increases with I_E, as expected, but becomes negative at an I_E of 1.8 mA, well below the peak f_T I_E (9 mA), because of self-heating. However, for medium I_E (region B), I_{Sub} increases proportionally with I'_{Ave}, and a constant I_{Sub}/I'_{Ave} ratio can be identified. Intuitively, this ratio can be viewed as the efficiency of substrate current generation due to avalanche, which we denote as η. Measurements show that η is independent of V_{CB} and V_{CS}, which accords with [19]. Note

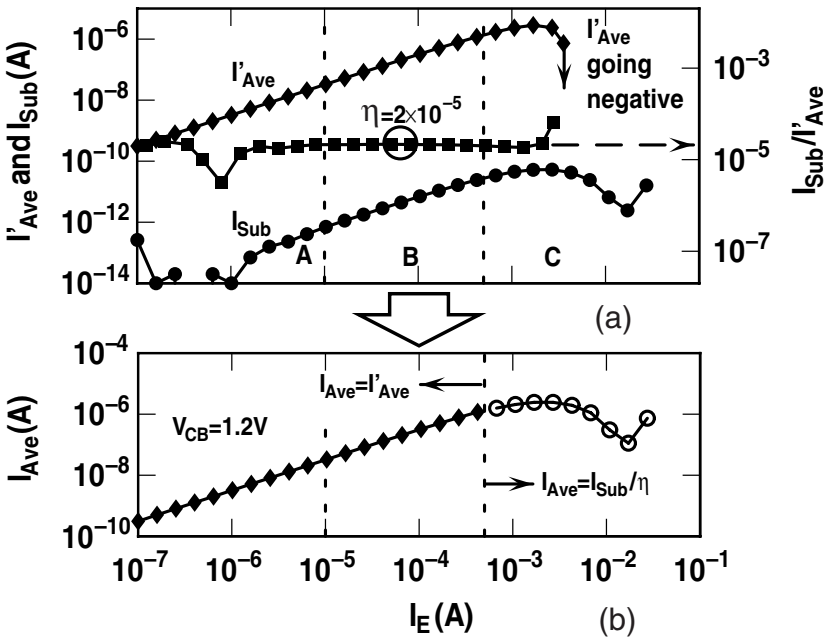


Figure 2.5: (a) I'_{Ave}, I_{Sub} and the I_{Sub}/I'_{Ave} ratio versus I_E. The right y-axis shows the I_{Sub}/I'_{Ave} ratio. V_{CB} = 1.2 V. A_E = 0.2 × 6.4 μm^2. The substrate current generation efficiency η is determined as the constant I_{Sub}/I'_{Ave} ratio identified in region B (10^{-5}A/μm^2 < I_E < 5 × 10^{-4}A/μm^2). (b) Avalanche current I_{Ave} versus I_E.

that η cannot be measured using the I_{Sub}/I'_{Ave} ratio in region C, because I'_{Ave} gives incorrect avalanche current. Caution, however, must be exercised in interpreting the I_{Sub} at very high I_E. A rapid increase of I_{Sub} is observed at very high I_E, due to hole injection resulting from the forward biasing of the internal CB junction. This is confirmed by MEDICI simulation of the HBTs measured. This, however, occurs at I_E values well above the I_E of peak f_T, and does not present a problem for practical application of the proposed method.

Assuming that the substrate current generation efficiency is independent of self-heating, the avalanche current in the high I_E region (region C) can be extracted from I_{Sub} as I_{Sub}/η. The

avalanche current becomes:

$$I_{Ave} = \begin{cases} I'_{Ave} & \text{region A and B,} \\ I_{Sub}/\eta & \text{region C.} \end{cases} \tag{2.9}$$

The I_{Ave} calculated using (2.9) is shown in Fig. 2.5 (b). Strictly speaking, the substrate current generation efficiency η is also a function of temperature, and hence a function of I_E in region C. To account for the temperature dependence of η, η is measured at various ambient temperatures in the medium I_E range, where self-heating is negligible. The measured η decreases from 2.1×10^{-5} to 1.5×10^{-5} in a linear fashion as temperature increases from 295K to 360K. For each I_E in region C, the junction temperature T_J is calculated as:

$$T_J = T_{amb} + P \times R_{th}, \tag{2.10}$$

where T_{amb} is ambient temperature, $P = I_B V_{BE} + I_C V_{CE}$ is power consumption, and R_{th} is thermal resistance. The η in (2.9) is now dependent on I_E in region C through T_J. R_{th} is determined using the method of [22]. By definition, $M - 1$ is obtained as:

$$M - 1 = \frac{I_{Ave}}{I_C - I_{Ave}}. \tag{2.11}$$

2.3.2 Results and Discussion

Fig. 2.6 shows $M-1$ versus J_E obtained using the conventional technique [37] and using the proposed technique. The corresponding I_{Sub} versus J_E characteristics is shown in Fig. 2.7. The results with and without accounting for the temperature dependence of η are both shown. The cut-off frequency f_T is measured and shown on the right y-axis. Observe that the conventional

technique gives negative and thus unphysical $M-1$ at J_E well below the J_E of peak f_T. The proposed technique works over a much wider J_E range that includes the peak f_T point. The upper applicable limit is set by hole injection from the collector, which occurs at J_E well above the peak f_T. As expected, the temperature dependence of η yields a correction that increases with J_E, which is 30% at peak f_T.

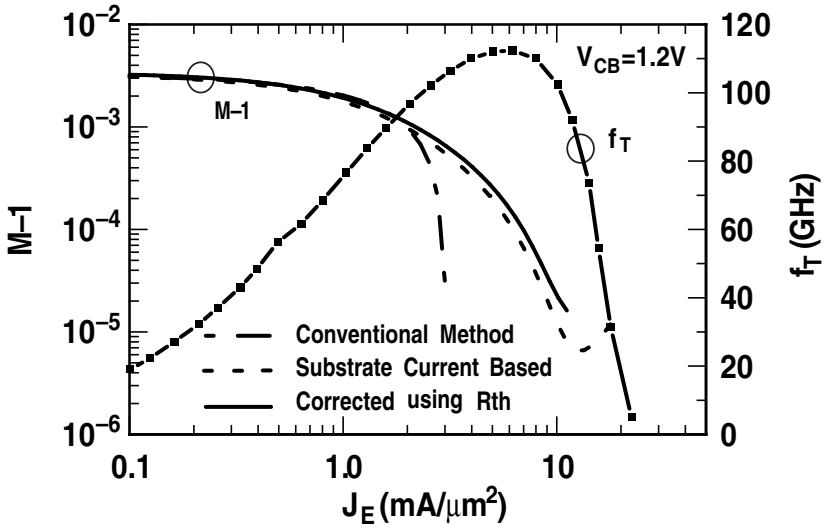


Figure 2.6: $M-1$ versus J_E obtained using conventional technique and substrate current based technique with and without considering temperature dependence of η. f_T is shown on right axis. V_{CB} = 1.2 V. $A_E = 0.2 \times 6.4\ \mu\text{m}^2$.

Note that the decrease of $M-1$ starts at $J_E = 0.1\ \text{mA}/\mu\text{m}^2$, which is much lower than the J_E of peak f_T, 7 $\text{mA}/\mu\text{m}^2$. $M-1$ decreases from 2×10^{-3} at $J_E = 0.1\text{mA}/\mu\text{m}^2$ to 10^{-4} at the J_E of peak f_T. Physically this is reasonable considering that f_T rolls off only when J_E is sufficient to cause base push out, while $M-1$ decreases as long as the J_E is sufficient to cause a decrease of the CB junction peak field. As a result, I_{Sub} reaches its peak at a J_E of 1.5 $\text{mA}/\mu\text{m}^2$, which is well below the J_E of peak f_T (7 $\text{mA}/\mu\text{m}^2$), as shown in Fig. 2.7.

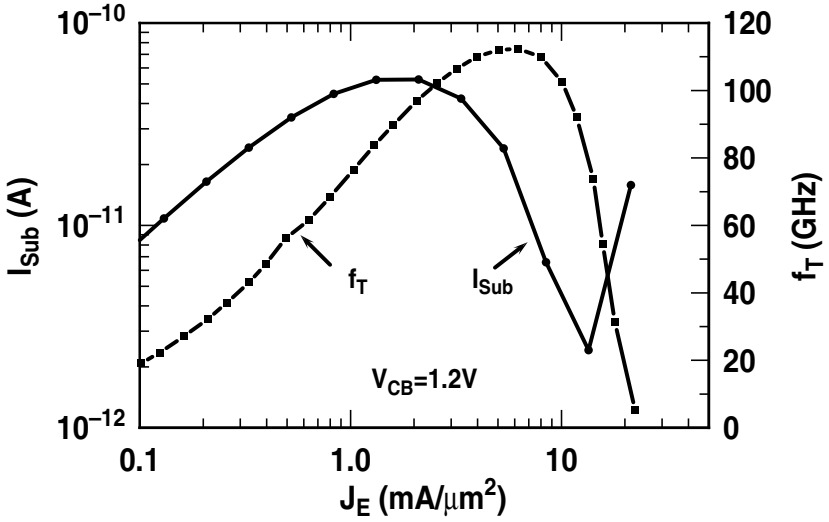


Figure 2.7: Substrate current I_{Sub} and f_T versus J_E.

For transistors used in RF power amplifiers, the maximum voltage handling capability depends on the details of $M-1$ vs V_{CB} characteristics. These applications require high J_E biasing for high speed, and high power density. It is therefore important to understand the $M-1$ vs V_{CB} characteristics at high biasing J_E. The breakdown voltage at high I_C is also an important concern for operating with mismatched load. Fig. 2.8 shows the measured $M-1$ versus V_{CB} at $J_E = 6.875$ mA/μm^2, at which $f_T = 110$ GHz. The measured data at $J_E = 0.1$ and 1 mA/μm^2 are also shown for comparison. The $M-1$ at $J_E = 6.875$ mA/μm^2 is 20 times smaller than the $M-1$ at $J_E = 1$ mA/μm^2. Operating the transistor at high J_E for high f_T effectively increases the breakdown voltage by 1 V, as can be seen from the lateral shift of the $M-1$ curves. A 1 V increase of the breakdown voltage is significant for these 120 GHz SiGe HBTs.

In an $M-1$ measurement, the applied collector base voltage is usually used. However, the avalanche is actually a function of the intrinsic voltage on the CB space-charge region rather

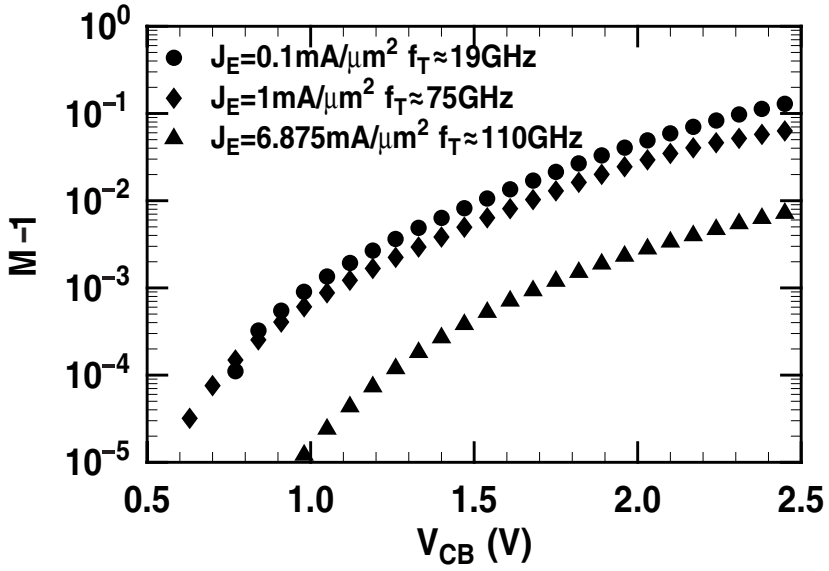


Figure 2.8: $M-1$ versus V_{CB} for $J_E = 6.875$, 1.0, and 0.1 mA/μm^2. $A_E = 0.2 \times 6.4$ μm^2.

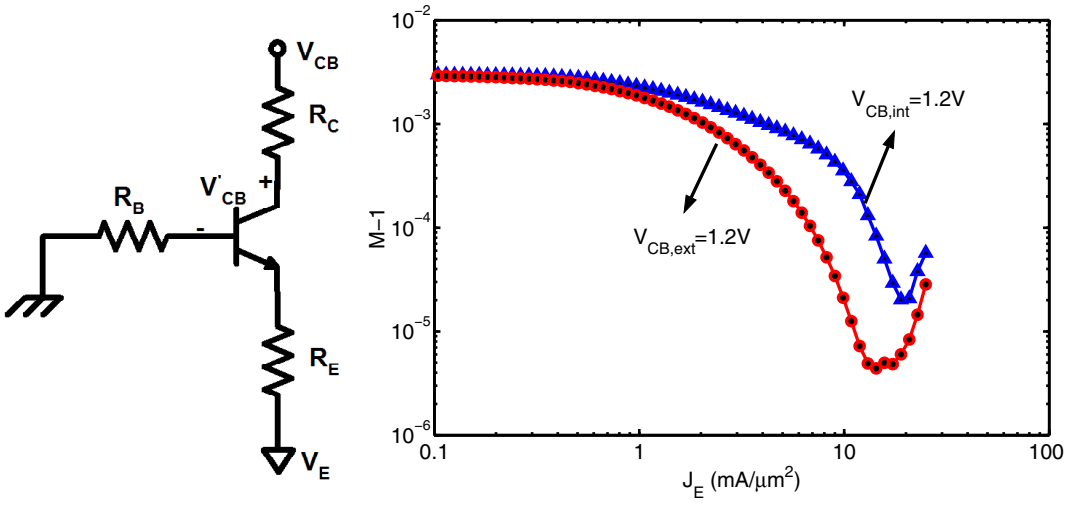
than the applied voltage. As shown in Fig. 2.9 (a), the intrinsic collector base voltage should be:

$$V'_{CB} = V_{CB} - I_C \cdot R_C + I_B \cdot R_B \tag{2.12}$$

Where V'_{CB} is the intrinsic collector base voltage and V_{CB} is the applied CB voltage. Usually $I_C \cdot R_C >> I_B \cdot R_B$,

$$V'_{CB} = V_{CB} - I_C \times R_C \tag{2.13}$$

Fig. 2.9 (b) shows the measured $M-1$ versus J_E for a fixed extrinsic V_{CB} =1.2 V and a fixed intrinsic V_{CB} =1.2 V. The difference shows clearly that the impact of current dependence is smaller after considering the voltage drop at the collector resistor.



(a) (b)

Figure 2.9: Illustration of a transistor with lumped extrinsic resistors and the $M-1$ corresponding to intrinsic V_{CB} and extrinsic V_{CB}.

2.4 Pulsed I-V Measurement Method

2.4.1 Measurement Technique

As self-heating is a slow process, which has a time constant on the order of microseconds, it is possible to avoid significant device temperature changes by using RF pulse I-V measurements. A quiescent biasing condition is perturbed with short, low duty rate (as short as 10 ns) pulsed stimuli, during which the change in terminal voltage and current is recorded. The pulse duration is short enough to prevent relatively slow processes such as self-heating and slow electron/hole trapping from occuring, but not too short to allow the junction capacitances to complete the charging/discharging process. The separation between pulses is long enough to allow the quiescent condition to recover from any perturbation that may occur during the short pulse. The I-V

measured using the pulse stimuli represents transistor RF behavior at the quiescent operating point (or biasing point).

Fig. 2.10 depicts the trajectory of the voltages and currents for a pulsed I-V measurement made on a SiGe HBT with 50 GHz peak f_T [23], including the biasing point. The Dynamic I-V Analyzer (DIVA) from Accent Optical Technologies was used for these measurements [24], with a pulse width of 100 ns. The separation between pulses was 1 ms. Note that the I-V curves measured using RF pulses depend on the quiescent biasing point chosen, and the device is at a constant temperature determined by the power consumption at the biasing point. The key difference from a conventional DC I-V (e.g. measurements taken using an HP 4155) is that none of the time dependent phenomena have a chance to reach the steady state, including self-heating. An dynamic I-V is more relevant than DC I-V for devices controlled by fast moving voltages, e.g. transistors in RF amplifiers [25].

In order to measure $M-1$, we must first choose a biasing point of interest, $V_{BE,bias}$ and $V_{CE,bias}$. V_{BE} and V_{CE} RF pulses are applied simultaneously, and the currents are recorded. The common-base biasing configuration used in conventional $M-1$ measurement is not used here, primarily because of the RF test signals involved. The S-parameter test structures must be used to apply RF signals to the transistor, where the emitter is typically the common ground. As the device temperature does not change for each pulse cycle, the base current difference between V_{CB}= 0 V and the V_{CB} of interest gives the avalanche current. $M-1$ is readily obtained as [26]:

$$I_{Ave} = I_B(V_{BE}, V_{CB} = 0) - I_B(V_{BE}, V_{CB}) \tag{2.14}$$

$$M - 1 = \frac{I_{Ave}}{I_C - I_{Ave}}. \tag{2.15}$$

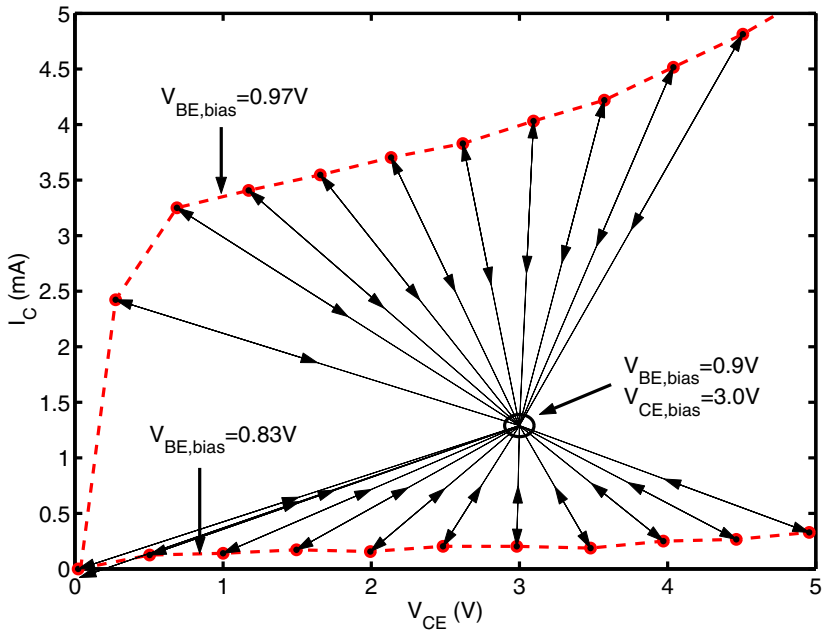


Figure 2.10: The trajectory of the currents and voltages during a RF pulsed I-V measurement. The device is a SiGe HBT with 50 GHz peak f_T.

2.4.2 Results and Discussion

The utility of the proposed method is demonstrated using SiGe HBTs with 50 GHz and 120 GHz peak f_T [15] [16]. For the 120 GHz HBT, a current density of 7 mA/μm^2 is necessary to reach peak f_T.

Operating Current and Voltage Dependence

Fig. 2.11 shows a plot of $M-1$ versus V_{CB} measured at V_{BE} =0.96 V using both DC I-V (HP 4155) and the RF pulse I-V (DIVA). In the pulse I-V measurement, $V_{CE,bias}$ =1.5 V and $V_{BE,bias}$=0.9 V. As discussed above, the conventional DC I-V result is inaccurate because of the device temperature change with V_{CB}. Observe that the collector current density J_C is

approximately 1.76 mA/μm^2 at V_{BE}=0.96 V in the pulse I-V measurement, and 2.04 mA/μm^2 at V_{BE}= 0.96V in the DC I-V measurement. The current density dependence of $M-1$ must be examined if the error made using the conventional method needs to be quantified for the same biasing current density.

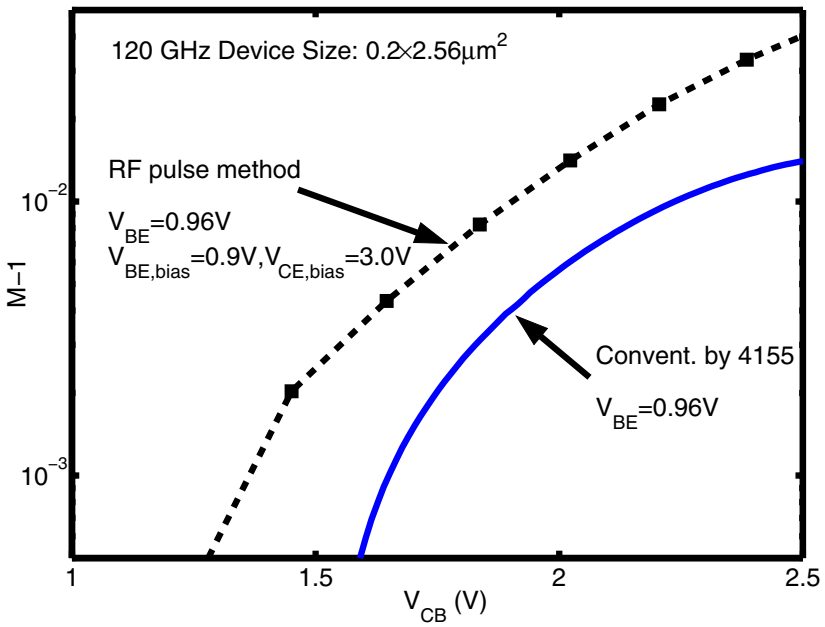


Figure 2.11: $M-1$ versus V_{CB} measured at $V_{BE}=0.96$V using both the conventional technique and the RF technique.

Fig. 2.12 shows the $M-1$ vs V_{CB} measured at pulse V_{BE} of 0.84, 1.0 and and 1.04 V. $V_{BE,bias}$ = 0.9 V and $V_{CE,bias}$ = 1.5 V. The collector current densities are approximately 0.5, 6 and 9 mA/μm^2, covering the practical operating region of this 120 GHz HBT. For the two larger V_{BE}, the conventional method fails to complete the measurement due to thermal runaway.

Fig. 2.13 shows $M-1$ versus J_E measured at V_{CB} =1.8 V using RF pulse I-V and conventional DC I-V, together with f_T shown on the right y-axis. Note that the conventional method gives inaccurate $M-1$ in the peak f_T region while the RF pulse method works well. The peak

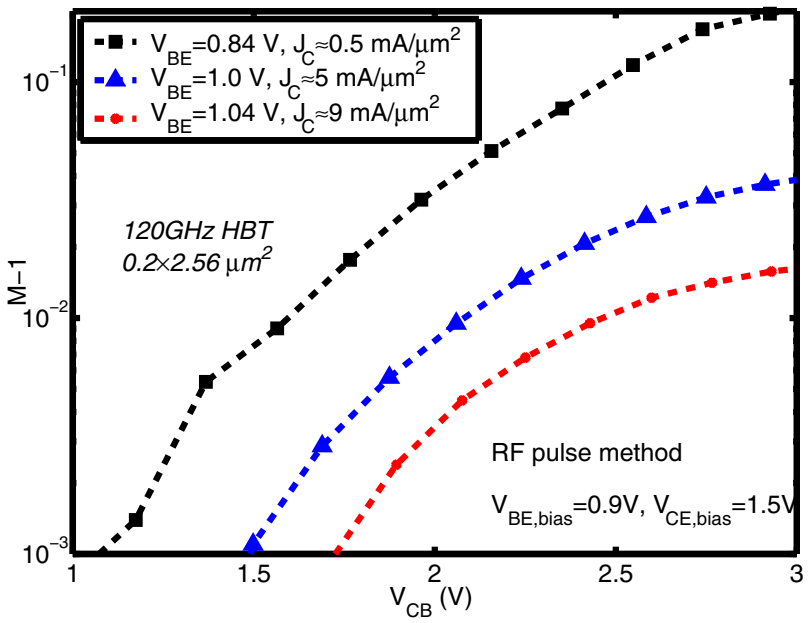


Figure 2.12: $M-1$ versus V_{CB} measured at pulse VBE =0.84 V, 1.0 V and 1.04 V using RF pulse I-V.

f_T region is important for practical circuit application, and the $M-1$ in this region can only be characterized using RF pulse I-V.

We mentioned that pulse I-V measurement allows measurement at high current and high voltage, where conventional method fails. However, if one is interested in measuring $M-1$ at low current density, the conventional DC I-V method should be used, because the resolution of RF pulsed I-V systems is not as good (resolution for fixed V_{BE} pulse is on the order of mV).

Biasing Current and Voltage Dependence

The RF pulse I-V characteristics depend on the DC biasing condition. It is thus interesting to examine the dependence of $M-1$ from pulse I-V measurement on the biasing condition.

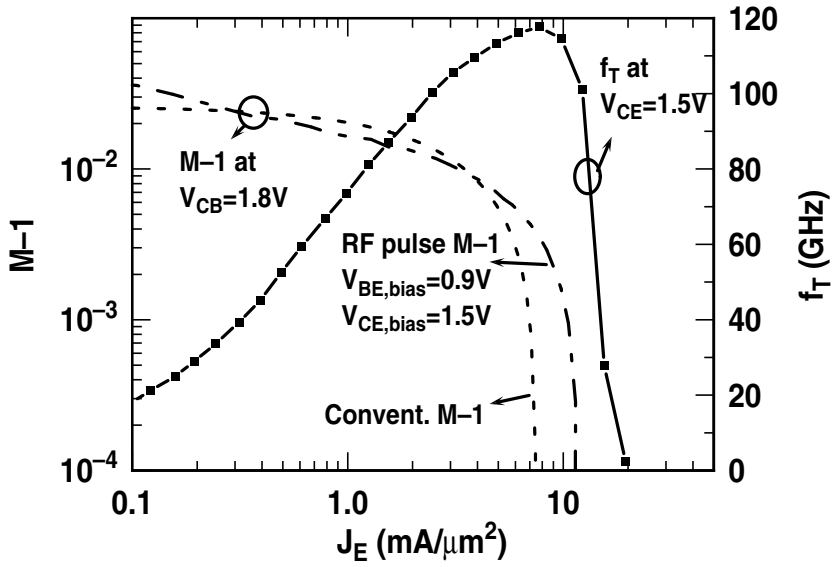


Figure 2.13: $M-1$ versus J_E measured at V_{CB} = 1.8 V using RF I-V and conventional DC I-V with f_T shown on the right y-axis.

Fig. 2.15 shows $M-1$ versus V_{CB} at pulse V_{BE} = 0.83 and 0.95 V. For each pulse V_{BE}, two biasing conditions are used: $V_{BE,bias}$ =0.9 V and $V_{CE,bias}$=0.5 V, at which the device temperature is 300 K, and $V_{BE,bias}$=0.93 V and $V_{CE,bias}$=1.5 V, at which the device temperature is 310 K. The $M-1$ is mainly a function of the operating current and voltage, and does not vary strongly with the biasing current and voltage. A slight increase of $M-1$ with increasing biasing power consumption (device temperature) is observed at lower operating current, which could be associated with the temperature dependence of the impact ionization coefficients of electrons. At higher operating current, however, the $M-1$ for the two biasing points is approximately the same.

Fig. 2.15 shows M-1 versus operating current density for a 50 GHz peak f_T HBT and a 120 GHz peak f_T HBT. The V_{CB} is 1.8 V and 3.5 V for the 120 and 50 GHz peak f_T HBTs. The results confirm that at higher operating current density M-1 is mainly a function of the operating

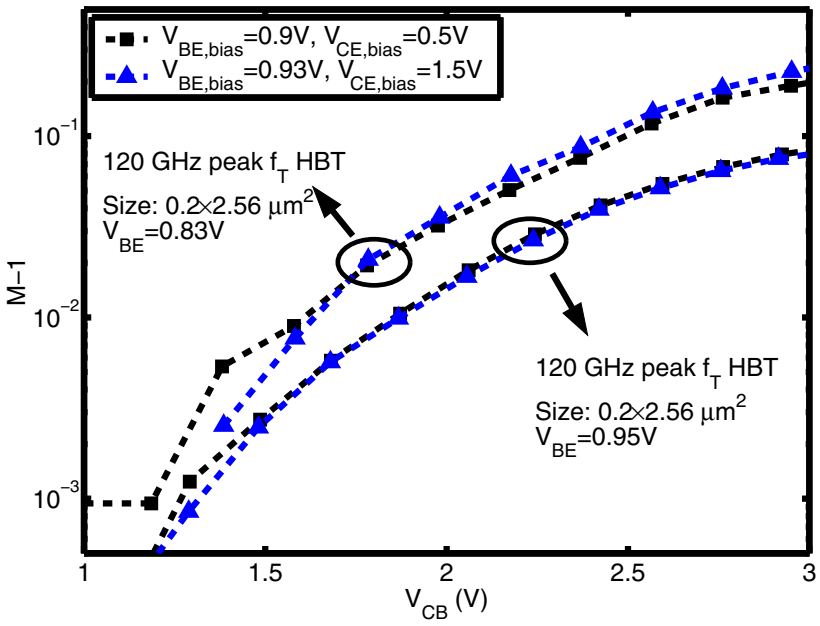


Figure 2.14: $M-1$ versus V_{CB} measured at pulse V_{BE} = 0.83 V and 0.95 V using RF I-V with two bias conditions.

current density and voltage (V_{CB}), while at lower current density, M-1 increases slightly with device temperature.

R_C effect

Applying the collector resistance to $M-1$ versus V_{CB} results, as shown in Fig. 2.16, a similar R_C effect is observed as that seen for the substrate current based method. $M-1$'s change versus V_{be} (I_C) at the same applied V_{CB} is much larger than $M-1$'s change at the same intrinsic V'_{CB}. This can be seen more clearly in Fig. 2.17. Fig. 2.17 shows $M-1$ versus collector current density at an applied V_{CB} of 1.8 V and an intrinsic V'_{CB} of 1.8 V. Clearly, the drop in $M-1$ is much faster in the applied V_{CB} curve. This implies the collector current dependence of

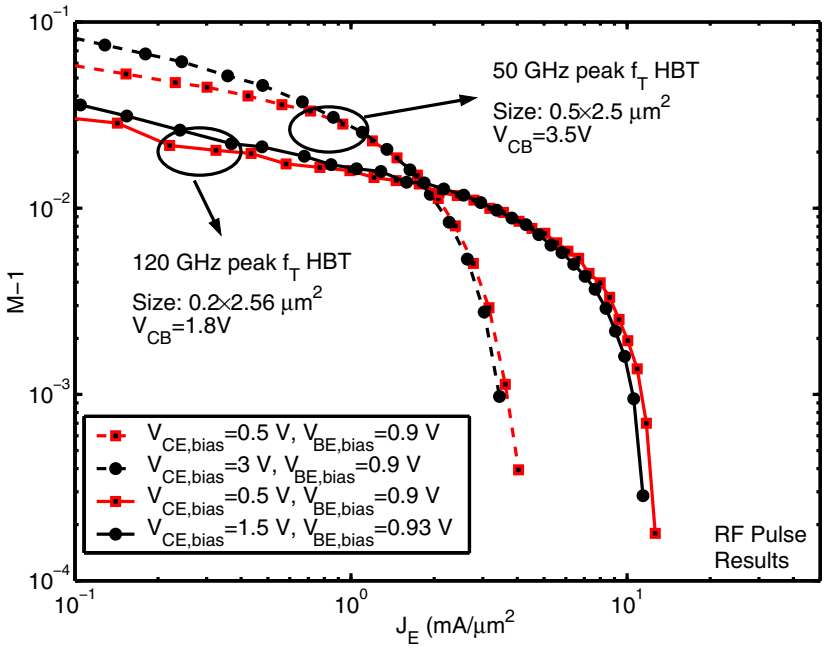


Figure 2.15: $M-1$ versus operating current density J_E for a 50 GHz peak f_T HBT and a 120 GHz peak f_T HBT.

$M-1$ is actually much lower when considering the collector resistance. $M-1$ modeling should thus include the collector resistance of the transistor.

2.5 Comparison of the Two New Methods

The two methods proposed in this chapter are both capable of measuring $M-1$ at relatively high current density at which the convential methods fail. However, due to their inherent properties, the two methods have their own preferred application range.

For the substrate current based method, the measurement can only be done when there is a substrate which can be used to collect the photo current. This substrate, however, is not always present. Secondly, although the measured $M-1$ of the substrate current method is the true

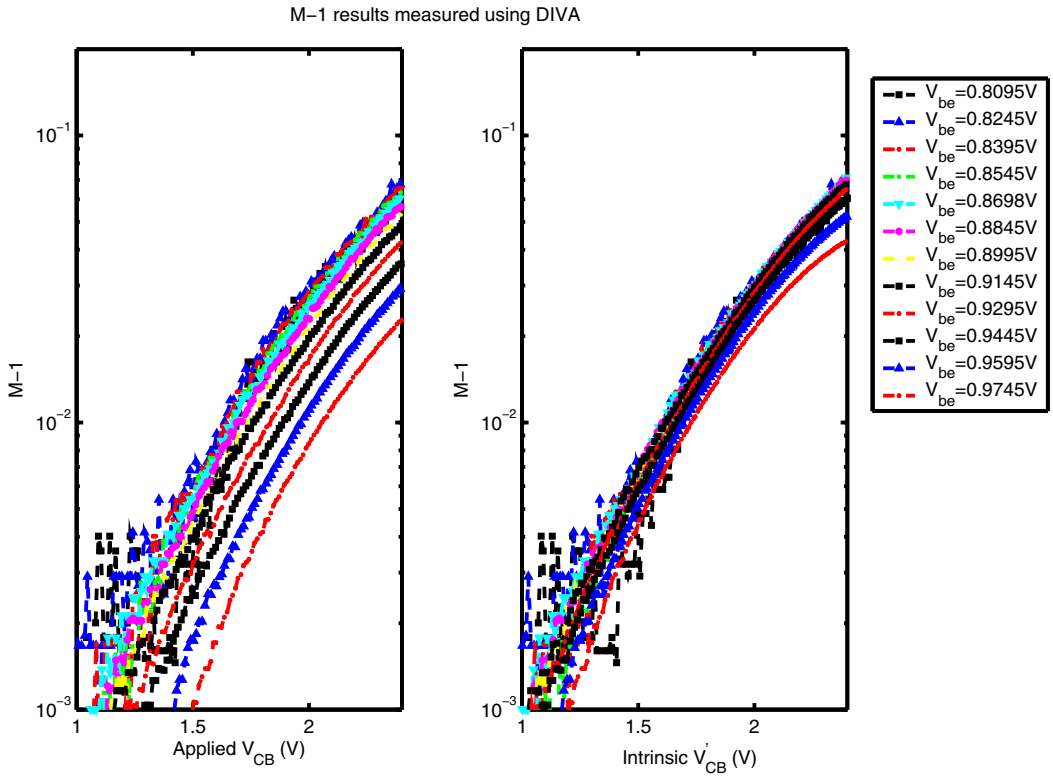


Figure 2.16: $M-1$ versus applied V_{CB} and intrinsic V'_{CB}.

$M-1$ of the device, the $M-1$ measured is at different temperatures as self-heating is inevitable during the measurement. Thus it is not the true case under small signal RF condition. Thirdly, the method assumes that the photo current dominates the collector substrate current, which is not always the case. In some devices, there may be a path between the emitter and substrate, mostly for ESD protection. This path may generate a comparable substrate current to that obtained form the photo current, leading to a erroneous photo current generation efficiency.

The pulsed-IV method does not suffer from the first two problems of the substrate method. The pulsed-IV method only needs three terminals for HBTs; access to a substrate terminal is not necessary. The pulsed-IV method also maintains the device at the same temperature because the

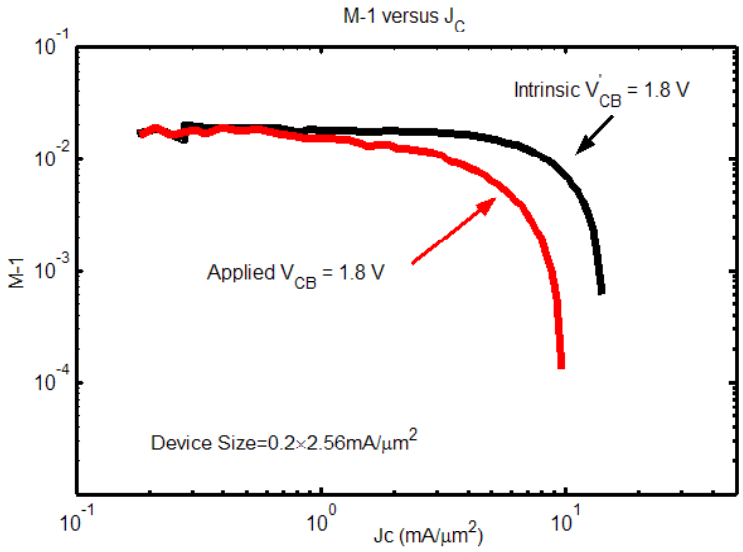


Figure 2.17: $M-1$ versus applied V_{CB} and intrinsic V'_{CB}.

device is always held at the bias condition before and after each pulse. This is the same condition as that for the small signal RF operation. The downside of this method is that the accuracy of the equipment is currently not very high. At low avalanche or low current, the measurement result is too noisy.

2.6 Summary

In this chapter, we reported the development of two new techniques that allow accurate $M-1$ measurement at the high operating current densities required for high speed, where conventional methods fail because of severe self-heating. The utility of the substrate current method was demonstrated on SiGe HBTs featuring 120 GHz peak f_T. The results show that the CB breakdown voltage at the J_E of peak f_T is higher than that at either low J_E or off state by a significant 1 V. We also presented RF pulse I-V based M-1 measurement at the high operating

current densities typically required for high speed operation of SiGe HBTs. Comparisons with conventional DC I-V based measurement were given and the dependences of M-1 on operating current and voltage as well as biasing current and voltage were examined. The impact of the collector resistance was examined for both methods. The collector current dependence of $M-1$ was found to be much smaller when the collector resistance was taken into account.

CHAPTER 3

ON-WAFER RF CHARACTERIZATION

RF characterization of SiGe HBTs is generally performed by on-wafer measurements. For on-wafer level measurement, special calibration procedures, as well as basic RF measurement procedures, must be followed in order to ensure the accuracy of the results. This chapter discusses the measurement process for on-wafer S-parameters characterization, large signal power characterization and third order intermodulation characterization. In general, S-parameters characterize small signal parameters and can be used to extract the base resistance R_B, the cut-off frequency f_T and the maximum oscillation frequency f_{max}. The large signal power measurement is able to provide the first order output power versus input power, with monitoring the DC voltage and current, the power added efficiency can be calculated. The third order intermodulation is one of the figure of merits for device linearity. The third order intermodulation characterization presented in this chapter will focus on the input 3rd order intercept point, IIP3. The equipments necessary for all three types of characterization will be discussed as well as special settings needed to improve the accuracy of the measurements.

3.1 S-parameter Measurement

3.1.1 Measurement Setup

The S-parameter measurement system consists of an HP8510C Vector Network Analyzer, an HP8517B S-parameter Test Set (45MHz-50GHz), an HP83651A Synthesized Sweeper (45MHz-50GHz), an Alessi REL-4300 microwave probing station, a pair of Infinity I40A GSG 150 microwave probes made by Cascade Microtech (dc-40GHz), and an HP6626A system DC power

supply. This system can be used to characterize any n-port network having a maximum output power of +17dBm (50mW) over the measurement frequency range. The HP83651A Synthesized Sweeper is capable of sourcing +10dBm, but the actual power at the ports is attenuated as a result of the system losses and is a function of frequency. Functionally, the HP83651A Synthesized Sweeper provides the RF source signals, while the HP8517B S-parameter Test Set separates the RF source signals into reference and test signals and down-converts the reference and test signals into separate 20MHz IF (intermediate frequency) signals. The 20MHz signals are further down-converted to 100kHz signals, which are then amplified, digitized, processed and displayed. All of these steps are performed by the HP 8510C Vector Network Analyzer. The HP 8510C Vector Network Analyzer controls the entire system via a dedicated GPIB (General Purpose Interface Bus, or IEEE-488) parallel interface bus, with a second GPIB connector which allows external control of the entire system. The HP8517B S-parameter Test Set is equipped with two coaxial test ports where the device under test (DUT) can be connected. If the DUT has more than two ports, the ports not connected to the test set should be terminated at the characteristic impedance Z_0 of the measurement system (Z_0=50Ω). [27]

3.1.2 Calibration Processes

In this work, S-parameter results were used to extract the small signal parameters for SiGe HBTs. Because of the nonlinearity of the HBTs, the input signal power level should be very small (-50 dBm in this work) and equal at different frequencies. The cable attenuation, however, is not constant over the range of frequencies. An HP437B Microwave Power Meter and HP 8487A Power Sensor were used to measure the actual power at the ends of the cables, and the 8510C built in function Power Flatness Correction is used adjust the actual power level to be the same for the frequencies interested. The calibration schematic is shown in Fig. 3.1.

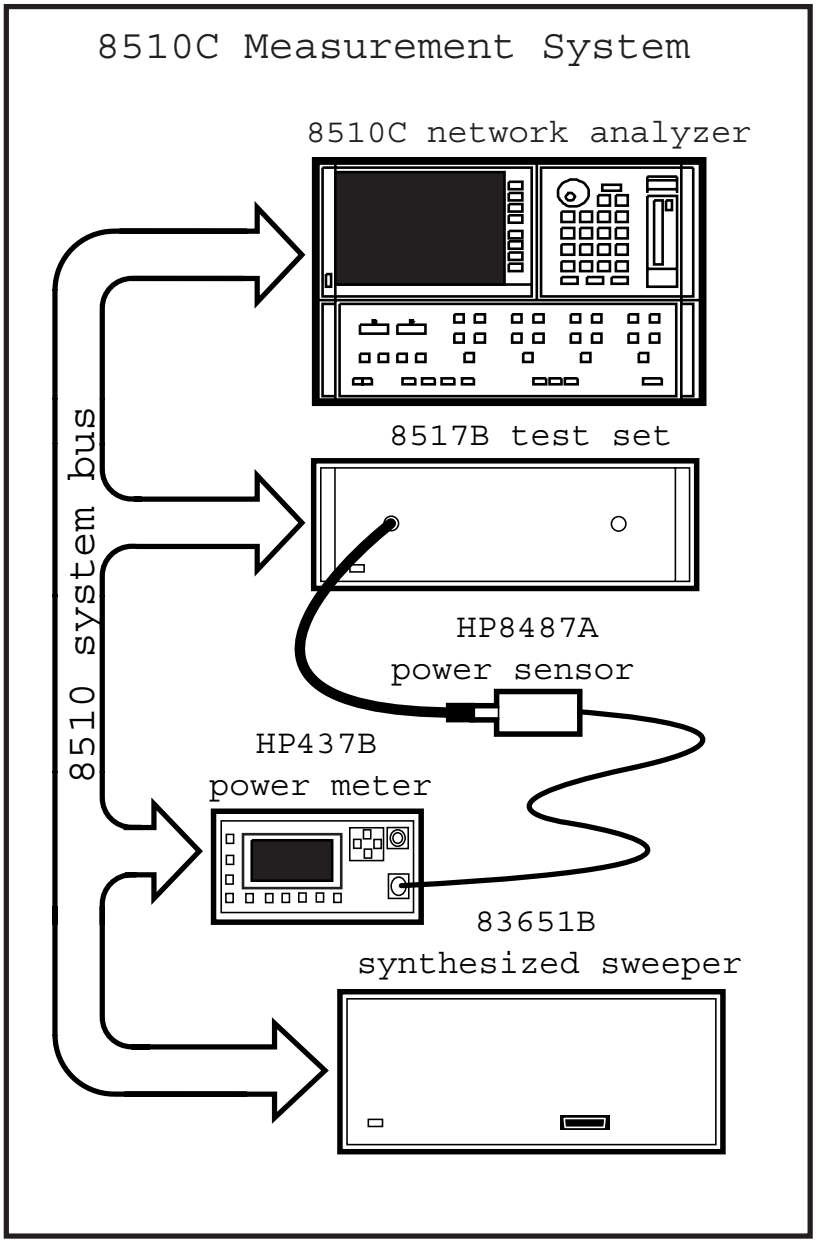


Figure 3.1: Power flatness calibration setup. [27]

When conducting RF measurement, the defectiveness of cables and probes is severe enough to affect the validity of the results. The parasitics of cables and probes will be added to the true response from the Device Under Test (DUT). Besides the errors from cables and probes, VNA itself generates systematic errors from imperfections within the internal fixtures. These parasitics and systematic errors can be quantified by measuring characteristics of known devices (standards) and then removed from the measurement results of DUT.

Fig. 3.2 shows a typical one-port VNA error model for reflection coefficient measurement [28]. The term S_{11M} represents the reflection coefficient measured by the receiver within the VNA. E_{DF}, E_{RF}, and E_{SF} are error model coefficients. The term E_{DF} accounts for directivity error, which causes the measured reflected signal fail to collect entire reflections caused by the DUT. The term E_{SF} is the error caused by source mismatching. The term E_{RF} describes the frequency tracking imperfections between reference and test channels. Through the error models, the relation between S_{11} of the DUT and S_{11M} becomes:

$$S_{11} = \frac{S_{11M} - E_{DF}}{E_{SF}\,(S_{11M} - E_{DF}) + E_{RF}} \tag{3.1}$$

If S_{11} and S_{11M} are both known, this equation becomes an equation with three unknowns: E_{DF}, E_{SF}, and E_{RF}. Since S_{11M} is the result measured by the VNA, it is always known. Obviously, S_{11} must be provided to quantify the error models, which is done by measuring standards with known S parameters.

Similarly, two-port VNA error models are shown in Fig. 3.3. Here, the forward and reverse error models combined together are able to provide the relationship between the internal S parameters (from the DUT) with the external S parameters (from the VNA) [29] [30].

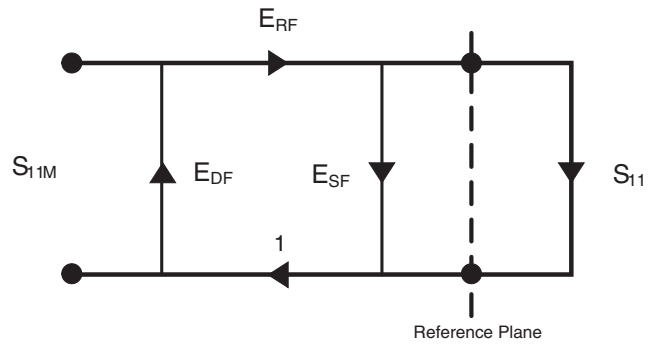


Figure 3.2: Typical one-port VNA error model for reflection coefficient measurement. [32]

The error models can be quantified in the same way as for the single port VNA. Several standards are needed to determine the error coefficients. In this work, open, short, load, and through (OSLT) standards were applied. Ideal OSLT standards are lossless and have no electrical length. Fig. 3.4 shows the electrical definitions for a set of ideal OSLT. Using ideal standards, the error coefficients can be easily calculated. Since it is impossible to fabricate the ideal standards, the S parameters of practical standards are provided by the manufacturer in a so called "Cal Kit". By loading the "Cal Kit" into the VNA and conducting the corresponding measurements, the error coefficients can be calculated automatically by the VNA and stored in its internal memory.

Other calibration methods are also available, for example, TRL (Through, Reflect, Line) and LRM (Line, Reflect, Match) methods. Detailed explanation of these methods are can be found in [31] and [32].

After the power flatness calibration and the OSLT calibration are implemented, the power at the probe tips are almost the same for all the frequencies and the measurement reference plane is moved from the ports of the VNA to the probe tips; the VNA is now ready to perform measurements on the SiGe HBTs.

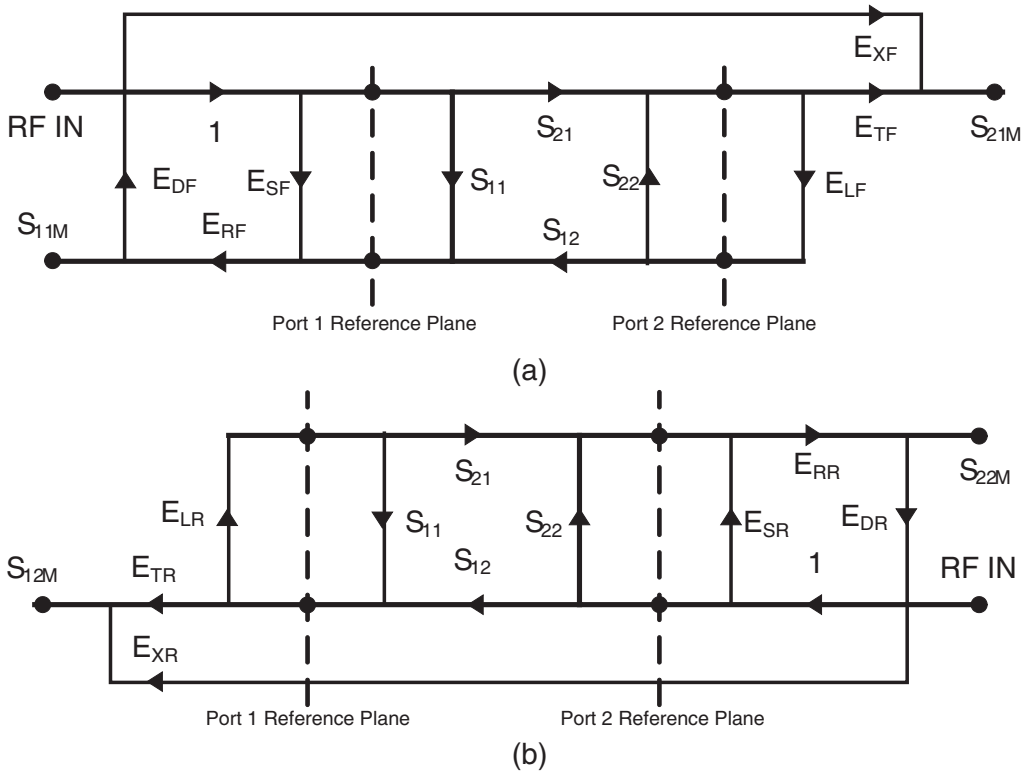


Figure 3.3: (a) Forward model for typical two-port VNA error model (b) reverse model. [32]

3.1.3 De-embedding

The size of these SiGe HBTs is typically on the order of microns, which is too small for the standard RF probes. To make the devices reachable by probes, probe pads and lead lines must be added on-wafer. Fig. 3.5 shows a picture of an actual device with connection pads on the wafer. This kind of layout is referred as a ground-signal-ground (GSG) pad, which is widely used in industry. Depending on the purpose of the measurement, connection pads can be assigned to any ports in a device. For common emitter configurations of HBTs, the top pads and the bottom pads are shorted and connected to the emitter, while the two middle pads are connected to the base and collector, respectively, allowing measurements to be performed on a common-emitter HBT

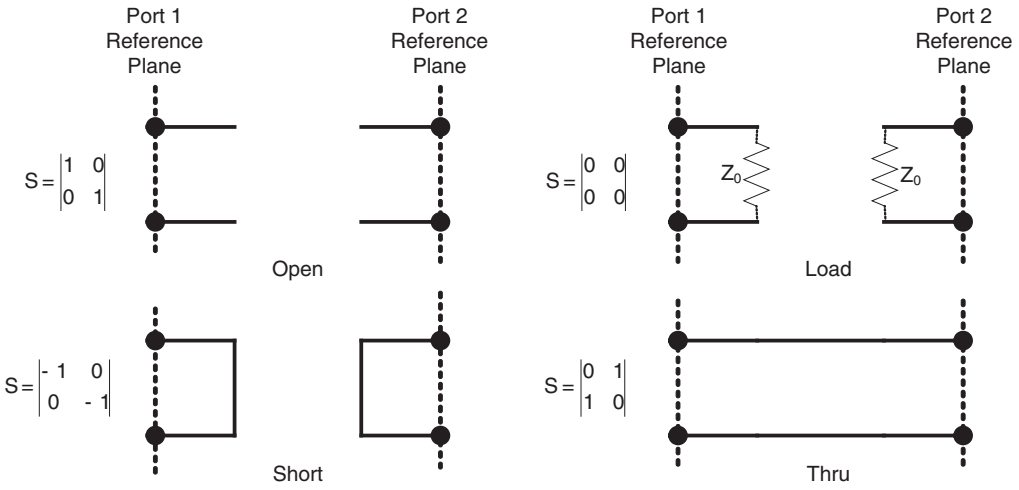


Figure 3.4: Electrical definition for ideal OSLT standards.

amplifier. Different size of HBTs come with different connection pads. At high frequencies, these probe pads and lead lines are like capacitors and inductors. The measured S-parameter is a combination of these parasitics and the device. Thus on-wafer de-embedding is another necessary step which must be taken in order to eliminate the effects of parasitics in the probe pads and lead lines.

Fig. 3.6 shows the equivalent circuit diagram of the DUT. In this model, the distributed parasitics are simplified into three parallel elements and three series elements between the ports and the device.

To eliminate the effects from these elements, each set of connection pads provides one short and one open structure. These two structures can be used to eliminate the effect of parasitics in the results by using the industry-standard “open-short” de-embedding method [33].

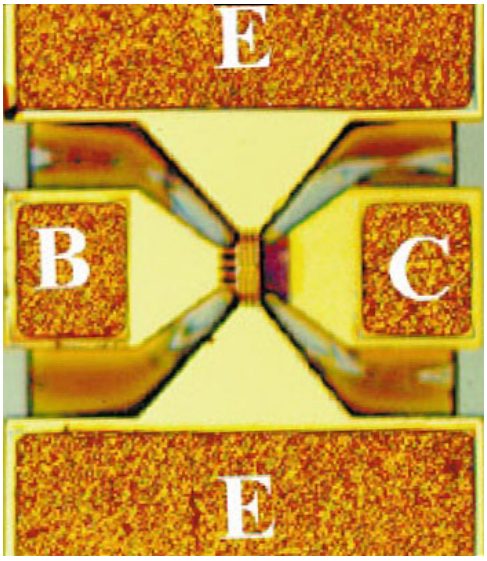


Figure 3.5: A microphotograph of a SiGe device with contact pads.

For the open pattern shown in Fig. 3.7, the Y-parameter matrix Y_{open} is:

$$Y_{open} = \begin{bmatrix} Y_{p1} + Y_{p3} & Y_{p3} \\ Y_{p3} & Y_{p2} + Y_{p3} \end{bmatrix} \tag{3.2}$$

Comparing the open structure with that of the DUT', one can easily see that the open structure is in parallel with DUT'. Thus,

$$Y_{DUT} = \begin{bmatrix} Y_{p1} + Y_{p3} + Y_{DUT',11} & Y_{p3} + Y_{DUT',12} \\ Y_{p3} + Y_{DUT',21} & Y_{p2} + Y_{p3} + Y_{DUT',22} \end{bmatrix} = Y_{DUT'} + Y_{open} \tag{3.3}$$

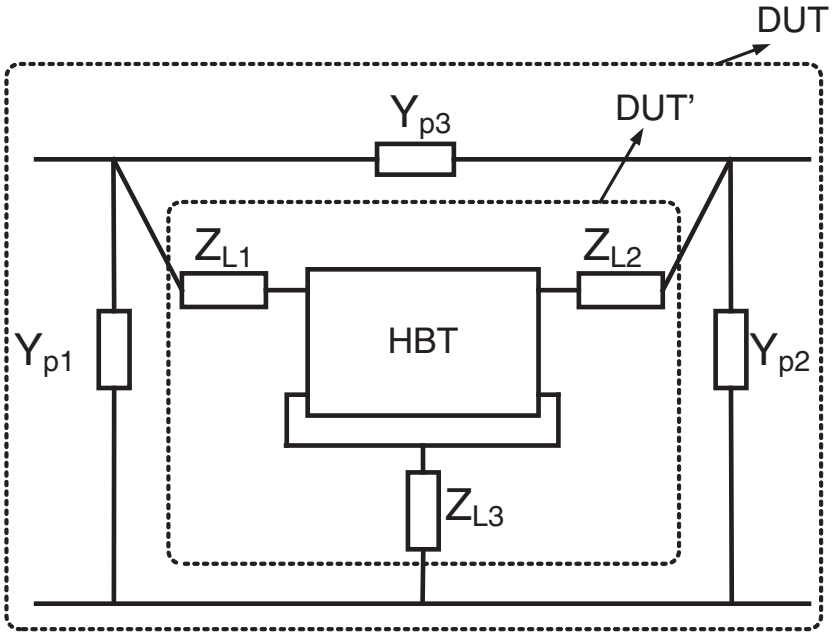


Figure 3.6: Equivalent circuit diagram used for "open-short" deembedding method, including both the parallel parasitics Y_{p1}, Y_{p2}, Y_{p3} and the series parasitics Z_{L1}, Z_{L2} and Z_{L3} surrounding the transistor.

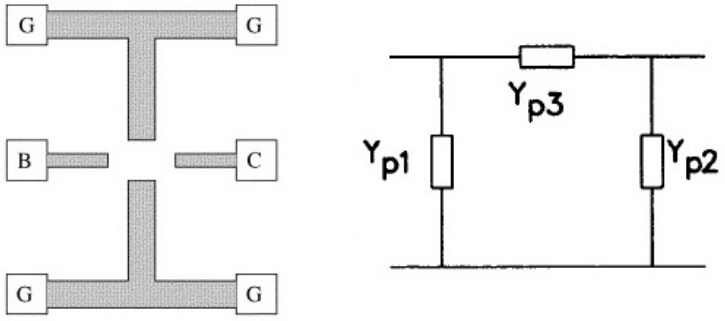


Figure 3.7: 'Open' pattern on wafer used to characterize the parallel parasitics. Also shown is the equivalent circuit diagram of this open pattern with parallel parasitics Y_{p1}, Y_{p2} and Y_{p3}.

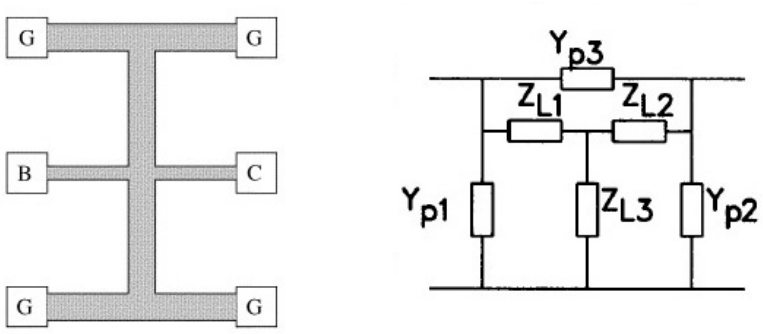


Figure 3.8: 'Short' pattern on wafer used to characterize the series parasitics. Also shown is the equivalent circuit diagram of this short pattern, with series impedances Z_{L1}, Z_{L2} and Z_{L3} embedded in parallel parasitics.

Similarly, for the short structure shown in Fig. 3.8, we obtain:

$$\begin{bmatrix} Z_{L1}+Z_{L3} & Z_{L3} \\ Z_{L3} & Z_{L2}+Z_{L3} \end{bmatrix} = \left(Y_{short}-Y_{open}\right)^{-1} \tag{3.4}$$

Since Z_{L1}, Z_{L2} and Z_{L3} are in series with the HBT, the Z parameter of DUT' can be written as:

$$\begin{bmatrix} Z_{L1}+Z_{L3}+Z_{HBT,11} & Z_{L3}+Z_{HBT,12} \\ Z_{L3}+Z_{HBT,21} & Z_{L2}+Z_{L3}+Z_{HBT,22} \end{bmatrix} = Z_{HBT}+\begin{bmatrix} Z_{L1}+Z_{L3} & Z_{L3} \\ Z_{L3} & Z_{L2}+Z_{L3} \end{bmatrix} \tag{3.5}$$

From Eq. 3.2-Eq.3.5, the Y parameter of the HBT can be calculated as:

$$Y_{HBT} = \left(\left(Y_{dut}-Y_{open}\right)^{-1}-\left(Y_{short}-Y_{open}\right)^{-1}\right)^{-1} \tag{3.6}$$

with Y_{dut} as the measured Y-parameter matrix of the HBT together with parasitics and Y_{HBT} as the actual Y-parameter matrix of the HBT. The "Open-Short" de-embedding method works well up to 40 GHz. After 40 GHz, modelling the parasitics using 6 lumped elements will introduce noticeable accuracy problems and a set of more complicated calibration structures will need to be fabricated [34].

The detailed measurement procedure can be found in Appendix.A.

3.2 Large Signal Measurement

The measurement setup for on-wafer power measurement is exactly the same as that used for the S-parameter measurement. The only difference is the use of the HP 8510C Vector Network Analyzer in power domain rather than the frequency domain, which is enabled by firmware version 7.16. Using the power domain allows the measurements of a DUT over the power range of interest at a constant frequency.

3.2.1 Power Calibration

Here, the 8510C system was used to measure the output power versus input power and the Agilent 6626A DC power supply was used for biasing the device. The RF power was supplied from port 1 of the VNA. An attenuator was inserted between the DUT and port 2 of the VNA in order to keep the input power to the VNA below specification (17 dBm). To make sure that the input power values read were the actual input powers to the DUT, a power flatness calibration as described in the S-parameter section must be performed. To ensure the output power values read are the actual output power out of the DUT, the receiver calibration is another necessary step. Otherwise, the power levels displayed are those determined by the source and do not account for

losses in the path between the source and the test ports [27]. A detailed description of receiver calibration can be found in Appendix.B.

3.2.2 Large Signal Figure of Merits

HBTs are nonlinear devices, which means the output RF signal will not be a linear replication of the input RF signal. At a very low input signal level, however, the output signal can be viewed as a proportional amplification of the input signal. Fig. 3.9 shows a typical output power versus input power curve for HBTs. As shown in the figure, when the input RF power is low, the HBT's power gain is a constant. This region is typically called the linear region, the output power follows the input power. A 1:1 ratio can be observed in this region. As the input power level increases, a point is reached where the power of the signal at the output is not amplified by the same amount as the smaller signal, which is defined as the onset of compression. At the point where the input signal is amplified by an amount 1 dB less than the small signal gain, the 1 dB Compression Point has been reached. A rapid decrease in gain will be experienced after the 1 dB compression point is reached. After a certain input power level, the output power no longer increases simply because of the operation range limits of the HBTs. The output power in this region is referred to as the saturated power, or the maximum output power. If the input power is increased to an extreme value, the HBT will be destroyed.

Another very important parameter is the efficiency. Active devices require one or more DC power supplies. Efficiency denotes how much of the supplied power is transferred into output power. The most commonly used efficiency measure is the power added efficiency (PAE). PAE is calculated as the output power minus the input power, divided by the DC power.

$$\text{PAE} = \frac{P_{out,RF} - P_{in,RF}}{P_{DC}}. \tag{3.7}$$

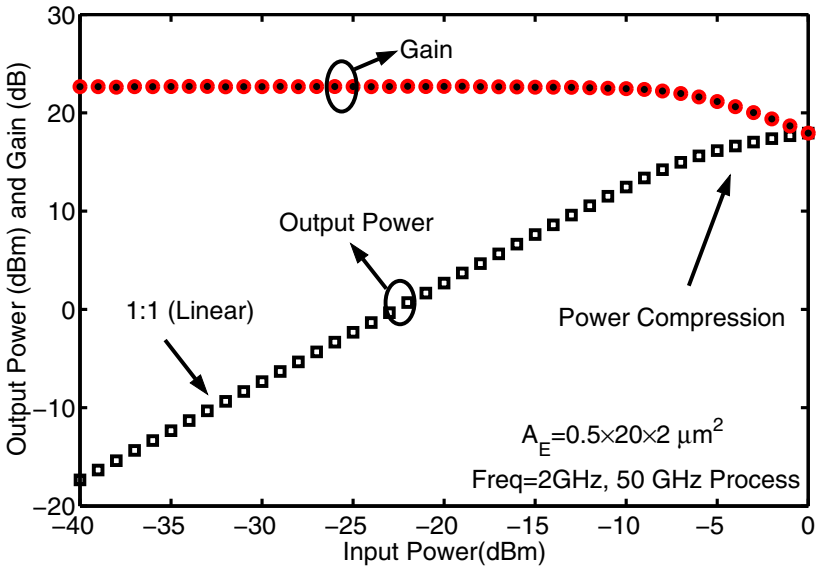


Figure 3.9: A typical output power versus input power for an active nonlinear device.

Efficiency is a function of the device operating conditions, including the operating voltage, class of operation (A, B, AB, etc), drive level, frequency, and temperature.

3.3 Third Order Intermodulation

3.3.1 Third Order Intermodulation Basics

In an ideal linear system, the signal is amplified without any distortion. However, in any real device the transfer function is usually a lot more complicated, generally due to the nonlinearity of either the active device or the power supply. To express this in mathematical terms, first write the transfer function of a nonlinear system as a series expansion of power terms:

$$v_{out} = k_0 + k_1 v_{in} + k_2 v_{in}^2 + k_3 v_{in}^3 + ... \tag{3.8}$$

Table 3.1: Outputs from Nonlinear Systems with Inputs at ω_1 and ω_2.

	Symbolic Frequency	Name	Comment
First order	ω_1, ω_2	Fundamental	Desired output
Second order	$2\omega_1, 2\omega_2$	HD2 (harmonics)	Can filter
	$\omega_2 - \omega_1, \omega_2 + \omega_1$	IM2 (mixing)	Can filter
Third order	$3\omega_1, 3\omega_2$	HD3 (harmonics)	Can filter
	$2\omega_1 - \omega_2, 2\omega_2 - \omega_1$	IM3 (intermod)	Hard to filter

To describe the nonlinearity perfectly, an infinite number of terms is required; however, in many practical circuits, the first three terms are sufficient to characterize the system with a fair degree of accuracy. For a two-tone input $v_{in} = A\cos\omega_1 t + A\cos\omega_2 t$, a simple expansion of Eq. 3.8 can show that the output will usually consist of signals at $\omega_1, \omega_2, 2\omega_1, 2\omega_2, 3\omega_1, 3\omega_2, \omega_1 + \omega_2, \omega_1 - \omega_2, 2\omega_1 + \omega_2, 2\omega_1 - \omega_2, 2\omega_2 + \omega_1,$ and $2\omega_2 - \omega_1$. The name and comment for each frequency component is summarized in Table. 3.1. The first order components are the desired outputs, referred to as fundamentals and the frequencies of these are the same as the input signals. The second order and the third order components include harmonics, mixing, and intermodulations of the two input tones. Usually the harmonics and mixing are not a problem, since they are far away from the desired signals. The third order intermodulations $(2\omega_2 - \omega_1, 2\omega_1 - \omega_2)$, however, are often major concerns.

As shown in Fig. 3.10, when ω_1 and ω_2 are closely spaced, the third-order intermodulation products are very close to ω_1 and ω_2 in the output and thus cannot be filtered out. Consider a weak desired signal channel, and two nearby strong interferers passing through an amplifier. As shown in Fig. 3.11, one of the two intermodulation products falls in the band of desired signals, and corrupts the desired component.

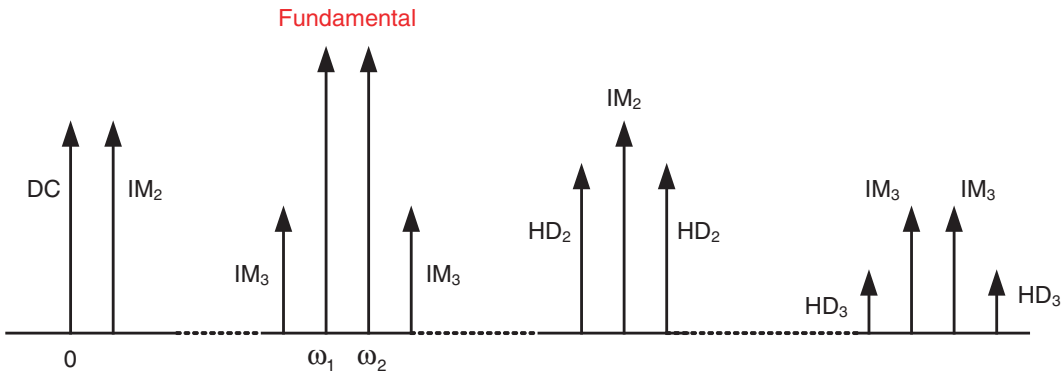


Figure 3.10: The different frequency components at the output of a nonlinear circuit for the input of two sinusoidal signals with the same amplitude.

The fundamental signal and intermodulation products in the output are given by [35]:

$$\begin{aligned} v_{out} =& (k_1 A + \frac{3k_3 A^3}{4} + \frac{3k_3 A^3}{2}) \cos \omega_1 t + \ldots && \text{fundamental} \\ &+ \frac{3k_3 A^3}{4} \cos(2\omega_2 - \omega_1) + \ldots && \text{3rd order intermod} \end{aligned} \tag{3.9}$$

We can then define the ratio of the amplitude of the IM product to the amplitude of the fundamental output as the **third-order intermodulation distortion IM3**. Neglecting the higher order terms added to $k_1 A_1$ in the amplitude of the fundamental term, one gets:

$$\text{IM3} = \frac{3k_3 A^3}{4} / k_1 A \tag{3.10}$$

Note that for small A, the fundamental rises linearly, and the IM3 terms rise as the cube of the input. Theoretically, there is an amplitude A that is large enough so that the IM3 terms would be equal to the fundamental. The input power at this point is called the *input third-order*

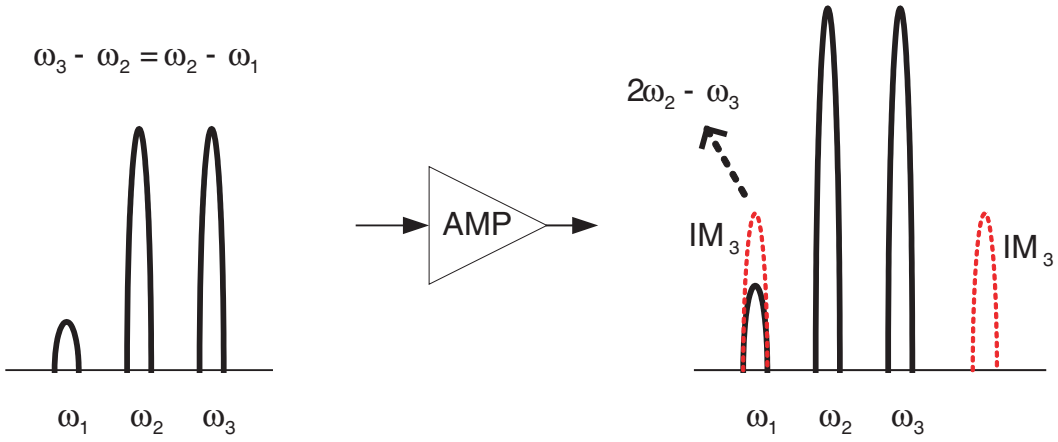


Figure 3.11: Illustration of how strong interference may cover a weak desired signal.

intercept point (IIP3). The output power at this point is called the *output third-order intercept point* (OIP3).

Of course, the third-order intercept point cannot actually be measured directly, since by the time the amplifier reached this point, it would be heavily overloaded. Therefore, it is useful to describe a quick way to extrapolate it at a given power level. Assume that a device with power gain G has been measured to have an output power of P_{1st} at the fundamental frequency and a power of P_{3rd} at the IM3 frequency for a given input power of P_{in}. Now, on a log plot (for example, when power is in dBm) of P_{3rd} and P_{1st} versus P_{in}, the IM3 terms have a slope of 3:1 and the fundamental terms have a slope of 1:1, as shown in Fig. 3.12. When input power is low, P_{3rd} and P_{1st} can be written as:

$$P_{1st} = P_{in} + G \tag{3.11}$$

$$P_{3rd} = 3 \times P_{in} + C \tag{3.12}$$

where G and C are constant. According to the definition of IIP3,

$$\text{IIP3} + G = 3 \times \text{IIP3} + C \tag{3.13}$$

$$\Rightarrow 2 \times \text{IIP3} = G - C \tag{3.14}$$

$$\Rightarrow \text{IIP3} = \frac{G - C}{2} \tag{3.15}$$

Since

$$P_{1st} - P_{3rd} = -2 \times P_{in} + G - C \tag{3.16}$$

one can easily get

$$\text{IIP3} = P_{in} + \frac{1}{2}[P_{1st} - P_{3rd}] \tag{3.17}$$

3.3.2 Measurement Setup

The IIP3 measurement system consists of an Agilent 8563EC Spectrum Analyzer (9kHz-26.5GHz), two Agilent E8247C PSG CW Signal Generators (20GHz), an Agilent 6625A Precision System Power Supply (25W or 50W, 2 outputs), an Alessi REL-4300 microwave probe station, Infinity I40A GSG 150 microwave probes made by Cascade Microtech (dc-40GHz), and power dividers, attenuators, and bias Tees.

Fig. 3.13 shows a schematic of the measurement setup. The two RF signals are provided by two E8247C signal generators. The signals pass through two attenuators and combine after the hybrid power combiner (divider). The Agilent E8247C signal generator can provide RF power over the range from -20 dBm to 20 dBm. Taking into account the cable attenuation, the lower

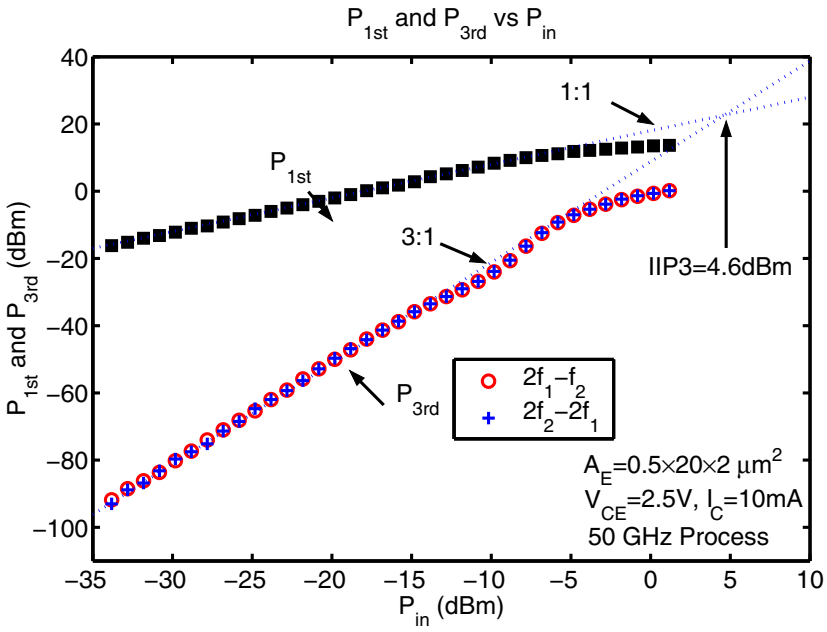


Figure 3.12: Illustration of the IIP3 extrapolation.

limit of power to the device is around -26 dBm, which is not low enough to measure IIP3. To reach at least -35 dBm of input power, attenuators must be used. The system has three sets of attenuators (6 dBm, 10 dBm and 20 dBm) to achieve different power ranges as needed. The attenuators also serve to isolate the two sources from each other: the signal from one source to another needs to pass through two attenuators while the signal to the DUT only needs to pass one attenuator. The hybrid power combiner provides an extra 19 dB isolation between the two sources.

DC bias to the DUT is provided by an Agilent 6625A Precision System Power Supply trough two bias Tees. The RF signals pass through the bias Tees and DUT, then reach the Agilent 8563EC spectrum analyzer.

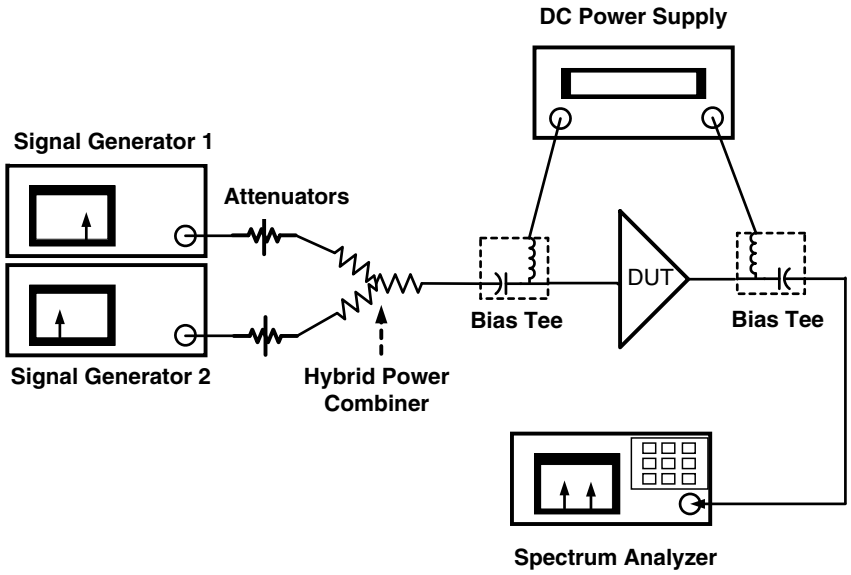


Figure 3.13: A schematic of the IIP3 measurement setup.

The system is controlled by an in-house VEE program which is able to automate the measurement and process the data to provide information on the IIP3, OIP3, power gain, harmonics, and DC bias.

3.3.3 Major Concerns in IIP3 Measurement

In an IIP3 measurement system, the system linearity is very important for the accuracy of the results. The relationship of the nonlinearity of the system to the nonlinearity of the device is like that of noise to signal; without careful adjustment, the measurement results are unjustified.

The IIP3 system includes two major non-linear parts, namely the Automatic Level Control (ALC) unit in the sources and the mixer in the spectrum analyzer. Other parts of the system

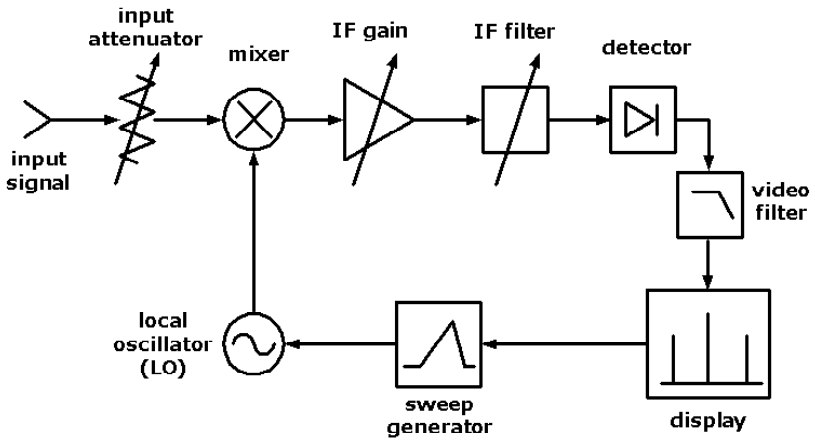


Figure 3.14: A simple block diagram of a typical superheterodyne spectrum analyzer.

are either passive or have negligible impact on the linearity of the system. In these two parts, the ALC unit can be easily shut off if the DUT is more linear than the ALC. The mixer of the spectrum analyzer, however, needs a more sophisticated treatment and it is first necessary to understand how a spectrum analyzer works.

Fig. 3.14 shows a simple block diagram of a typical superheterodyne spectrum analyzer. The spectrum analyzer consists of an input attenuator, a mixer which combines the input signal with the local oscillator, a filter with a gain stage before it, a peak detector, a video filter, a sweep generator and a display.

The input attenuator allows us to control the signal level into the spectrum analyzer to adjust its operating range and to keep from damaging the instrument. The mixer translates the input signal to an intermediate frequency (IF) that the spectrum analyzer can filter, amplify and detect. In addition to the input signal, the mixer receives a signal from the local oscillator whose frequency is controlled by the sweep generator. The IF gain stage adjusts the input level of the

IF filter and is attached to the input attenuator. When the input signal is attenuated, the IF signal is amplified by the IF gain stage, so that the signal level on the display stays the same. The IF filter is a fixed band-pass filter. Only the signal in the band of the IF filter goes to the detector section. The detector is a rectifier whose output follows the envelope or peak variation of the IF filter output. The purpose of the detector is to process the frequency components of the signal in order to display them. The last two parts of the block diagram are the video filter and the display. The video filter's job is to smooth the display by averaging the signal so that the fast, randomly changed spectrum components such as noise are suppressed on the display. Reduction of the noise makes low level signals easier to read on the display because they are less obscured.

There are three issues that must be considered in order to obtain the most accurate IIP3 result from a spectrum analyzer– resolution, sensitivity, and internal distortion.

Resolution is the ability to distinguish between closely spaced signals, which is usually determined by the bandwidth of the IF filter (RBF). The bandwidth usually means the 3 dB bandwidth, which guarantees a 3 dB dip at the edge of the bandwidth compared to the center in the filter's response. As shown in Fig. 3.15, a 30 kHz bandwidth is enough to separate two signals of equal amplitude spaced 30 kHz from each other. A larger bandwidth makes it harder to distinguish between two close signals.

In an IIP3 measurement, however, the amplitude of the third tone is usually many orders smaller than that of the fundamental tone. As shown in Fig. 3.16, a 10 kHz filter is able to distinguish the two fundamental tones, but the adjacent third tone signal is totally covered by the skirt of the 10 kHz filter. To resolve two closely spaced unequal amplitude signals, a smaller bandwidth RBF must be used. For the example in Fig. 3.16, a 1 kHz filter can distinguish two signals spaced 10 kHz from each other in frequency, with a 60 dB difference in power.

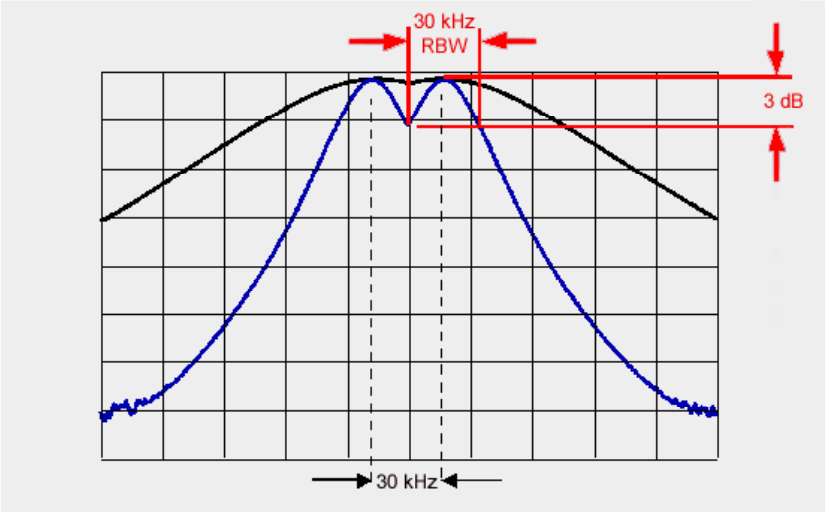


Figure 3.15: Illustration of using filters on signals of equal amplitude.

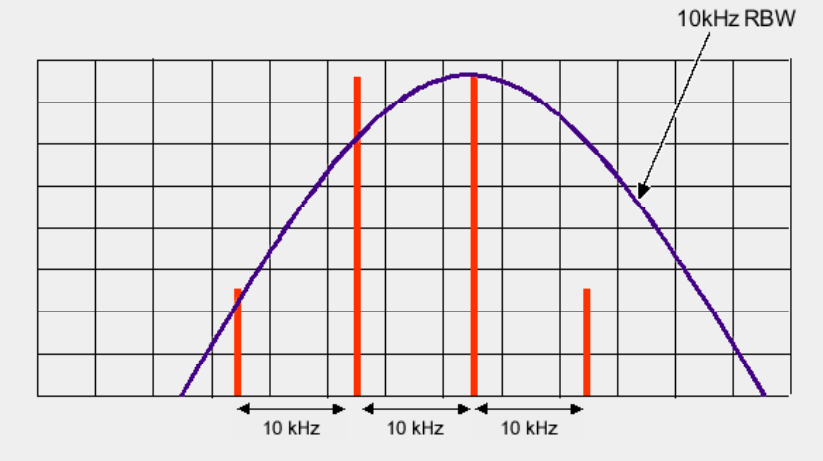


Figure 3.16: Illustration of using filters on signals of unequal amplitude.

Sensitivity determines the spectrum analyzer's ability to detect low level signals. A perfect receiver adds no additional noise to the input signal. The minimum noise floor is determined by

$$\text{Minimum Noise Floor} = 10\log kTB = -174\text{dBm/Hz @ 25C}, \tag{3.18}$$

where k is Boltzmann's constant, T is the temperature in Kelvin, and B is the bandwidth. However, because the spectrum analyzer is not an ideal receiver, the minimum noise floor of the spectrum analyzer is higher than -174 dBm/Hz. Since the internal noise of the spectrum analyzer is thermal in nature, it is random and has no discrete spectrum components. In addition, its level is flat over a very wide frequency range, which is certainly why it is comparable to the resolution bandwidth. This means the total noise reaching the detector is related to the resolution bandwidth selected. Since the noise is random, it is added on a power basis. So the relation is still a 10log function as shown in Fig. 3.17. Ten times resolution bandwidth generates 10 dB increase in the minimum noise floor.

Internal distortion is generated by the mixer in the spectrum analyzer. In fact, the distortion products generated by the mixer are at the same frequency as the distortion products coming from the DUT. At worst, the internal distortion can completely mask the external distortion products of the device. Even the level of the internal distortion is below the level of the external distortion, there could still be some error added to the results. For IIP3 measurements, the third order product of the internal distortion is the most important. The level of the third order intermodulation is directly related to the input level of the mixer. The relation is the same as in most of the nonlinear devices: the third order intermodulation increases as the cubic of the fundamental, which means on a log scale the level of the third order intermodulation changes three times as fast as the fundamental. By increasing the value of the input attenuator, the

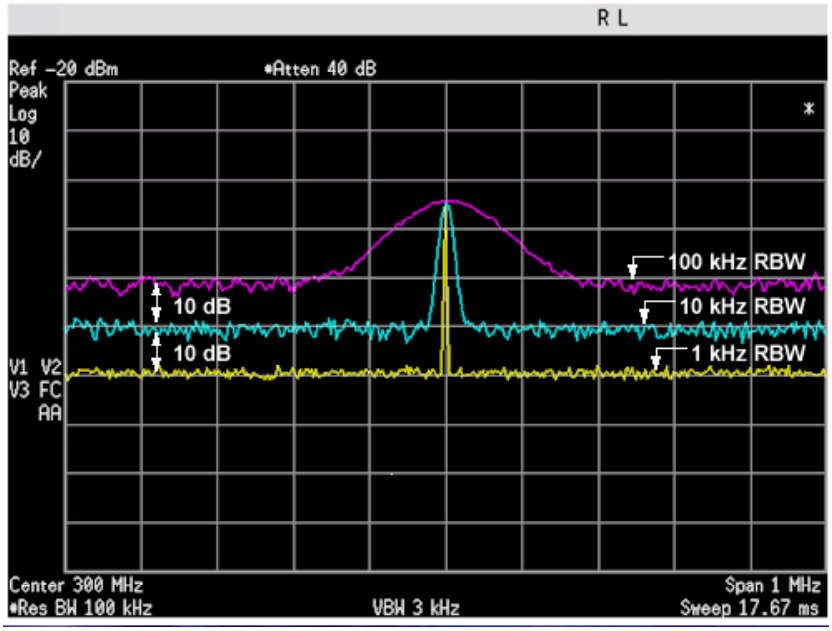


Figure 3.17: Illustration of the relation between resolution bandwidth and spectrum analyzer's noise floor.

internal third order intermodulation can be decreased to ensure accuracy reading of third order intermodulation by the device.

So to get high resolution and high sensitivity, the filter's bandwidth need to be small. The disadvantage of choosing a small bandwidth filter is that the measurement time will be increased accordingly. IIP3 measurements are time-consuming. A complete IIP3 measurement for a HBT device could take about 10 hours depending on how many bias points needed. That is why a just small enough filter is always favorable. Most of the internal noise is generated in the first active stage after the mixer. So the RF input attenuator has no effect on the actual noise level. However, the RF input attenuator does affect the signal level at the input and therefore the attenuation decreases the signal to noise ratio of the analyzer. So the best signal to noise ratio is achieved with the least amount of RF attenuation. This, however, conflicts with getting the least internal distortion because the input attenuator's value need to be high to get low internal distortion. As shown in Fig. 3.18, the noise to signal ratio decreases with increasing input power, while the ratio of the internal third order intermodulation to the fundamental increases with increasing input power. If we define Dynamic Range as the ratio of the largest and smallest signals that can be measured simultaneously, the point where these two curves cross corresponds to the input power level at which the dynamic range is the largest. Note in Fig. 3.18, if the bandwidth of the filter is changed, the noise to signal ratio will also change as discussed before.

By choosing the best dynamic range using the above method, the measurement system is at its best condition for measuring IIP3. To avoid the possibility that the measurement system is still not linear enough for the DUT, a simple test can be used to verify the validity of the results. Remember the input attenuator and the IF gain is attached, increase the input attenuator by 10 dB will not change the power on the display because the IF gain automatically compensate the attenuation. The input power level to the mixer, however, decreased by 10 dB. This way, the

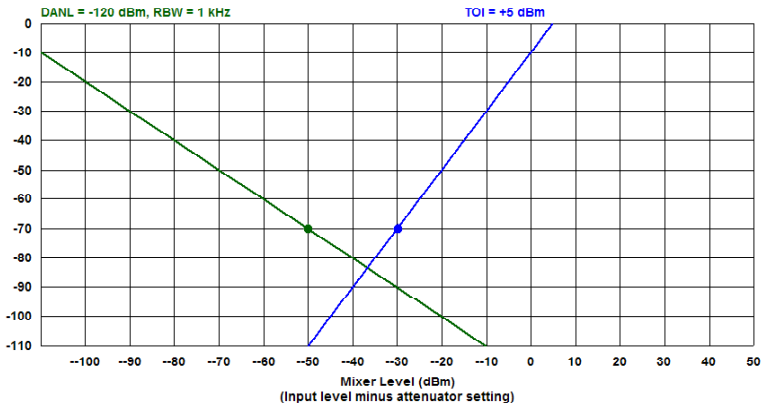


Figure 3.18: Determination of the best dynamic range.

internal distortion is decreased. If the third tone signals on the display changed noticeably, then the internal distortion gives error in the results. If the third tone signals do not change with the attenuator, then the internal distortion is small enough for the measurement [36].

3.3.4 IIP3 System Verification

To ensure that the system is able to measure highly linear active devices, several measurements were conducted on two linear elements, one through and one 12.5 Ω load. The connections for the two linear elements are shown in Fig. 3.19. The through has little reflection so that the power from two sources has little problem reaching the spectrum analyzer. The IIP3 measured on the through is thus able to tell the spectrum analyzer's limit. For the 12.5 Ω load, however, some reflections are expected which is a good simulation of real measurement conditions as the DUTs are not always matched to 50 Ω.

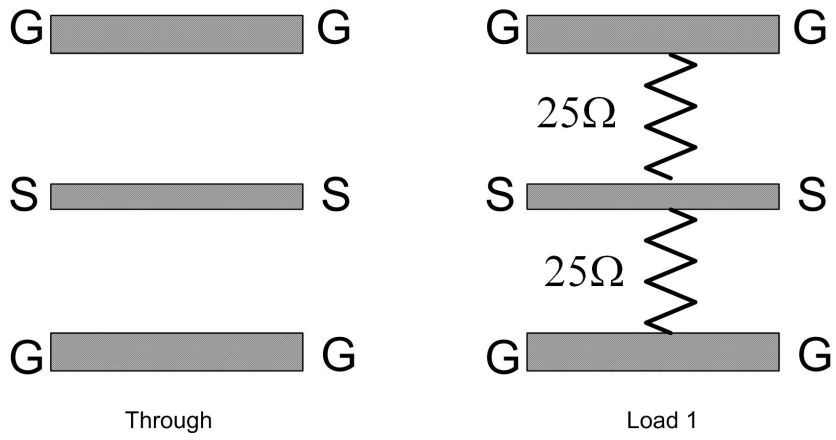


Figure 3.19: Illustration of the two linear elements used.

Fig. 3.20 shows the measured results for both elements. The first order output power shows a very nice 1:1 relationship to the input power while the third order output power of both elements is in the form of steps. This is because the third order output power is always under the noise level of the system. The noise level increases when the attenuator in the spectrum analyzer automatically increases its value with input power to the spectrum analyzer in 10 dBm steps. These results verify that the nonlinearity of the system is low enough to be useful for any active devices which have a third order output power over the noise level of the system in the input power range used. The detailed measurement process and program used can be found in Appendix.C

3.4 Summary

In this chapter, RF characterization methods were discussed, including the S-parameters, large signal power characterization and third order intermodulation. In general, the S-parameters

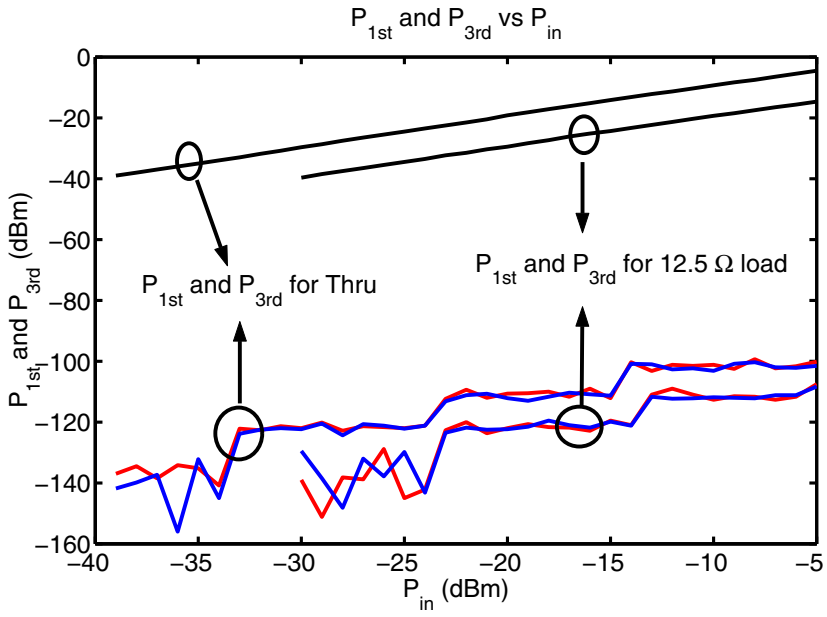


Figure 3.20: P_{1st} and P_{3rd} versus P_{in} for One through and One 12.5 Ω load.

are used to characterize the small signal parameters and can be used to extract the base resistance R_B, the cut-off frequency f_T and the maximum oscillation frequency f_{max}. The large signal system built utilizes the same equipment setup as for the S-parameter system, while being able to measure device performance at large input powers. By monitoring the DC voltage and current, the power added efficiency can be calculated. The third order intermodulation system constructed in this work is more complex than the previous two in the sense that the system distortion level can have a significant effect on the accuracy of the measurement result. By carefully setting the system, accurate IIP3 measurements are possible for both HBTs and MOSFETs [37] [38]. All three of the systems are controlled by in-house programs written in VEE. The program for the S-parameter system was originally written by William E. Ansley [39], but

the program was modified in this work in order to increase the stability of operation and application range. Programs for the large signal system and the IIP3 system were written during the construction of both systems. All of these programs are now capable of measuring both HBT and MOSFET devices with high accuracy, while requiring little attention from operators.

CHAPTER 4

LARGE SIGNAL PERFORMANCE OF SIGE HBTS

SiGe BiCMOS has recently become the mainstream technology for implementing high performance and cost effective radio transceivers. The core of the technology is the SiGe Hetero junction Bipolar Transistor (HBT), which has demonstrated excellent small-signal gain and noise performance at RF and microwave frequencies. The RF performance of SiGe HBTs depends heavily on the device-level layout and profile design. SiGe profile optimization for small signal performance such as current gain β, base resistance r_b, cut off frequency f_T, maximum oscillation frequency f_{max}, minimum noise figure NF_{min}, and $1/f$ noise corner frequency f_c has been discussed extensively [40]- [42].

For practical applications, not only the small signal performance but also the large signal performance must be considered. For example, the 1 dB compression point of a low noise amplifier is very important besides the the gain and noise performance. For the transistors used in power amplifiers, which are currently the weakest link in applying the SiGe BiCMOS technology to radio transceivers, large signal performance is the prime concern. An examination of large signal performance versus SiGe profile design will greatly enhance the device level design ability. This will be discussed below using SiGe HBTs featuring multiple SiGe profiles. We note that the impact of SiGe profile design on large signal linearity was examined using numerical simulation in [43].

A common concern when using SiGe HBTs for large signal RF applications (e.g. power amplifiers) is its low breakdown voltage. A typical SiGe BiCMOS process provides multiple breakdown voltages through selective collector implantation for design leverage. Technology

scaling, on the other hand, has led to devices with peak f_T value of as high as 375 GHz [2]. The nature of bipolar transistor operation dictates the use of higher current density J_C for higher f_T, which leads to an inevitable decrease in the breakdown voltage. In this chapter, we will discuss how these tradeoffs affect the large signal performance of SiGe HBTs.

Finally, the impact of different bias conditions on PAE is shown for circuit design reference. Higher bias voltages shift the peak PAE location, while higher bias currents increase the peak PAE value.

4.1 Experimental Setup

The SiGe HBTs are biased in the common-emitter configuration using two bias Tees with an RF termination of 50 Ω, as shown in Fig. 4.1. The output power (P_{out}) as a function of input power (P_{in}) is measured using the Agilent 8510C Vector Network Analyzer (VNA). This ability to operate the VNA in power domain is enabled by firmware ver. 7.16. The RF power is supplied from port 1 of the VNA and an attenuator is inserted between the collector and port 2 of the VNA to keep the input power to the VNA below specification (17 dBm). Power flatness calibration and receiver calibration [44] are performed in order to guarantee accuracy in the RF power readings at both the input and output. DC power is provided by an Agilent 6626A precision system power supply through two bias Tees. During the RF power sweep, the base and collector DC bias voltages ($V_{B,bias}$ and $V_{C,bias}$), are fixed. The DC base and collector currents ($I_{B,dc}$ and $I_{C,dc}$), however, increase with increasing RF power. The $I_{C,dc}$ at low input power (-80 dBm) is denoted as $I_{C,bias}$. P_{out}, P_{in}, $V_{B,bias}$, $V_{C,bias}$, $I_{B,dc}$, and $I_{C,dc}$ are recorded using an in-house program written in Agilent VEE instrument control programming language. The measurements are done on-wafer.

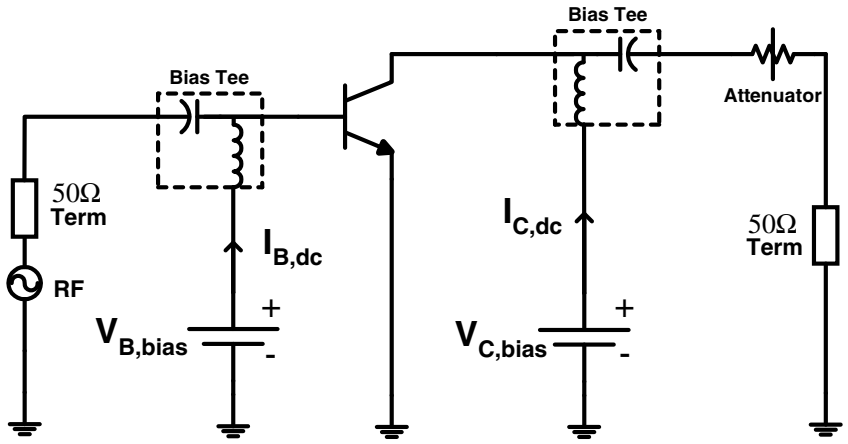


Figure 4.1: A diagram of the experimental setup.

The validity of the measurement results is checked by comparing the simulation results with the measured results for the same device. Fig. 4.2 shows that the measured results are quite dependable.

4.2 Results and Discussion

4.2.1 Impact of SiGe Profile Design

For a given SiGe BiCMOS process, the emitter width and base sheet resistance are fixed, thus r_b is fixed. RF noise figure can still be improved by increasing β and f_T, which reduces the amount of base current shot noise as well as the input referred noise due to the collector current shot noise [45]. The reduction of I_B for the same I_C also directly translates into a smaller $1/f$ noise corner frequency and lower phase noise. By managing the SiGe profile, the need for high β and high f_T can be fulfilled.

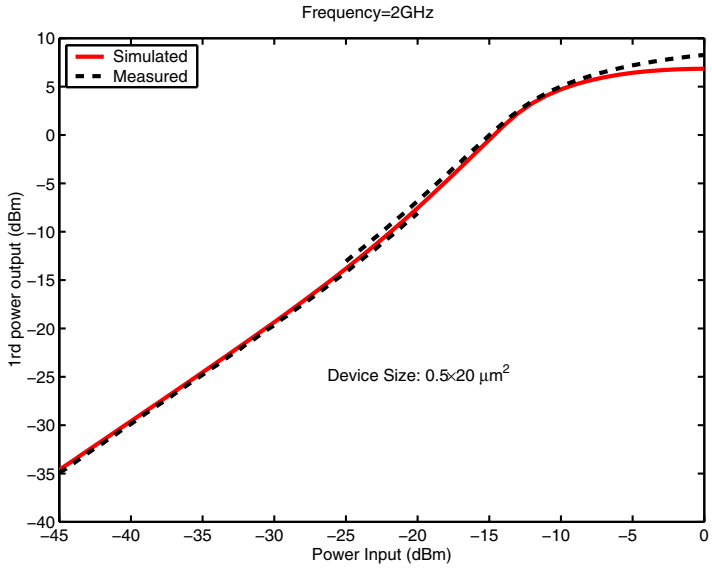


Figure 4.2: Comparison of cadence simulation result with the measured result for a 0.5×20 μm^2 SiGe HBT.

However, the total integrated Ge content that can be put into a HBT is limited by the SiGe film stability. In order to obtain a higher β and higher f_T, more Ge and a steeper Ge gradient can be applied to the neutral base, forcing a larger Ge gradient into the collector-base space charge region. Fig. 4.3 shows two such low-noise Ge profiles (LN1 and LN2) which maintain the stability and the slightly higher peak f_T compared to the peak f_T of the SiGe control profile (10% peak Ge percentage in the base), but have significantly lower NF_{min} in simulation (by 0.2 dB).

The retrograding of Ge in these designs does not have an impact on device operation at low injection levels, because of the carrier depletion in the CB space-charge region. Problems arise, however, at high injection levels. At a high collector current density J_C, the minority carrier charge is sufficient to compensate the ionized depletion charge in the CB space-charge region. At sufficiently high J_C, the neutral base pushes out (due to the Kirk effect), exposing the SiGe-Si

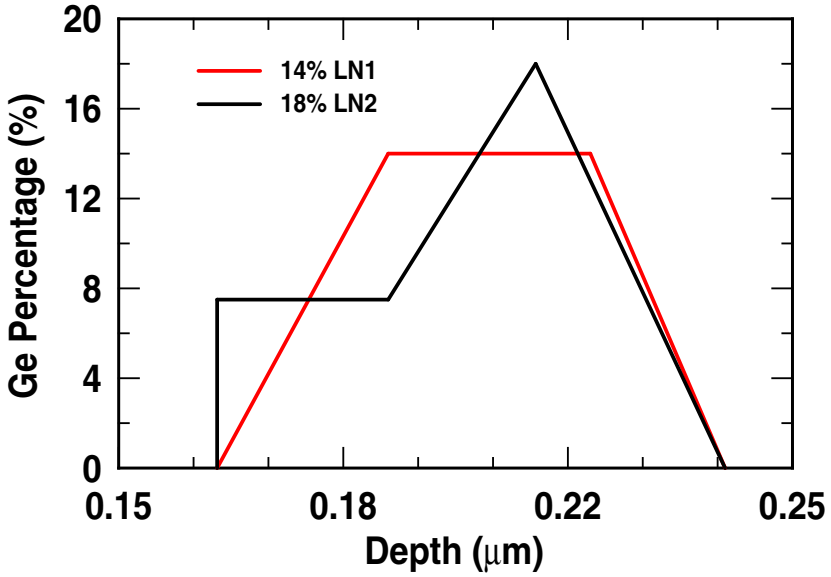


Figure 4.3: Schematic of the two optimized low-noise profiles that are both unconditionally stable.

heterojunction, which induces a conduction band barrier, and thereby strongly degrades both β and f_T [46] [47]. This is a tradeoff between high-J_C f_T performance and improved NF_{min} and $1/f$ noise.

To better illustrate the high-injection design trade-off, f_Ts are shown in Fig. 4.4. Three Ge-profile designs were examined, including a 10% peak SiGe control, a 14% peak low-noise design (LN1), and a 18% peak low-noise design (LN2) [40]. A Si BJT is also measured for comparison. All the SiGe HBTs have higher f_Ts compared to the Si BJT. Among the SiGe HBTs, LN1 and LN2 have higher f_Ts. In contrast, the 10% SiGe control design has a weaker f_T roll-off at high injection due to the deeper Ge retrograding to the collector. Si BJT has an even weaker f_T roll-off compared to the SiGe control HBT because there is no heterojunction barrier effect in a pure silicon device.

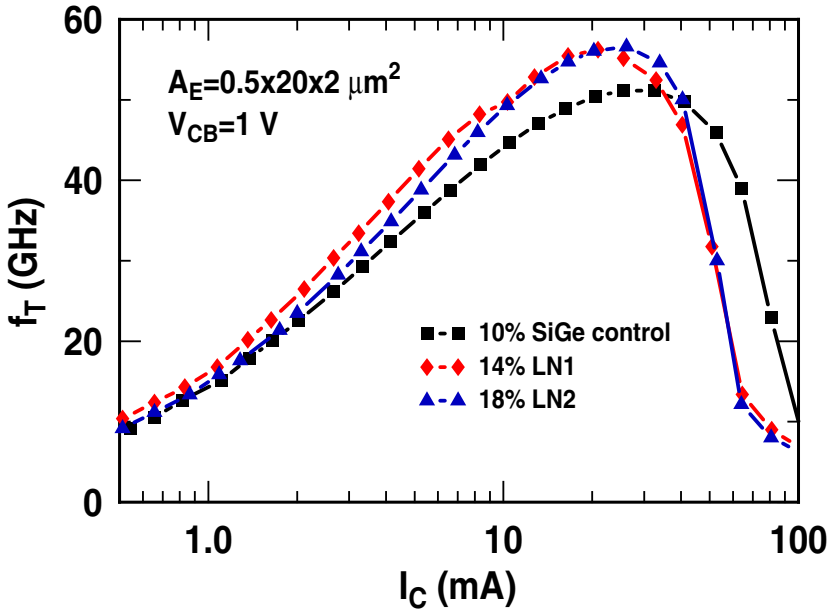


Figure 4.4: f_T versus I_C for Si, SiGe control, LN1, and LN2. V_{CB}=1 V.

As we will show below, the SiGe HBT can be driven into the high-injection f_T rolloff region at high input power levels. In that case, the large signal output power in SiGe designs optimized for low noise are likely to be worse (smaller).

Table. 4.1 shows the small signal performance of four devices. In addition to a much higher β, a modest increase in f_T is achieved in the two low-noise designs, primarily due to an increased Ge gradient in the neutral base. The f_{max} of the two low-noise designs are comparable to that of the SiGe control, indicating that the high power gain in the SiGe control design point is retained. The improvement of β and f_T translate into a clear improvement of the NF_{min} over the Si BJT and SiGe control profiles. The measured $1/f$ noise corner frequency is much lower in the two low-noise designs.

Fig. 4.5 compares the P_{out}, gain and power added efficiency (PAE) versus P_{in} characteristics at 2 GHz. $V_{C,bias}$=3 V, emitter area A_E=0.5 × 20 × 2 μm^2. $I_{C,bias}$=8 mA. At low P_{in}, P_{out} is

Table 4.1: Small signal performance of four profiles. I_C= 8 mA. Device size are $0.5\times20\times2$ μm^2.

Profile	Si	10%	LN1	LN2	(V_{CB})
β	51	73	250	225	0 V
f_T (GHz)	33	43	46	46	1 V
f_{max} (GHz)	53	56	58	57	1 V
NF_{min} (dB)	1.2	0.85	0.57	0.57	0 V
f_C (kHz)	26	15	4	4	0 V

about the same for all devices because the f_Ts are enough for 2 GHz. At high P_{in}, the SiGe control profile shows a higher maximum output power and a higher PAE than LN1 and LN2, while Si BJT has the largest maximum output power and highest PAE among all the devices.

Fig. 4.6 shows the $I_{C,dc}$ as a function of P_{in}. At high P_{in}, during part of the cycle, the transistor is driven into high injection. For SiGe HBTs, the input diffusion capacitance is thus higher in LN1 and LN2, resulting in more losses of power delivered to the transistor base. Furthermore, high injection barrier effect also degrades transconductance. The degradation is worse in LN1 and LN2 than in the SiGe control, leading to a reduced large signal output power, and lower PAE. This is confirmed by the Si BJT's result. Because there is no heterojunction barrier effect, although the Si BJT is driven harder into the high injection region, the degradation is the smallest among all devices, which gives the best large signal performance for the Si BJT. Therefore, there is a SiGe profile design tradeoff between small signal gain/noise performance and large signal performance. Circuits such as low noise amplifiers (LNA) require both good small signal and large signal performance. For these applications, the SiGe profile should be designed for not only sufficient gain and low noise, but also high large signal output power and PAE.

4.2.2 Speed Breakdown Tradeoffs

Typical SiGe processes offer devices with multiple collector profiles (breakdown voltages) in order to provide leverage in circuit design. A popular belief is that a higher breakdown voltage

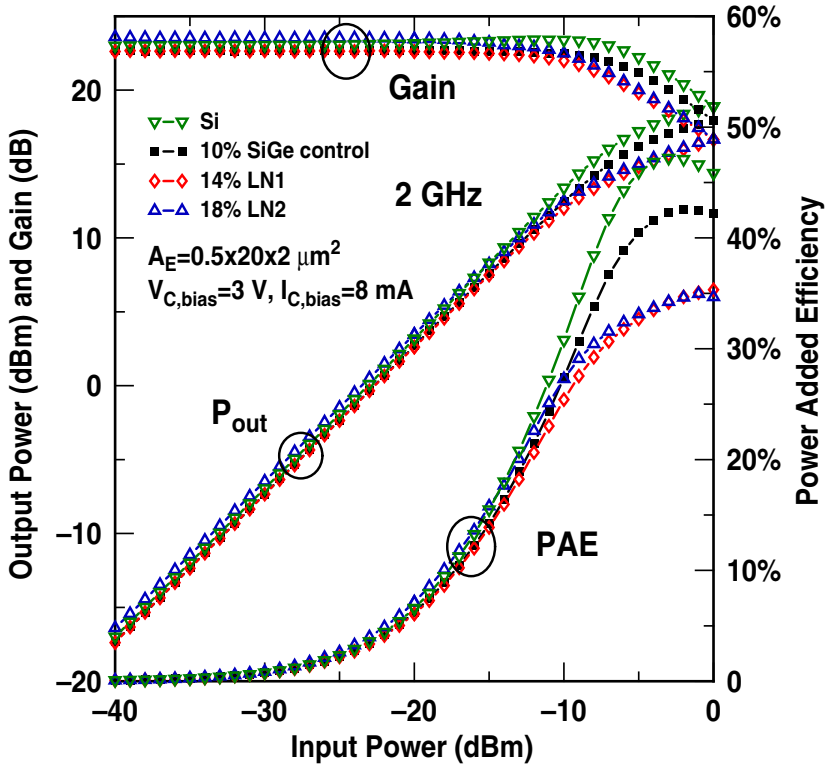


Figure 4.5: Gain, P_{out} and PAE versus P_{in} for SiGe control, LN1, and LN2.

enables a larger voltage swing and hence a higher output power [48] [49]. This, however, is only true for ideal transistor operation. In practice, high injection barrier effect occurs at a much smaller current in high breakdown voltage devices, because of the lower collector doping. A higher output power can only be effectively realized with an increase of both current and voltage. If the resulting instant current at high input power is high enough to cause high injection, the output power of the high breakdown voltage (HBV) device can be severely limited. One may then be forced to increase the device size as well to reduce biasing current density, which may be an issue if a large amount of output power is required. An increased size will also increase the parasitic input capacitance.

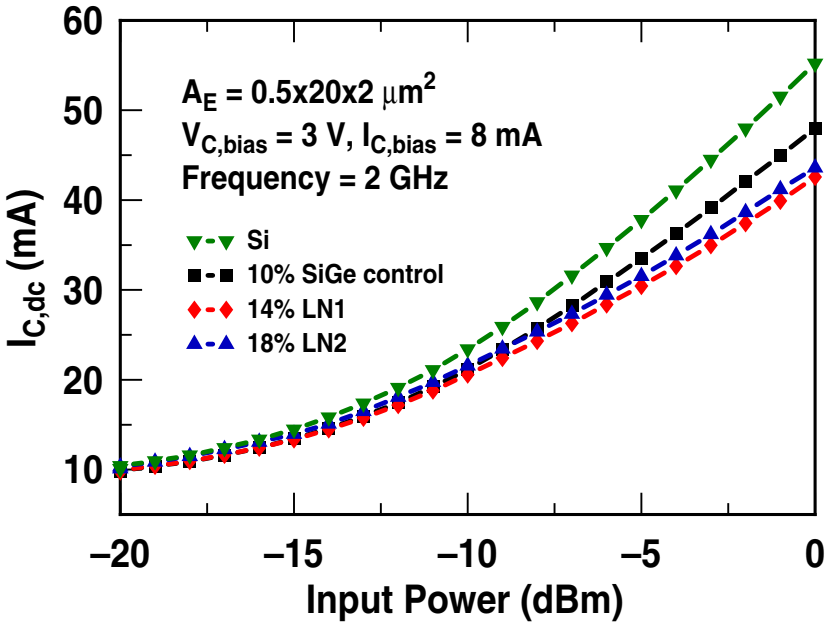


Figure 4.6: $I_{C,dc}$ versus P_{in} for SiGe control, LN1, and LN2.

Fig. 4.7 shows the gain, P_{out} and PAE versus P_{in} for the HBV device and the standard breakdown voltage (SBV) device. A_E=0.5 × 20 × 2 μm^2. $I_{C,bias}$=8 mA. $V_{C,bias}$=3 V for the SBV device. $V_{C,bias}$=3 V and 5 V are used to examine the impact of $V_{C,bias}$. The HBV device has a peak f_T of 28 GHz and a BV_{CEO} of 5.3 V. The SBV device has a peak f_T of 50 GHz and a BV_{CEO} of 3.3 V. A higher $V_{C,bias}$ in general helps suppressing the high injection f_T roll off by limiting the amount of base push-out at high current density. The $P_{out} - P_{in}$ data clearly shows that a higher $V_{C,bias}$ improves large signal output power, however, at the expense of reduced efficiency. Overall, the standard and high breakdown voltage devices show similar small signal performance. The large signal performance, however, is considerably worse in the high breakdown voltage device, because of the early onset of high injection effect caused by the low

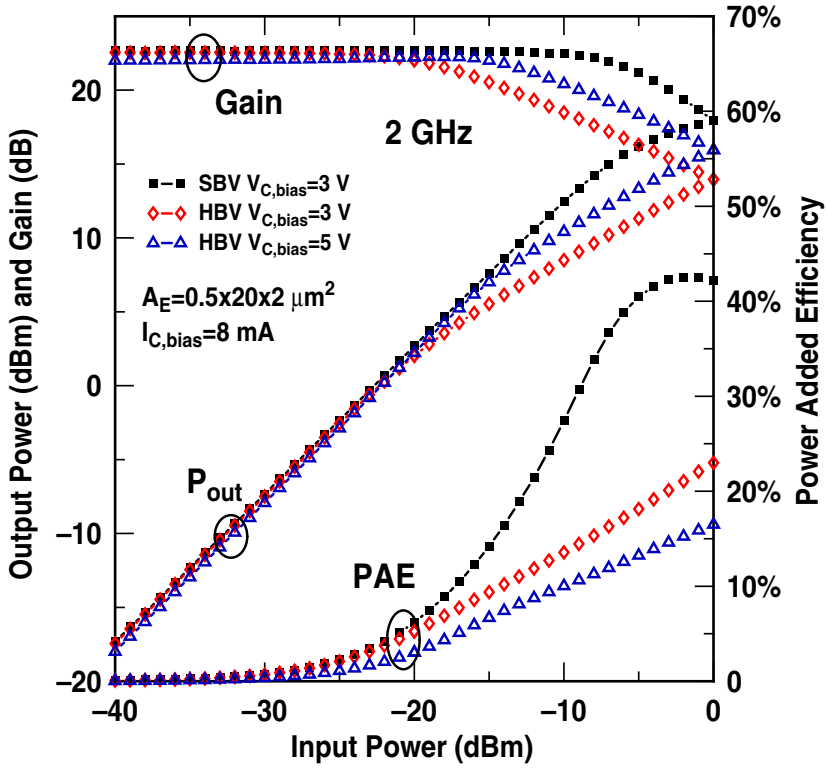


Figure 4.7: Gain, P_{out} and PAE versus P_{in} for the HBV and SBV devices.

collector doping. Both the 1 dB compression point and power added efficiency are much higher in the standard breakdown voltage device.

Fig. 4.8 shows the $I_{C,dc}$ vs P_{in} for the HBV and SBV devices. The DC current is clearly sufficient to drive the HBV device into high injection for P_{in} > -10 dBm according to the f_T curves shown in Fig. 4.9. Note that the actual instant collector current is higher than the DC current for a significant part of one signal period. For the same $V_{C,bias}$ of 3 V, the $I_{C,dc}$ at higher P_{in} is clearly larger for the SBV device than for the HBV device. This is also caused by the higher input capacitance and lower transconductance in the HBV device after high injection occurs.

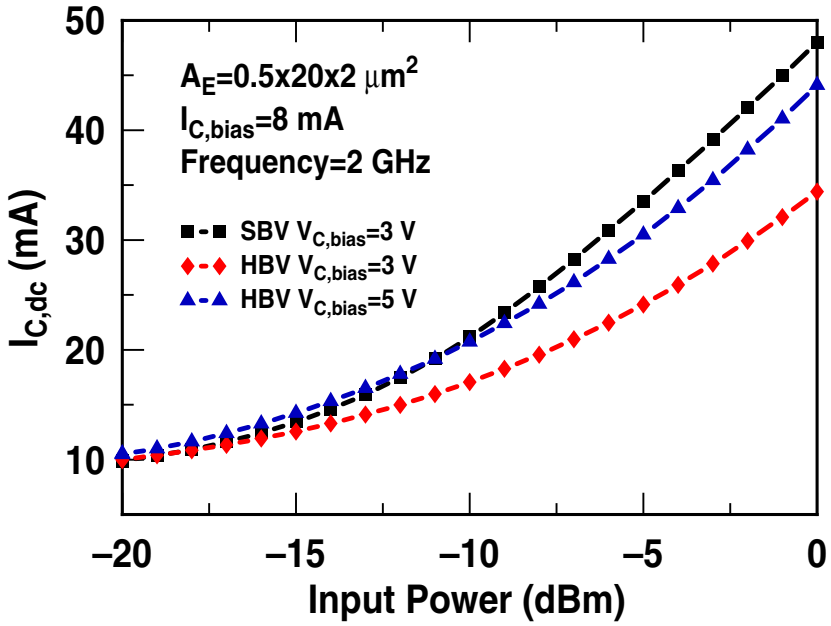


Figure 4.8: $I_{C,bias}$ versus P_{in} for the HBV and SBV devices.

A popular practice in SiGe RFIC design is to use the high breakdown voltage device as opposed to the standard breakdown voltage device as long as the f_T of the high breakdown voltage device gives sufficient gain or comparable gain [48] [50]. The advantage is that the CB capacitance is smaller, and thus there is less feedback, which is often undesired [48] [50]. Our results, however, show that the large signal performance must also be taken into account. The use of a high breakdown voltage device may inadvertently affect large signal performance. For instance, the 1 dB compression point of a LNA could be degraded when a high breakdown voltage device is used in place of a standard breakdown voltage device.

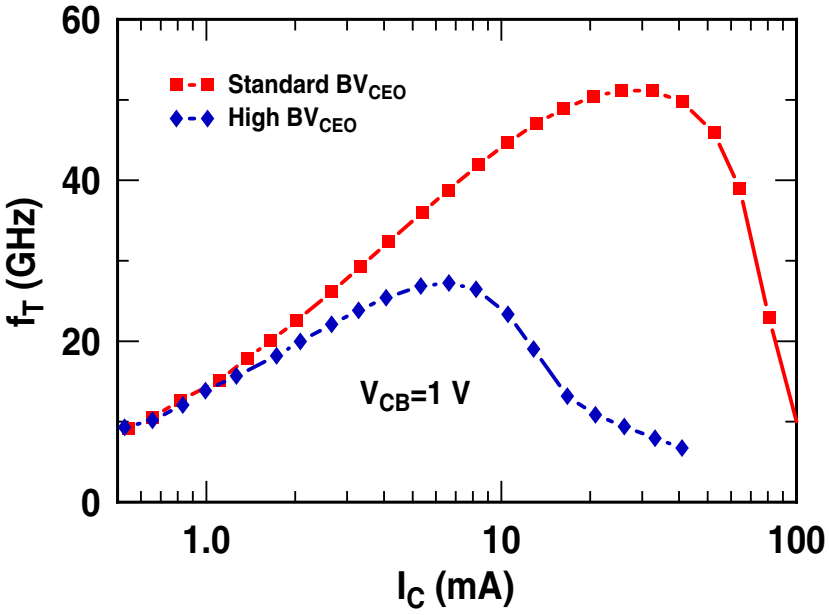


Figure 4.9: f_T versus I_C for the HBV and SBV devices.

4.2.3 Technology Scaling

One of the major concerns in SiGe HBT scaling for higher speed is the inevitable decrease of breakdown voltage, which could affect amplifier large signal performance. Here we measure the large signal performance of a SiGe HBT with a 200 GHz peak f_T. Fig. 4.10 shows the gain, P_{out} and PAE versus P_{in} measured at 20 GHz (1/10 of the peak f_T). For comparison, we also show results obtained on a 50 GHz peak f_T SiGe HBT. The 200 GHz peak f_T HBT is biased at $I_{C,bias}$=10 mA and $V_{C,bias}$=1.1 V. The 50 GHz peak f_T HBT is biased at $I_{C,bias}$ =10 mA, with $V_{C,bias}$=3 V and $V_{C,bias}$=1.1 V. The emitter area A_E for the 200 GHz device is 0.12 × 10 μm^2 and A_E for 50 GHz device is 0.5 × 20 × 2 μm^2. As can be seen from Fig. 4.11, at $I_{C,bias}$=10 mA, both devices are biased for achieving 80% of the peak f_T, which leaves room for current swing at high input power. As shown in Fig. 4.10, the small signal gain is remarkably higher

in the 200 GHz device than in the 50 GHz device (15 dB vs 4 dB), in part due to the much higher f_T, as expected. The 50 GHz device shows a flat gain in the whole range simply because of the linearization by the larger parasitic capacitances. However, with only 4 dB gain and less than 10% PAE, the 50 GHz peak f_T device is of limited use at 20 GHz. The maximum output power for the 200 GHz device is not very high in this case, as a small size device is used. The maximum output power can be increased with a larger size and a matching network. The peak PAE, an important figure-of-merit for large signal applications, is 36% at 20 GHz without any matching for a supply voltage of only 1.1 V.

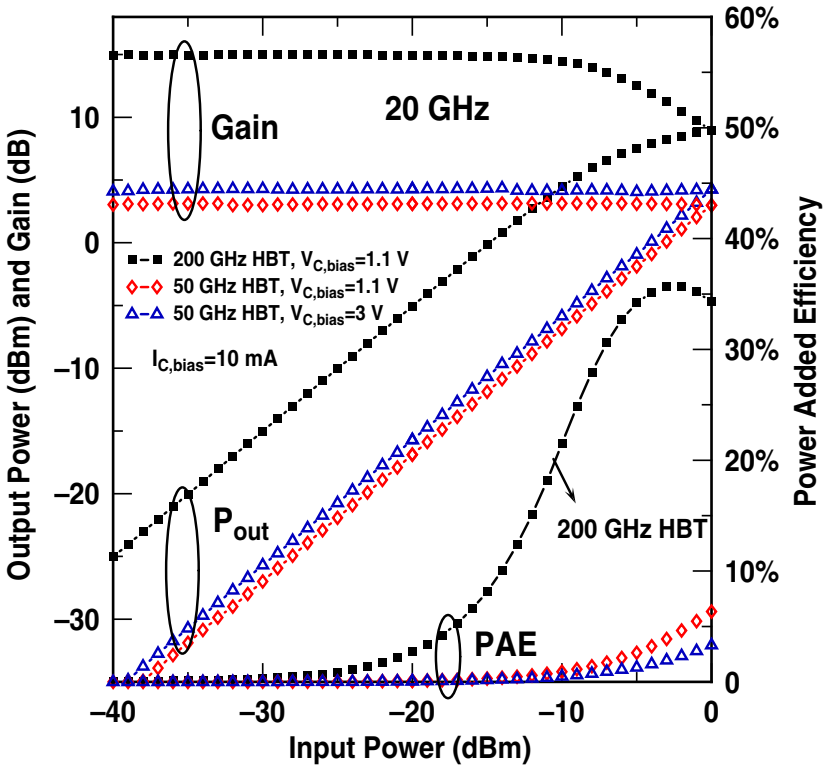


Figure 4.10: Gain, P_{out} and PAE versus P_{in} at 20 GHz for a 50 GHz peak f_T device and a 200 GHz peak f_T device.

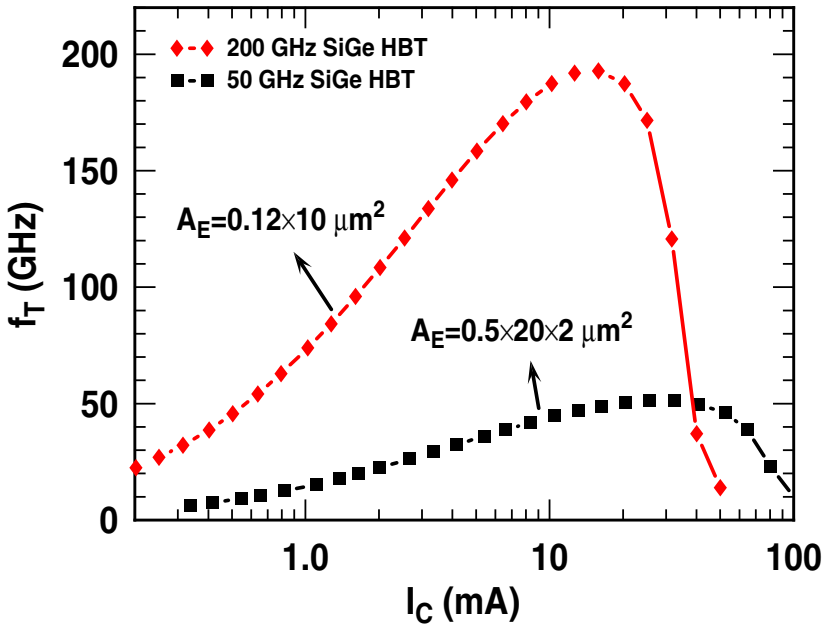


Figure 4.11: f_T versus I_C for a 50 GHz peak f_T device and a 200 GHz peak f_T device.

Fig. 4.12 shows the power gain versus frequency for both SiGe HBTs. As shown in the figure, the power gain at an input power of -5 dBm keeps dropping from 1 GHz to 21 GHz. The gain decrease for the 50 GHz peak f_T device is nearly 15 dB and 20 dB from 1 GHz to 21 GHz when biased at V_{CE} = 1.1 V and V_{CE} = 3 V. While the 200 GHz peak f_T device shows an incredibly stable performance with less than a 3 dB gain decrease over this very large frequency range. This is impressive, and demonstrates the superior high frequency large signal performance of the 200 GHz HBTs.

4.2.4 Impact of Bias Conditions on PAE

PAE is one of the figure of merits for amplifiers. Depending on the input power range, PAE requirements for different applications could change on where should the peak PAE locate.

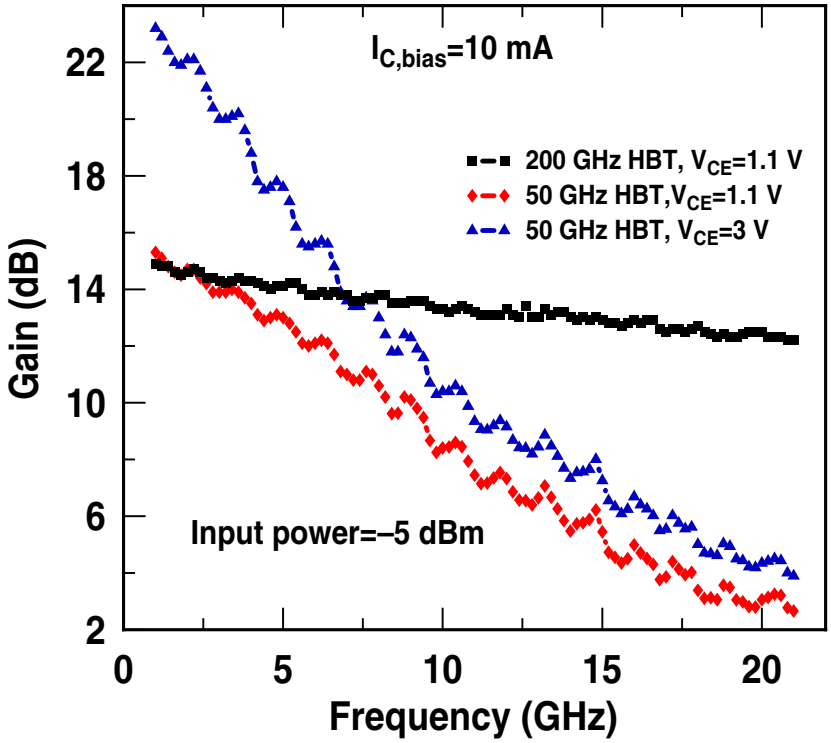


Figure 4.12: Gain versus frequency at input power of -5 dBm for a 50 GHz peak f_T device and a 200 GHz peak f_T device.

Fig. 4.13 shows Gain, P_{out} and PAE versus P_{in} and supply voltage at 2 GHz for a Si 0.5 × 20 × 2 μm^2 50 GHz device, $I_{C,bias}$=8 mA. In the small signal range, output power and gain are almost the same for all bias voltages, showing that the bias voltages have little impact on small signal performance. At input powers over -30 dBm, however, the difference is quite apparent. As shown in Fig. 4.13, a higher bias voltage helps increase the maximum output power and 1 dB compression point, and shifts the peak PAE into a higher input power range. This indicates that for low input power applications, the bias voltage should be chosen to be low in order to get high efficiency for the system, and vice versa for high input power applications.

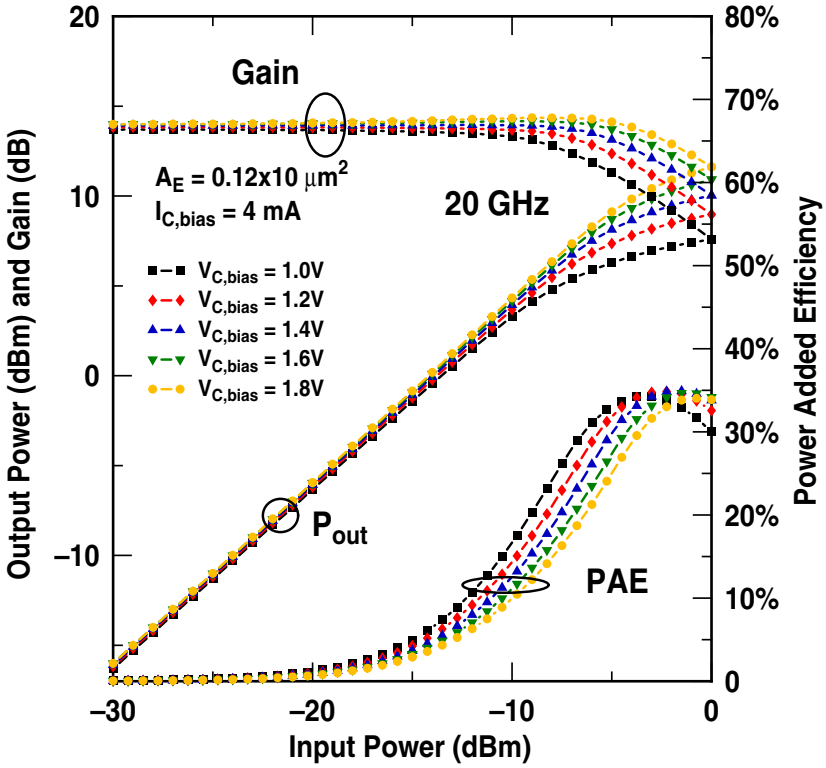


Figure 4.13: Gain, P_{out} and PAE versus P_{in} and supply voltage at 2 GHz for a Si 0.5 × 20 × 2 μm^2 50 GHz device.

The impact of bias current on small signal power performance of HBTs has been characterized by f_T and f_{max}. Fig. 4.14 shows Gain, P_{out} and PAE versus P_{in} and bias current at 2 GHz for a Si 0.5 × 20 × 2 μm^2 50 GHz device, $V_{C,bias}$ = 3V. As predicted by the trend in f_T and f_{max}, small signal gain increases with bias current and is expected to decrease eventually if the bias current is larger than the peak f_T f_{max} current. In the high input power range, however, the output powers for different bias currents merge with each other and gain and output power are almost the same for 0 dBm input power. This is because the DC currents are pushed up by the nonlinearity of the device at high input powers, and the difference in DC currents are smaller at large input powers as shown in Fig. 4.15. With similar large signal performance and lower

DC current consumption for lower bias currents, the PAE of the lowest bias current is thus the highest. This indicates that for applications where the major concern is PAE at high input power, a low bias current is preferable.

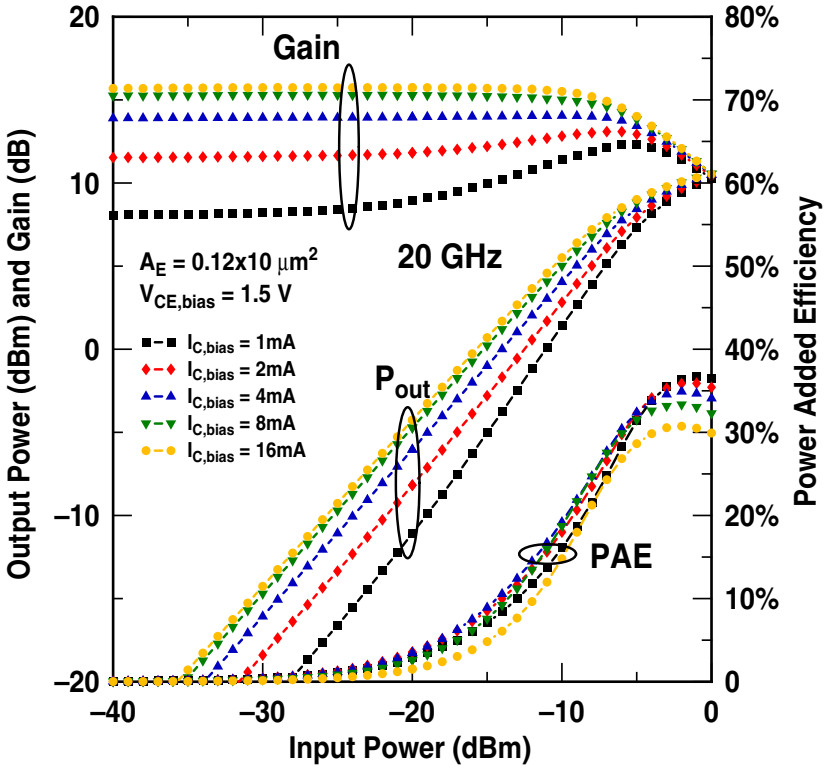


Figure 4.14: Gain, P_{out} and PAE versus P_{in} and bias current at 2 GHz for a Si 0.5 × 20 × 2 μm^2 50 GHz device.

4.3 Conclusions

This chapter reports an experimental investigation of SiGe profile and collector profile optimization from a large signal performance standpoint, as well as the impact of technology scaling. The results show that device and circuit designs that only consider the optimum small

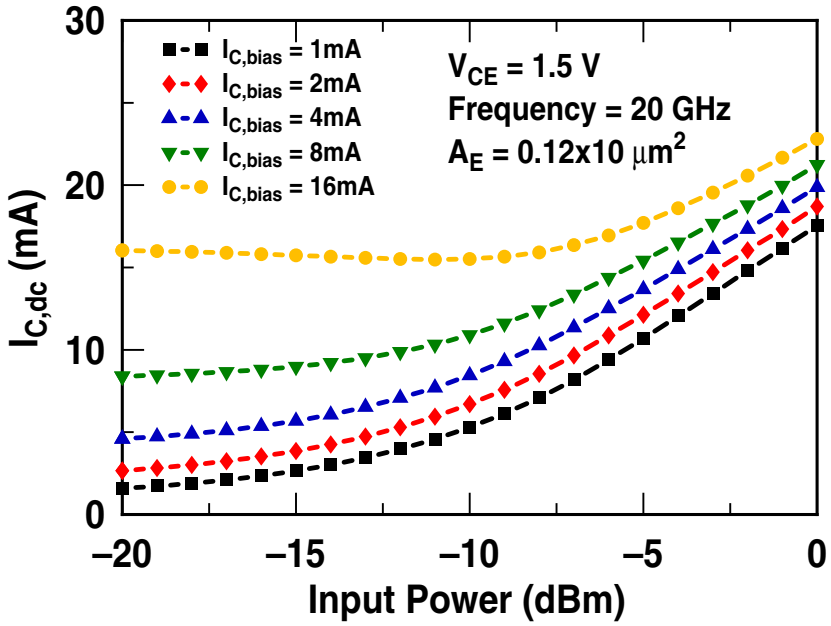


Figure 4.15: f_T versus I_C for a 50 GHz peak f_T device and a 200 GHz peak f_T device.

signal performance may inadvertently degrade large signal performance. The tradeoffs in SiGe profile design between small signal and large signal performance, as well as the impact of speed-breakdown tradeoff on large signal performance, were experimentally examined. The SiGe HBTs from a 200 GHz technology showed impressive small and large signal performance at 20 GHz, demonstrating the benefits of technology scaling, despite the accompanying breakdown voltage decrease.

CHAPTER 5

THIRD ORDER INTERMODULATION

Intermodulation linearity is an important figure-of-merit for RF devices, as it relates to the selectivity of an RF receiver and the spectral purity of an RF transmitter. Various theories, simulations, and experimental investigations of linearity have been reported for Si, SiGe and III-V bipolar transistors [51]- [56]. Accurate simulation and modeling of linearity is challenging, as it requires an accurate description of all current and charge nonlinearities in the device. This calls for systematic experimental characterization to better understand the intermodulation linearity behavior of bipolar transistors, which is practically nonexistent for SiGe HBTs.

The purpose of this chapter is to systematically characterize the intermodulation linearity of SiGe HBTs in order to gain insights into the device physics underlying their linearity behavior, and to provide guidelines for optimal sizing, biasing, and device selection, for instance, the choice of high breakdown versus low breakdown versions. The input 3rd order intercept point, IIP3, is used as a figure-of-merit for intermodulation linearity. IIP3 is measured on $I_C - V_{CE}$ plane for devices of various sizes, breakdown voltages, Ge profiles, and technology generations. The f_T is also measured, as in general f_T rolloff is a good indicator of the onset of high injection, and increase of collector charge storage. We will show that IIP3 rolls off as well at biasing currents near high injection f_T rolloff. Having the f_T data measured will also allow us to examine the detailed correlation between $f_T - I_C$ and IIP3 $-I_C$, as it was recently suggested that the IIP3 $-I_C$ curve is primarily determined by $f_T - I_C$ characteristics in [56]. We will show that this is not the case in SiGe HBTs. Later in this chapter, problems of VBIC model for simulating IIP3 are presented. Improvements for base collector capacitance and avalanche modelling in the

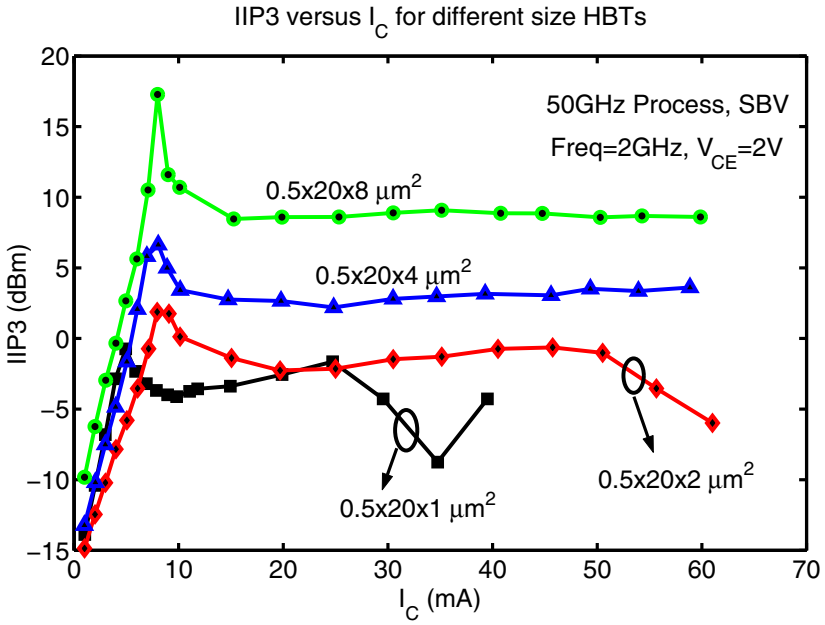
VBIC model are suggested and implemented in Verilog-A in order to provide a much better fit to the measurement results.

5.1 Experimental Results for IIP3

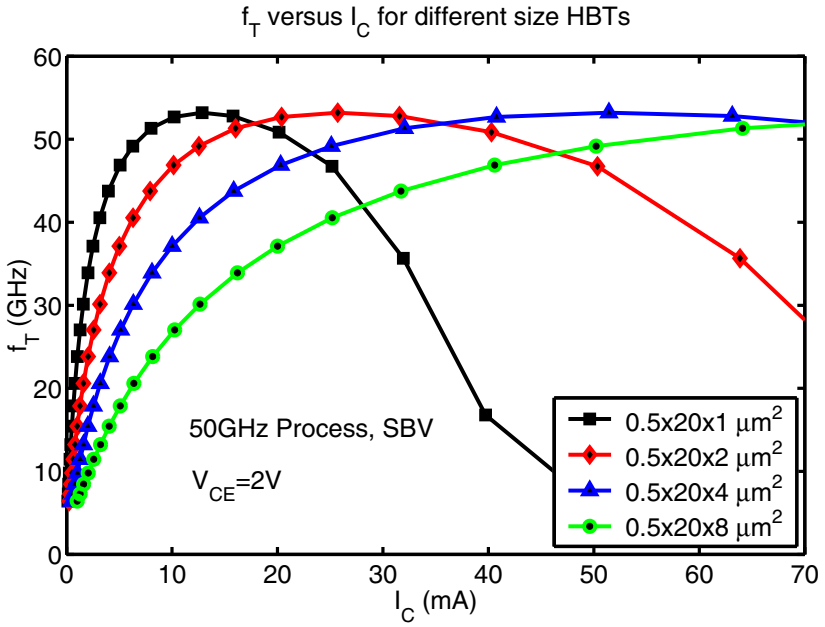
5.1.1 Current and Size Dependence

Fig. 5.1 (a) shows the measured 2 GHz IIP3 versus I_C for 50 GHz HBTs of various sizes. V_{CE} = 2.0 V. The corresponding $f_T - I_C$ data are shown in Fig. 5.1(b). The s-parameters are measured using an HP8510C Vector Network Analyzer (VNA), from which f_T is extracted from h_{21}. Note that IIP3 first increases with I_C, reaches a peak, and then drops to an almost constant level. For 2 finger, 4 finger and 8 finger HBTs, the IIP3 $-I_C$ curves have similar shapes. The peak IIP3 occurs at about the same I_C, which is well below their respective peak f_T I_C's. This indicates that the peak IIP3 current for these large HBTs is mostly determined by the factors relating to I_C, as opposed to derivatives of f_T with respect to I_C suggested in [56]. Clearly, the derivatives of f_T with respect to I_C are different for different device sizes, as can be seen from Fig. 5.1(b). If the IIP3 versus I_C behavior is governed by the $f_T - I_C$ behavior, the peak IIP3 would occur at I_C's corresponding to the same current density J_C, which would then increase with the number of fingers. The results here clearly indicate that the IIP3 behavior of bipolar transistors cannot be simply described using the $f_T - I_C$ behavior, and good modeling of $f_T - I_C$ does not necessarily guarantee good IIP3 modeling.

The exact mechanisms responsible for the similar IIP3 $-I_C$ behavior for different device sizes need further investigation. In general, a peak of IIP3 is a result of complex cancellation between individual physical nonlinearities [52] [53]. A larger device has larger depletion capacitances but smaller series resistances. A larger C_{CB} for instance, helps IIP3 as a linear feedback



(a)



(b)

Figure 5.1: (a) IIP3 versus I_C; (b) f_T versus I_C at $V_{CE} = 2.0$ V for different size SiGe HBTs in a 50 GHz technology.

element, but may degrade IIP3 as the C_{CB} nonlinearity is stronger. We have also applied the IIP3 equations in [54], and found that they cannot predict the same trend as measured.

The vertical shift in the IIP3 curves of the three larger devices is likely due to the varying degree of mismatch to the 50 Ω source, which will lead to a smaller internal v_{be} and hence higher IIP3 for the larger device. This is also consistent with the decreasing gain with increasing device size (not shown). Fig. 5.2 shows the OIP3 (IIP3×gain) versus I_C for all four devices. The larger devices still show a higher OIP3.

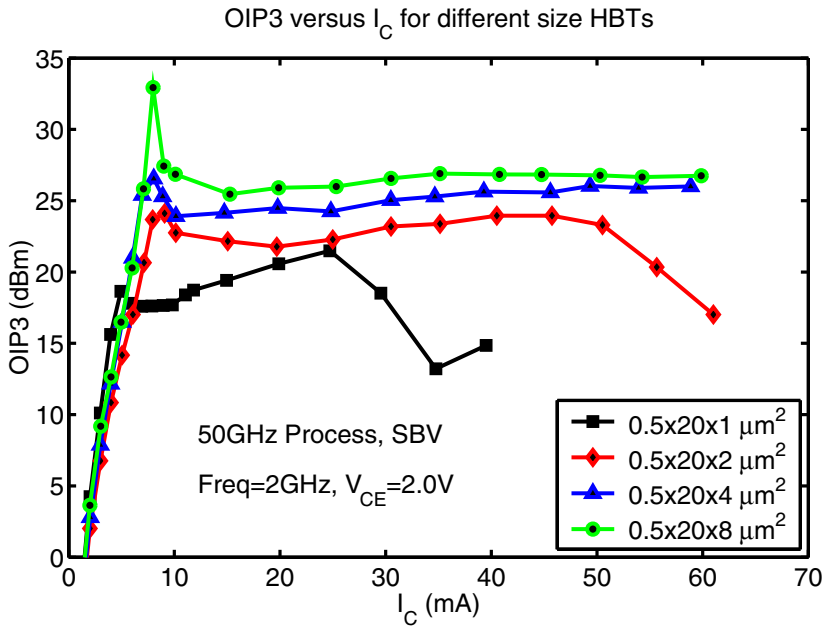


Figure 5.2: OIP3 versus collector current I_C at V_{CE} = 2.0 V for different size SiGe HBTs in a 50 GHz technology.

The peak IIP3, however, occurs at a smaller I_C for the smallest 1 finger HBT. IIP3 starts to drop with increasing I_C at 25 mA, where appreciable f_T rolloff occurs, as can be seen from Fig. 5.1(b). This IIP3 decrease is likely related to the increased collector charge storage due to

Kirk effect. The same effect can be observed for the 2 finger HBT, for which the IIP3 decreases at $I_C \approx 50$ mA, twice of the I_C at which IIP3 drops in the 1 finger HBT.

5.1.2 V_{CE} Dependence

Fig. 5.3 (a) and (b) show the measured 2 GHz IIP3 and f_T versus I_C at different V_{CE} for the 2 finger device. The emitter area is $0.5 \times 20 \times 2$ μm^2. A peak of IIP3 is observed for all V_{CE}'s. The peak IIP3 I_C and the peak IIP3 value both increase with V_{CE}. At higher current level (still well before high injection), IIP3 is nearly a constant for $V_{CE} < 3.0$ V. Due to the cancellations between nonlinearities and their current dependence, e.g. M-1 and CB charge storage, the change of IIP3 with V_{CE} is not monotonic, and depends on I_C, as shown in Fig. 5.4. At low I_C, IIP3 decreases with increasing V_{CE}, indicating that M-1 dominates. At a higher $I_C = 20$ mA, M-1 is reduced, IIP3 increases with V_{CE}, likely due to more linear C_{CB}. At $I_C = 50$ mA, the device is in high injection, and the change of IIP3 with V_{CE} is not monotonic.

5.1.3 Collector Profile and Breakdown Voltage

Typical SiGe processes offer devices with multiple collector profiles to provide devices with multiple breakdown voltages. High breakdown voltage (HBV) devices have a lower peak f_T than the standard breakdown voltage (SBV) high performance device, because of early onset of high injection due to lower collector doping. Fig. 5.5 (a) and (b) show the IIP3 and f_T comparison between the standard and high breakdown HBTs respectively. The 4 finger device with an emitter area of $0.5 \times 20 \times 4$ μm^2 is used. Other device sizes show similar trends. For a fair comparison, $V_{CE} = 2$ V and $V_{CE} = 4$ V are used for the HBV device. $V_{CE} = 2$ V for the SBV device. The BV_{CEO} is 3.3 and 5.1 V for the SBV and HBV devices.

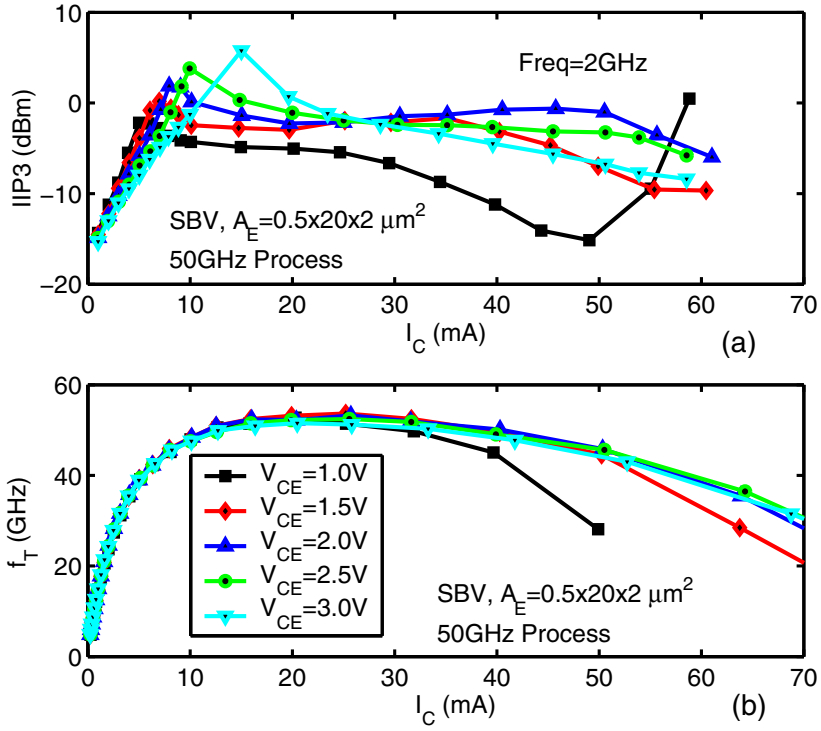


Figure 5.3: IIP3 and f_T versus I_C. 50 GHz process.

For the same $V_{CE} = 2.0$ V, IIP3 is nearly identical for the HBV and SBV devices before the IIP3 peak of the HBV device, which occurs at a smaller I_C. The peak IIP3 is higher in the SBV device. The IIP3 for the HBV device drops rapidly after the peak (at 8 mA) even if f_T continues to increase. In this case, the rolloff of f_T, however, does not occur by an appreciable amount until I_C is above 20 mA. Interestingly, the IIP3 for the HBV and SBV devices are identical below 8 mA, even though their f_T have shown appreciable differences near 8 mA. This also indicates that $f_T - I_C$ or $f_T - J_C$ characteristics does not determine IIP3.

The HBV device is intended for use with higher V_{CE}, which also helps suppressing high injection f_T rolloff. We thus also measured IIP3 at a higher V_{CE} of 4 V for the HBV device.

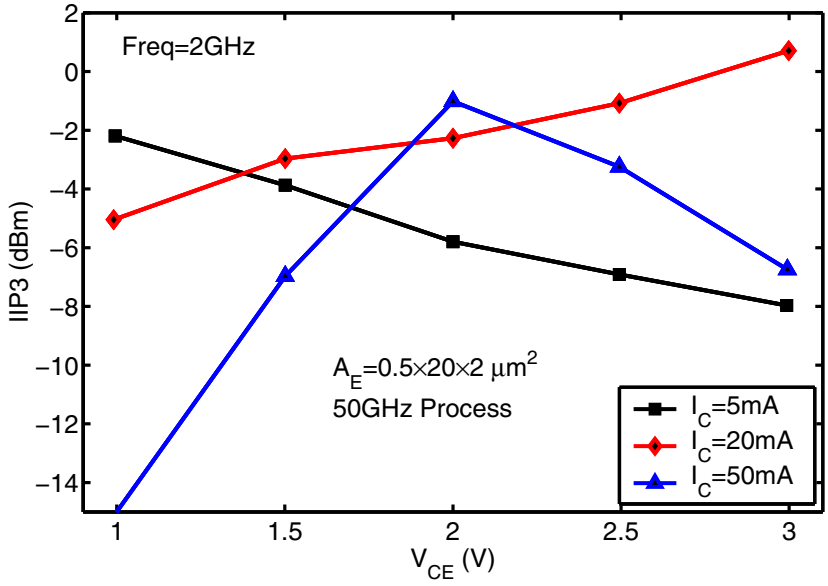


Figure 5.4: IIP3–V_{CE} at different I_C. 50 GHz process.

Compared to $V_{CE} = 2$ V, the IIP3 peak is now shifted to a higher I_C, and the peak value is higher than in the SBV device at $V_{CE} = 2$ V. For both V_{CE}, IIP3 rises with increase I_C after the decrease. Overall, the standard breakdown voltage device shows higher IIP3 and higher gain across a much wide biasing current range. An interesting observation is that even though f_T is identical for $V_{CE} = 2$ and 4 V below 13 mA, IIP3 is clearly different for $V_{CE} = 2$ and 4 V.

5.1.4 Ge Profile Dependence

To examine the impact of SiGe profile on IIP3 $-I_C$ characteristics, we measured the IIP3 and f_T of three SiGe profiles, including a 10% peak SiGe control, a 14% peak low-noise design (LN1), and a 18% peak low-noise design (LN2), with a Si BJT control. Details of the SiGe profile differences can be found in [40]. They all have identical SiGe film stability, but different shapes of Ge. LN1 and LN2 have more Ge content and higher Ge gradient in the neutral base,

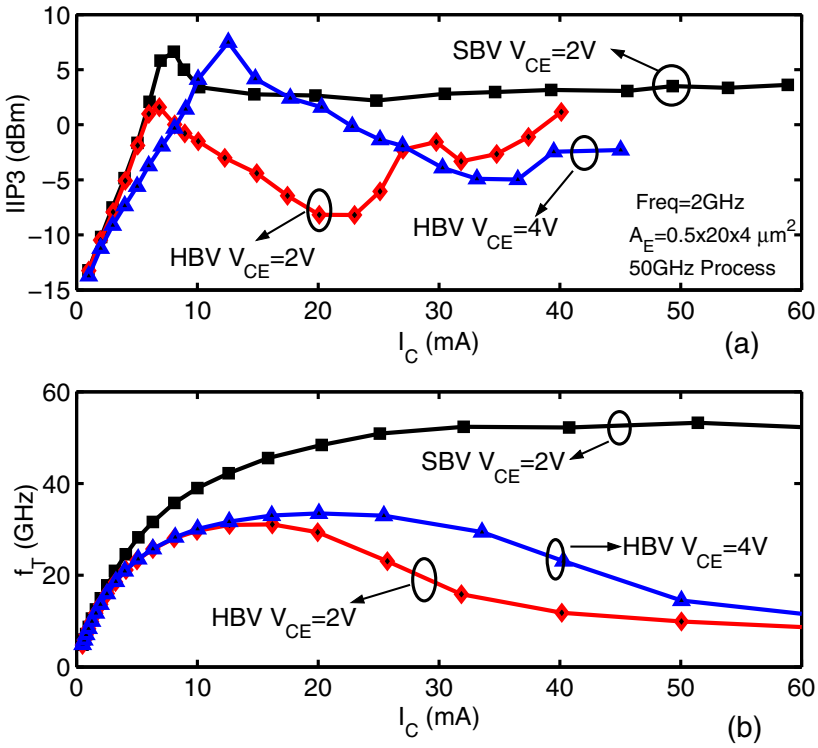


Figure 5.5: Comparison of IIP3 and f_T for standard and high breakdown HBTs. 50 GHz process.

but less retrograding into the collector, and consequently have much higher β, slightly higher f_T, but more rapid high injection f_T roll off.

Fig. 5.6 (a) and (b) show the IIP3 and f_T versus I_C at V_{CE} = 1.5 V. The 4 finger device is used here for illustration. Overall, the IIP3 is similar for all of the SiGe profile designs and the Si BJT control. As shown in [40], the optimum load varies with SiGe profile, and the peak IIP3 for optimum load is comparable for all profiles. The IIP3 of the LN1 and LN2 HBTs starts to drop at the same I_C of 40 mA, even though their f_T rolloffs occur at different I_C. The f_T rolls off at 40 mA in the LN1 profile, but at 50 mA in the LN2 profile. The IIP3 rolloff, however, does not occur in the 10% SiGe control and the Si control at I_C = 60 mA, despite visible f_T rolloff.

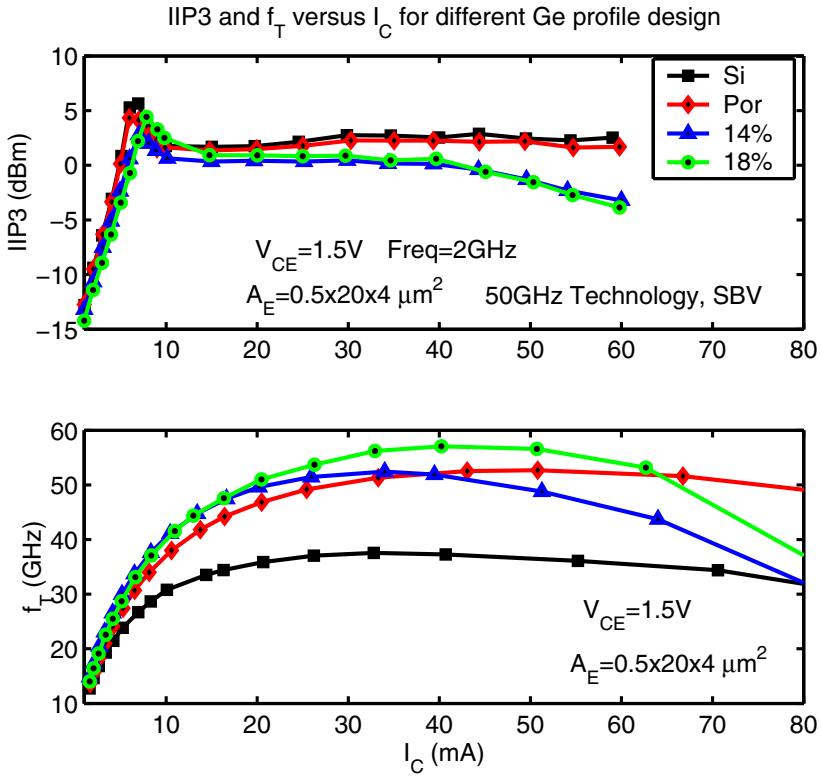


Figure 5.6: IIP3 and f_T for HBTs with different SiGe profiles. V_{CE}=1.5 V. 50 GHz process.

5.1.5 Technology Scaling

Fig. 5.7 (a) and (b) show the measured 5 GHz IIP3 and f_T of a 0.12×18 μm^2 HBT from a 200 GHz HBT process [57]. The IIP3-I_C behavior is qualitatively similar to that in the 50 GHz HBTs, but different in certain details. IIP3 reaches a peak at lower I_C, then drops to a relatively constant value. At higher I_C, IIP3 drops rapidly at a I_C before the high injection f_T rolloff. The breadth of the flat IIP3 region does not coincide with the breadth of the peak (flat) f_T region. For instance, the IIP3 at $V_{CE} = 1$ V decreases rapidly as I_C increases from 20 to 30 mA, while f_T only decreases by a small amount. However, a wider flat IIP3 region qualitatively corresponds to

a wider peak f_T region. The IIP3 peak is less obvious than in the 50 GHz HBTs. The lower IIP3 value compared to the 50 GHz HBTs is in part due to the smaller device size used. Fig. 5.7 (c) shows the OIP3 versus I_C. The OIP3 is found to increase with device size, as was in the 50 GHz HBTs. For similar device size and same frequency, the IIP3 is lower, but gain is higher for the scaled 200 GHz process. The OIP3 is comparable for the 200 GHz and 50 GHz processes. The scaled HBTs are obviously capable of operating at much higher frequencies.

5.2 SiGe HBT Nonlinearities

In a typical SiGe HBT, there are numerous nonlinearities which need to be considered for third order intermodulation analysis. Fig. 5.8 shows a simplified large signal HBT model. None of the elements shown in the figure are linear except the source impedance R_S and the load impedance R_L. The lumped resistors R_b, R_e, and R_c are functions of the bias due to either current-crowding effect or epi-layer modulation. The base current I_b, base emitter capacitor C_{be}, collector current I_c, collector base capacitor C_{cb} and avalanche current I_{ave} are all functions of bias. Normally in a compact model like VBIC, the nonlinear equations are incorporated inside the model. In the subsections below, the nonlinear C_{CB} and avalanche current are discussed as well as the equations for implementing them in VBIC. Some physically true relations about C_{CB} and avalanche are not implemented in these equations, which are shown later to be the reason why the IIP3 simulation results deviate from the experimental results. These relations are also discussed in the subsections below.

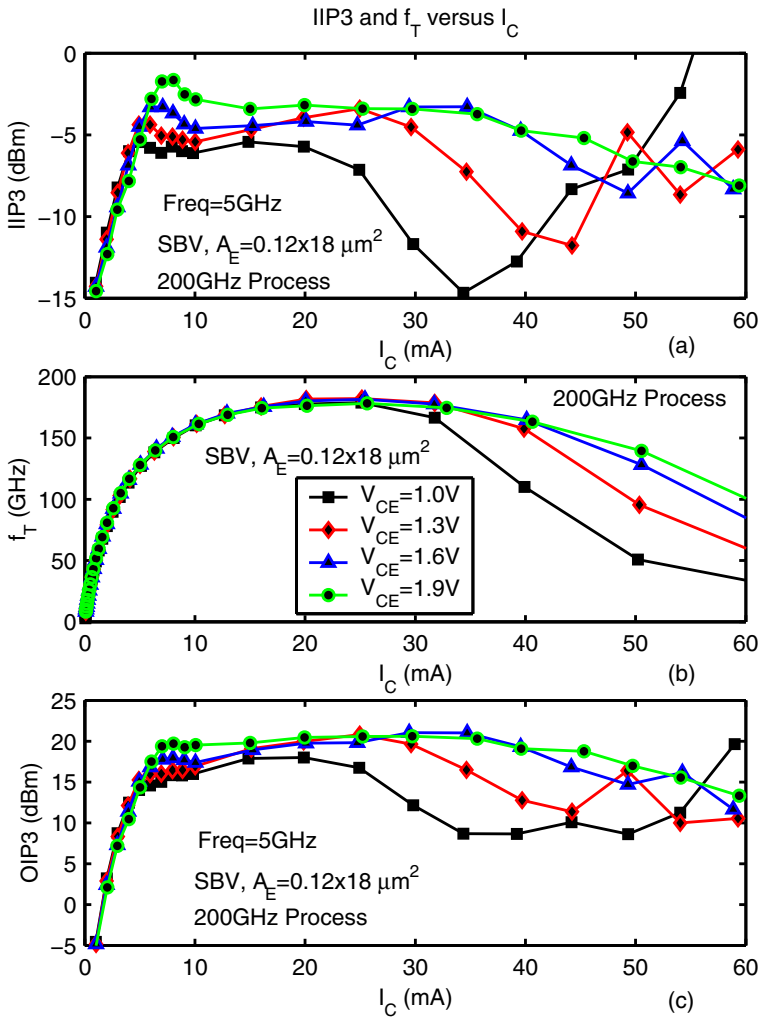


Figure 5.7: IIP3 and f_T of a $0.12 \times 18\mu\text{m}^2$ HBT from a 200 GHz process.

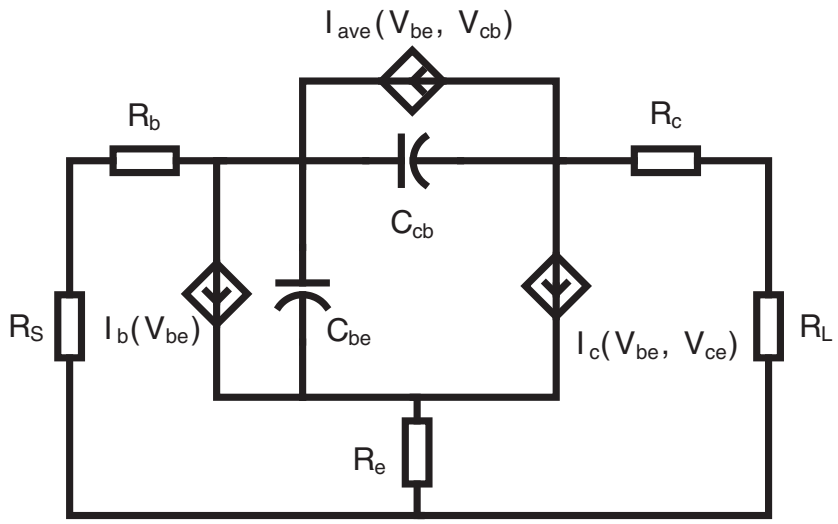


Figure 5.8: A simplified large signal HBT model.

5.2.1 Collector Base Capacitor

A collector base junction is usually reverse biased. Thus, in most cases the collector base capacitor is a depletion capacitor. The equation for this depletion capacitor is

$$C_t(V_f) = \frac{C_0}{\left(1 - \frac{V_f}{V_j}\right)^{m_j}} \tag{5.1}$$

where C_0, V_j, and m_j are known model parameters. For a flat collector doping profile, m_j is around 1/2. For a graded doping profile as used in many practical diffused junctions, m_j is around 1/3.

In the VBIC model, C_{CB} includes both the internal capacitance and the external parasitic capacitance. Here we discuss only the internal capacitance. The internal C_{CB} is described in VBIC through the CB junction charge Q_{CB}, which is composed of three parts. The part related

to C_{CB} is implemented as [58]

$$Q_{jc} = C_0 V_j \frac{1 - \left(1 - \frac{V_f}{V_j}\right)^{1-m_j}}{1 - m_j} \tag{5.2}$$

This Q_{jc} is a direct integration of the depletion capacitor, as shown in Eq. 5.1.

In an HBT, the collector base charge is not only a function of CB voltage but also a function of collector current. In [59], a Q_{jc} model is proposed for an one-sided step doping junction:

$$Q_{jc} = C_0 V_j \frac{1 - \left[\left(1 - \frac{V_f}{V_j}\right)\left(1 - \frac{I_C}{qcc \cdot ICRIT0}\right)\right]^{1-m_j}}{1 - m_j} \tag{5.3}$$

where qcc is a correction to the electron velocity and the physical meaning of $ICRIT0$ is given by $ICRIT0 = qA_e N_D v_{sat}$, where A_e is the junction area, N_D is the doping concentration, and v_{sat} is the saturation velocity. The depletion capacitance is now an equation of both V_f and I_C,

$$C_t = \frac{\partial Q_{jc}}{\partial V_f} = \left[\left(1 - \frac{V_f}{V_j}\right)\left(1 - \frac{I_C}{qcc \cdot ICRIT0}\right)\right]^{-m_j} \tag{5.4}$$

For a linearly graded doping such as the one shown in Fig. 5.9 (a), assuming $N_D = bx$, the depletion width is [60]

$$W = \left[\frac{3K_S\varepsilon_0}{qb}\left(V_{bi} - V_f\right)\right]^{1/(m+2)} \tag{5.5}$$

Applying a finite current I_C to a junction with the same profile, the electron density will turn part of the positive charge region into a negative charge region, as shown in Fig. 5.9 (b). The width W_1 of the converted part is given by

$$W_1 = \frac{I_C}{qA_e v_{sat} b} \tag{5.6}$$

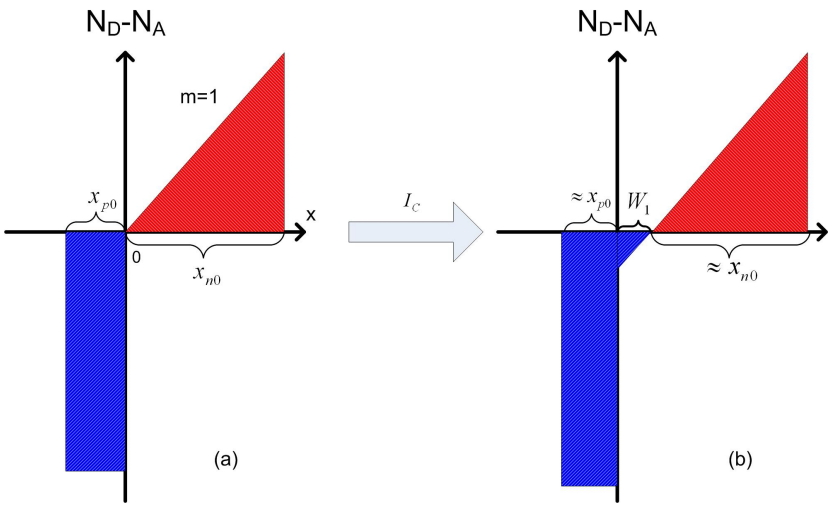


Figure 5.9: Examples of one-sided linearly graded doping profiles: (a)no current, (b)finite current.

To maintain the same potential difference, the width of the rectangle base side space charge region x_{p0} and the triangle collector side space charge region x_{n0} must decrease. Since the doping level in the left side is very high, the decrease in the left edge is very small compared to W_1 caused by a relatively high I_C. On the right side, the doping is much higher at the edge than the converted part. So the right side width is around $x_{n0} + W_1$. This a good estimate especially at high reverse voltage. Thus a good approximation of the total width change due to I_C is W_1. Since the depletion capacitance can be calculated as

$$C_J = \frac{K_S \varepsilon_0 A}{W} \tag{5.7}$$

where $K_S\varepsilon_0$ is the dielectric permittivity of the material, A is the area, and W is the depletion region width. A simple model for current dependent C_J can be derived as

$$C_J = \frac{K_S\varepsilon_0 A}{W + W_1} \tag{5.8}$$

To implement this model into VBIC, C_0 in Eq. 5.2 is changed to be current dependent according to Eq. 5.8. Fig. 5.10 shows the original VBIC CJC and the modeled CJC versus I_C. The VBIC CJC increases with I_C due to self-heating, as CJC increases with temperature. The new model models the current dependent effect. The result is similar to what was reported in Fig.3 of [59]. Notice although the CJC changes about one third, the total C_{CB} does not change as much because total C_{CB} includes the the parasitic capacitance which follows CJCP. As shown in Fig. 5.10, CJCP is not modified and still increases with I_C. Strictly speaking, a new model of Q_{jc} should be derived based on an integration over the whole region. However, since the doping in the real device is not strictly linearly graded, replacing C_0 with a current dependent one is a good approximation to the real case and, as shown later, serves well for IIP3 simulation.

5.2.2 Avalanche Current

As we noted earlier, avalanche current is not solely a function of V_{CB} but also of collector current. Under the assumption that avalanche is a linear relation to the peak electrical field, a new avalanche equation can be developed based on the empirical $M - 1$ equation

$$M = \frac{1}{1 - \left(\frac{V_{CB}}{BV_{CBO}}\right)^m} \tag{5.9}$$

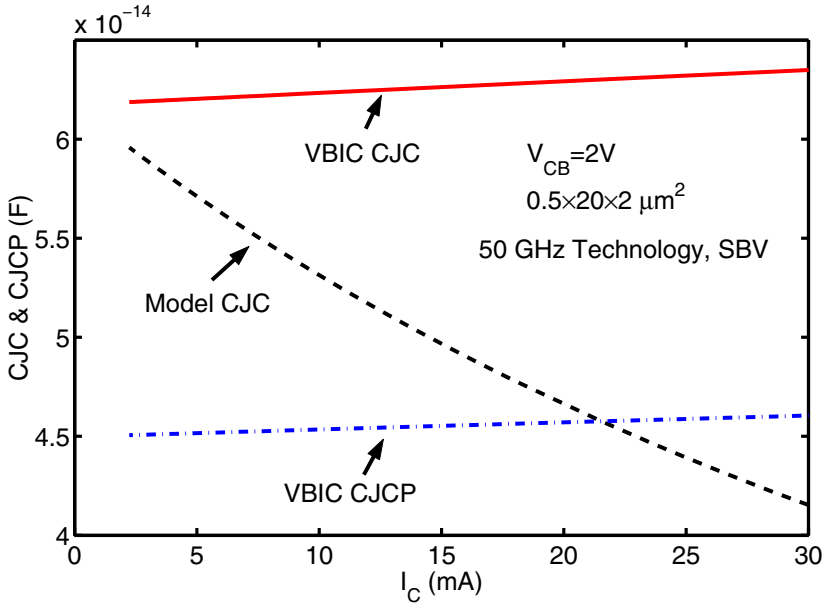


Figure 5.10: The VBIC CJC, the modeled CJC, and the VBIC CJCP versus I_C.

where V_{CB} is the applied CB voltage, and BV_{CBO} and m are fitting parameters. For a one-sided step CB junction, the peak electrical field under an external bias V_{CB} is given by [60]

$$|E_{max}| = \frac{qN_D}{K_S\varepsilon_0}x_n \tag{5.10}$$

where N_D is the doping concentration of the lightly doped side and

$$x_n = \left[\frac{2K_S\varepsilon_0}{qN_D}(V_{bi} + V_{CB})\right]^{1/2} \tag{5.11}$$

is the depletion width of the lightly doped side. Applying a finite current I_C to this junction, the effective doping will change to

$$N_{D,new} = N_D - \frac{I_C}{qA_e v_{sat} N_D} \tag{5.12}$$

Obviously the peak electrical field will also be decreased as a result of this change in the effective doping. The net change in the peak electrical field can be described by using a "current dependent" effective CB voltage $V_{CB,model}$

$$V_{CB,model} = V_{CB}\left(1 - \frac{I_C}{qA_e v_{sat} N_D}\right) - V_{bi}\frac{I_C}{qA_e v_{sat} N_D} \tag{5.13}$$

By replacing the V_{CB} in Eq. 5.11 with $V_{CB,model}$, the peak electrical field will be calculated to be the same while keeping the doping the same. Similarly, the M factor can be calculated by replacing the V_{CB} with $V_{CB,model}$.

Using a pulsed measurement system, the dependence of the avalanche multiplication factor on I_C can be characterized as in Fig. 2.17. Fig. 5.11 shows the measured $M-1$ versus J_C, VBIC simulated $M-1$ versus J_C, and the modeled $M-1$ versus J_C at an applied $V_{CB} = 3$ V. VBIC simulated $M-1$ cannot reproduce the $M-1$ dropping due to current dependence. Overall, the model does a good job. Notice in this model, we used a constant doping profile instead of a more realistic linear profile. The reason is that the model for linear profile is too complex and brings much trouble in integration into the Verilog-A model. As shown later in the chapter, the main point is to shown the trend of IIP3 improvement through modifying the avalanche model, not to accurately simulate IIP3. The simplified model shown here will serve well for presenting qualitatively the impact due to current dependent $M-1$.

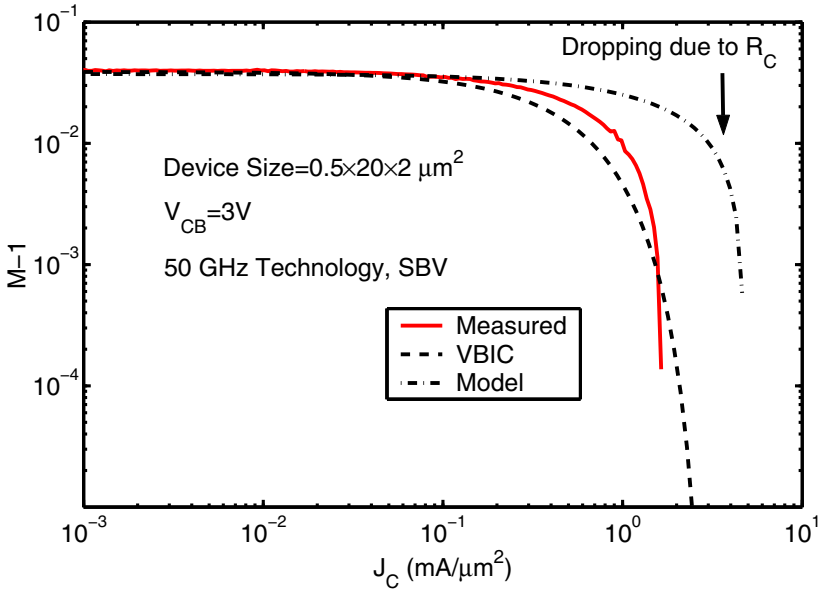


Figure 5.11: The measured $M-1$, the modeled $M-1$, and the VBIC simulated $M-1$ versus J_C.

5.3 Simulation of IIP3

Simulation of IIP3 for SiGe HBTs was performed in Advanced Design System (ADS) using compact models written in Verilog-A. Compact models are the set of mathematical equations that describe the performance of a device. In this work, VBIC was used for SiGe HBTs. Verilog-A is a procedural language, with constructs similar to C and other languages. With Verilog-A, it is possible to create and use modules that describe the high-level behavior of components and systems. The programmer provides the constitutive relationship of the inputs and outputs, the parameter names and ranges, while the Verilog-A compiler handles the necessary interactions between the model and the simulator–ADS [61]. Implementing the VBIC model in Verilog-A provided the freedom of adding the device physics that is not included in the original VBIC code. In the following subsections, simulations are done first using the original VBIC model to verify

that the VBIC model is capable of giving fairly good results for most DC and AC simulations on SiGe HBTs. Then, the simulated IIP3 is compared to the measurement results, which shows a large deviation. Finally, better simulation results of IIP3 are shown as a result of applying the improved model based on VBIC.

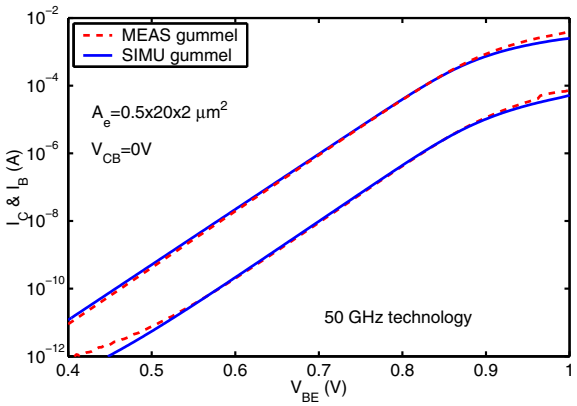
5.3.1 Performance of VBIC Model

Fig. 5.12 (a) shows the measured Gummel curves versus the simulation results using VBIC for a 0.5 × 2.5μm^2 50 GHz SiGe HBT. The Gummel curves show an excellent fit to the measurement results by simulation in the middle V_{BE} range, the most important range for RF circuit applications. Fig. 5.12 (b) shows the measured f_T versus I_C for a 0.5 × 20 × 2μm^2 device. The figure shows that the VBIC model is able to simulate the f_T very well. Besides these two figures, Fig. 4.2 in Chapter 4 also shows that VBIC has done a good job for the first order large signal fitting. Overall, VBIC is adequate for most of the DC, AC, and large signal simulation.

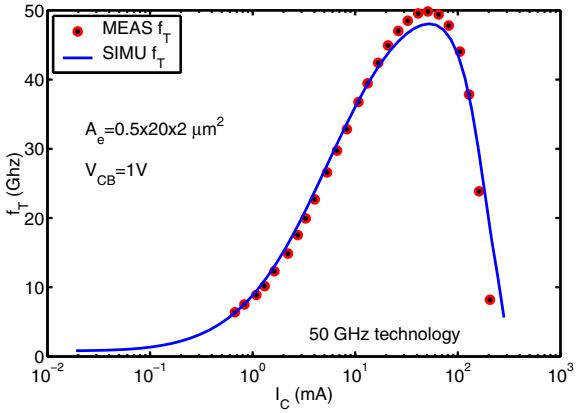
5.3.2 IIP3 Simulation Using VBIC and Verilog-A

Fig. 5.13 shows the simulated IIP3 using VBIC versus the measured IIP3 on a 0.5 × 20 × 2μm^2 SiGe HBT. Measurement and simulation were all performed at 2 GHz. As shown in Fig. 5.13, VBIC simulations on IIP3 do not fit the measurement data well. The shape of the measured IIP3 is always rising to a peak then falling and staying flat. The simulated IIP3 curves, however, show no obvious peaks and the curves are not flat at relatively high currents.

As stated in Section 5.1, the VBIC model does not take into account the current dependence of the collector base capacitor or the current dependence of avalanche. As shown in Fig. 5.8, both the CB capacitor and avalanche current are between the internal collector and base nodes, which is the most important feedback path. For a slight change in this path, the impact on the first



(a)



(b)

Figure 5.12: (a) Gummel curves; (b) f_T versus I_C at V_{CB} = 1.0 V for a 0.5 × 20 × 2 SiGe HBT in a 50 GHz technology.

order output might not be huge and can be compensated by adjusting the parameters. The impact on the third order output, however, could be enormous since the third order output is related to the second and third order derivatives of the nonlinear capacitance and avalanche current.

Fig. 5.14 shows the simulated and measured IIP3 versus I_C for the 0.5 × 20 × 2μm^2 SiGe HBT at V_{CE} = 2.0 and 3.0 V. Fig. 5.14 (a) shows the measured IIP3. Fig. 5.14 (b) shows the simulated IIP3 in ADS using the original VBIC model and the same VBIC model implemented

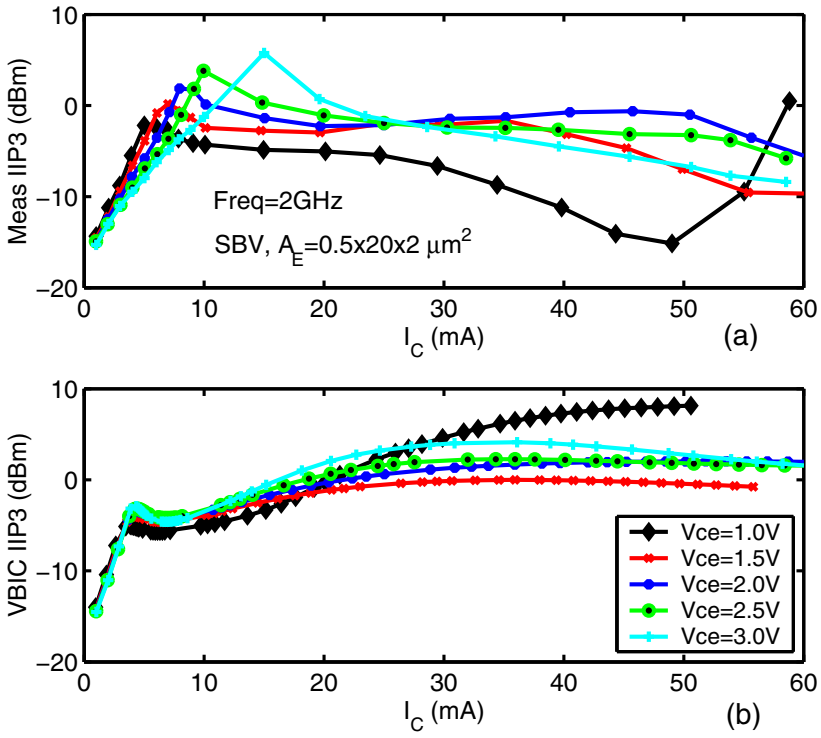


Figure 5.13: (a) Measured IIP3 versus I_C at different V_{CE}; (b) Simulated IIP3 using VBIC in ADS.

in Verilog-A. The difference in the original VBIC result and the Verilog-A VBIC result arises due to the different version of VBIC model used. The built in VBIC model in ADS is version 1.1.4, while the Verilog-A implemented one is version 1.2, which is not fully compatible with the 1.1.4 version. Although there is a value shift, the shape of the IIP3 is well preserved, which makes the Verilog-A module a good base for improvement. Fig. 5.14 (c) shows the simulated IIP3 using the Verilog-A VBIC model with the modification for C_{CB}. By introducing the current dependent C_{CB}, the shape of the simulated IIP3 is changing favorably toward the measured IIP3. The simulated IIP3 curve now has a very nice peak and remains flat at high currents, especially at $V_{CE} = 2$ V. However, with only C_{CB} modulation, the high IIP3 peak at V_{CE}= 3 V is

not reproduced. In Fig. 5.14 (d), current dependent avalanche is implemented in the Verilog-A code. The results shows that the avalanche at V_{CE} = 2 V does not affect the IIP3 result much, which is expected since the avalanche current is very low at that V_{CE}. At V_{CE} = 3.0 V, the current dependent avalanche makes a significant difference because of the high V_{CE} bias. At high current, neither model can simulate the dropping IIP3 since in that region Kirk effect or high injection barrier effect is the dominating factor.

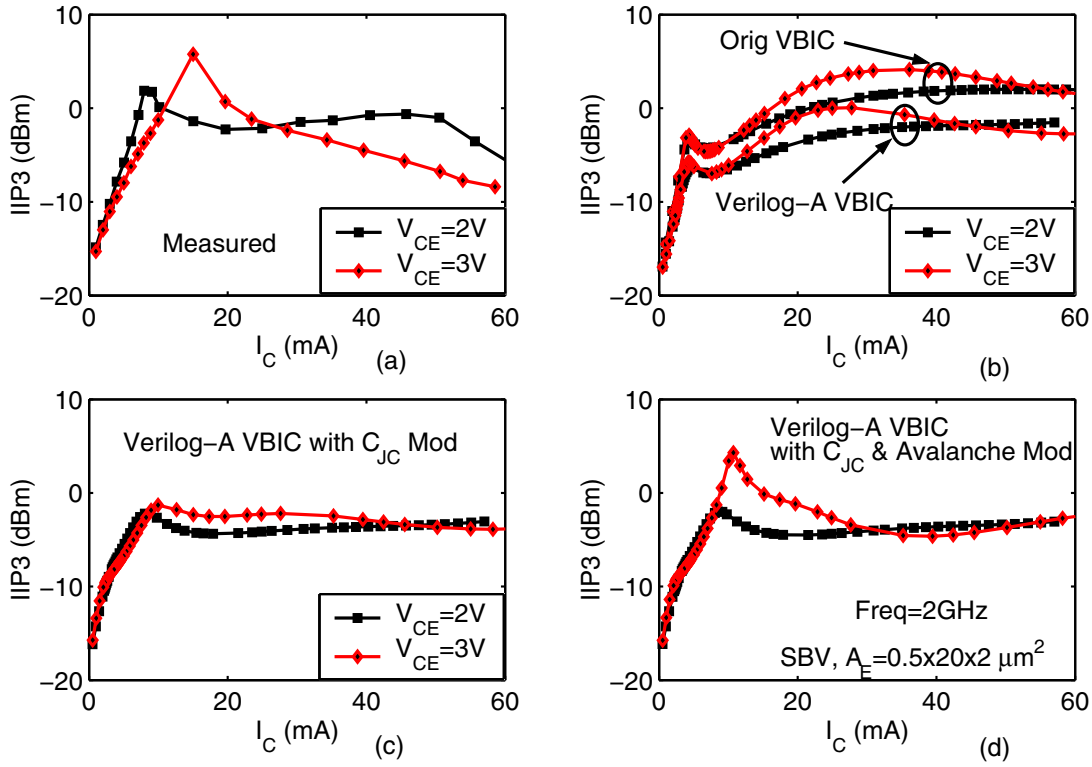


Figure 5.14: Simulated and measured IIP3 versus I_C for the 0.5 × 20 × 2μm^2 SiGe HBT at V_{CE} = 2.0 and 3.0 V.

Obviously, the modified VBIC model is not perfect. By introducing modified models for C_{CB} and avalanche, DC and AC simulation results will deviate from the original simulation

results. To achieve good simulations for DC, AC and IIP3 requires coordination of a lot of parameters and re-extraction of model parameter which is a hard job with our current measurement equipments. So the main point of applying improved C_{CB} and avalanche model into Verilog-A VBIC is to find the direction for improving IIP3 simulation.

5.4 Conclusions

In this chapter, the results of an experimental investigation of the intermodulation linearity of SiGe HBTs as a function of biasing current and voltage, device size, breakdown voltage, SiGe profile, and technology scaling were reported. The IIP3-I_C characteristics were compared with the f_T-I_C characteristics, and some correlation between high injection f_T rolloff and IIP3 rolloff was observed. The high performance standard breakdown HBT shows better IIP3 over a wider biasing current range than the high breakdown HBT. For the same size and frequency, scaled HBTs show lower IIP3 but higher gain, and comparable OIP3. Simulations were conducted using VBIC and Verilog-A modules based on VBIC. By introducing current dependence C_{CB} and avalanche to VBIC, simulated IIP3 data gained a much better fit to the measured IIP3 results.

CHAPTER 6

CONCLUSIONS AND FUTURE WORK

This work focused on the characterization and modeling of some important SiGe HBT parameters, with a discussion of the implications for analog circuit design. Chapter 2 gives a review of measurement methods for characterizing avalanche multiplication in SiGe HBTs. With increased device scaling, conventional methods fail at practical bias. New methods were proposed to accurately measure avalanche multiplication factor (M-1) even in the severe self-heating region. Current dependence of $M-1$ was demonstrated. The results show that the CB breakdown voltage at the J_E of peak f_T is higher than that at either low J_E or off state by a significant 1 V in a 120 GHz peak f_T device. Also in Chapter 2, the current dependence of $M-1$ was found to be considerably smaller by taking into account the extrinsic collector resistance. Later, in Chapter 5, a simplified model for the current dependent $M-1$ was proposed.

In Chapter 3, RF characterization methods are discussed, including S-parameters, large signal power characterization and third order intermodulation. In general, S-parameters characterize small signal parameters and can be used to extract base resistance R_B, cut-off frequency f_T and maximum oscillation frequency f_{max}. The large signal system built utilized the same equipment setup as for the S-parameter system, while being able to measure device performance at large input power. By monitoring the DC voltage and current, power added efficiency can be calculated. The third order intermodulation system built was more complex than the previous two in the sense that the system distortion level largely affects the accuracy of the measurement results. With careful setting, accurate IIP3 measurements can be performed for HBTs and MOSFETs. All three of the systems are controlled by in-house programs written in VEE. The program

for the S-parameter system was written by William E. Ansley, and modified here in order to increase the stability of operation and the application range of the program. Programs for the large signal system and the IIP3 system were written during the construction of both systems. All of the programs are now capable of measuring both HBT and MOSFET devices with high accuracy while requiring little attention from operators.

Chapter 4 reports experimental investigations of SiGe profile and collector profile optimization from a large signal performance standpoint, as well as the impact of technology scaling. The results show that device and circuit designs that only consider optimum small signal performance could inadvertently degrade large signal performance. The tradeoffs in SiGe profile design between small signal and large signal performance, as well as the impact of speed-breakdown tradeoff on large signal performance, were experimentally examined. The SiGe HBTs from a 200 GHz technology showed impressive small and large signal performance at 20 GHz, demonstrating the benefits of technology scaling, despite decreased breakdown voltage.

Chapter 5 presents a systematic characterization of the intermodulation linearity for SiGe HBTs to provide insights into the device physics underlying linearity behavior, and to generate guidelines to optimize sizing, biasing, and device selection (e.g. high breakdown versus low breakdown versions). The input 3rd order intercept point, IIP3, was measured on $I_C - V_{CE}$ plane for devices of various size, breakdown voltage, Ge profile, and technology generation. In the same chapter, problems with the VBIC model for simulating IIP3 were presented. Improvements for base collector capacitance and avalanche modelling in the VBIC model were suggested and implemented in Verilog-A to give a much better fit to the measurement results.

For future work, the device level characterization and modeling will continue. For example, a more practical model for $M - 1$ should be developed in order to provide a better fit to the experimental data; large signal performance of SiGe HBTs will be investigated within circuit

simulators using Verilog-A; and Mextram and HICUM models should be used for comparison with the VBIC model in the area of linearity simulation.

BIBLIOGRAPHY

[1] D.C. Ahlgren, M. Gilbert, D. Greenberg, J. Jeng, J. Malinowski, D. Nguyen-Ngoc, K. Schonenberg, K. Stein, R. Groves, K. Walter, G. Hueckel, D. Colavito, G. Freeman, D. Sunderland, D.L. Harame, B. Meyerson, "Manufacturability demonstration of an integrated SiGe HBT technology for the analog and wireless marketplace," *Tech. Dig. International Electron Devices Meeting*, pp. 859-862, 1996.

[2] J. Rieh, B. Jagannathan, H. Chen, K. Schonenberg, S. Jeng, M. Khater, D. Ahlgren, G. Freeman, and S. Subbanna, "Performance and design considerations for high speed SiGe HBTs of f_T/f_{max}=375 GHz/210 GHz," *Proc. IEEE Indium Phosphide and Related Materials Conference*, pp. 374-377, 2003.

[3] Z. Ma, S. Mohammadi, P. Bhattacharya, L. P. B. Katehi, S. A. Alterovitz, and G. E. Ponchak, "A high-power and high-gain X-band Si/SiGe/Si heterojunction bipolar transistor," *IEEE Tran. Microwave Theory and Techniques*, vol. 50, no. 4, pp. 1101-1108, April 2002.

[4] S. S. Iyer, G. L. Patton, S. S. Delage, S. Tiwari, and J. M. C. Stork, "SiliconCgermanium base heterojunction bipolar transistors by molecular beam epitaxy," *Int. Electron Device Meeting Tech. Dig.*, pp. 874-876, 1987.

[5] G. L. Patton, J. H. Comfort, B. S. Meyerson, E. F. Crabbe, G. J. Scilla, E. de Fresart, J. M. C. Stork, J. Y.-C. Sun, D. L. Harame, and J. N. Burghartz, "75-GHz f_T SiGe-base heterojunction bipolar transistors," *IEEE Electron Device Letters*, vol 11, no. 4, pp. 171-173, April 1990.

[6] D. L. Harame, J. H. Comfort, J. D. Cressler, E. F. Crabbé, J. Y.-C. Sun, B. S. Meyerson, and T. Tice, "Si/SiGe epitaxial-base transistors: part I-material, physics and circuits," *IEEE Transactions on Electron Devices*, vol. 40, no. 3, pp. 455-468, March 1995.

[7] B. S. Meyerson, "Low-temperature silicon epitaxy by ultrahigh vacuum/chemical vapor deposition," *Appl. Phys. Lett.*, vol. 48, pp. 797-799, 1986.

[8] R. Lodge, "Advantages of SiGe for GSM RF front ends," *Electron. Eng.*, vol. 71, no. 865, pp. 18-19, February 1999.

[9] R. Gotsfried, F. Beisswanger, S. Gerlach, A. Schuppen, H. Dietrich, U. Seiler, K. H. Bach, and J. Albers, "RFIC's for mobile communication systems using SiGe bipolar technology," *IEEE Transactions on Microwave Theory and Technology*, vol. 46, no. 5, pp. 661-668, May 1998.

[10] T. Ellermeyer, U. Langmann, B. Wedding, and W. Pholmann, "A 10Gb/s eye-opening monitor IC for decision-guided optimization of the frequency response of an optical receiver," *ISSCC Tech. Dig.*, pp. 50-51, 2000.

[11] G. Freeman, B. Jagannathan, S. Jeng, J. Rieh, A. D. Stricker, D. C. Ahlgren, and S. Subbanna, "Transistor design and application considerations for >200-GHz SiGe HBTs," *IEEE Transactions on Electron Devices*, vol. 50, no. 3, pp. 645-655, March 2003.

[12] G. Niu, J. D. Cressler, W. E. Ansley, C. S. Webster, R. B. Anna, and N. King, "Intermodulation characteristics of UHV/CVD SiGe HBTs," *Proc. IEEE Bipolar/BiCMOS Circuits and Technology Meeting*, pp. 50-53, 1999.

[13] D.L. Harame, D.C. Ahlgren, D.D. Coolbaugh, J.S. Dunn, G. Freeman, J.D. Gillis, R.A. Groves, F.N. Hendersen, R.A. Johnson, A.J. Joseph, S. Subbanna, A.M. Victor, K.M. Watson, C.S. Webster, P.J. Zampardi, "Current status and future trends of SiGe BiCMOS Technology," *IEEE Transactions on Electron Devices*, vol. 48, no. 11, pp. 2575-2594, November 2001.

[14] A. Schild, H.-M. Rein, J. Mullrich, L. Altenhain, J. Blank, and K. Schrodinger, "Amplifier array for 12 parallel 10 Gb/s optical-fiber links fabricated in a SiGe production technology," *Tech. Digest of IEEE Radio Frequency Integrated Circuits Symposium*, pp. 89-92, 2002.

[15] D. C. Ahlgren, G. Freeman, S. Subbanna, R. Groves, D. Greenberg, J. Malinowski, D. Nguyen-Ngoc, S. J. Jeng, K. Stein, K. Schonenberg, D. Kiesling, B. Martin, S. Wu, D. Harame, and B. Meyerson, "A SiGe HBT BiCMOS technology for mixed signal RF applications," *Proc. IEEE Bipolar / BiCMOS Circuits and Technology Meeting*, pp. 195-197, 1997

[16] A. Joseph, D. Coolbaugh, M. Zierak, R. Wuthrich, P. Geiss, Z. He, X. Liu, B. Orner, J. Johnson, G. Freeman, D. Ahlgren, B. Jagannathan, L. Lanzerotti, V. Ramachandran, J. Malinowski, H. Chen, J. Chu, P. Gray, R. Johnson, J. Dunn, S. Subbanna, K. Schonenberg, D. Harame, R. Groves, K. Waston, D. Jadus, M. Meghelli and A. Rylyakov, "A 0.18 μm BiCMOS technology featuring 120/100 GHz (f_T/f_{max}) HBT and ASIC-compatible CMOS using copper interconnect," *Proc. IEEE Bipolar / BiCMOS Circuits and Technology Meeting*, pp. 143-146, 2001.

[17] H. Ishiuchi, N. Tamba, J. D. Shott, C. J. Knorr, and S. S. Wong, "Hot-electron-induced minority-carrier generation in bipolar junction transistors," *IEEE Electron Device Letters*, vol. 11, pp. 490-492, Nov 1990.

[18] M. Rasras, I. De Wolf, G. Groeseneken, B. Kaczer, R. Degraeve, and H. E. Maes, "Photocarrier generation as the origin of Fowler-Nordheim-induced substrate hole current in thin oxides," *IEEE Trans. Electron Devices*, Vol. 48, pp. 231-238, Feb 2001.

[19] J. H. Klootwijk, J. W. Slotboom, M. S. Peter, V. Zieren, and D. B. de Mooy, "Photo carrier generation in bipolar transistors," *IEEE Trans. Electron Devices*, vol. 49, pp. 1628-1631, Sep 2002.

[20] P. F. Lu and T. C. Chen, "Collector-base junction avalanche effects in advanced double-poly self-aligned bipolar transistors," *IEEE Trans. Electron Devices*, vol. 36, pp. 1182-1188, June 1989.

[21] G. Niu, J. D. Cressler, S. Zhang, U. Gogineni, and D. C. Ahlgren, "Measurement of collector-base junction avalanche multiplication effects in advanced UHV/CVD SiGe HBT's," *IEEE Trans. Electron Devices*, vol. 46, pp. 1007-1015, May 1999.

[22] M. G. Adlerstein and M. P. Zaitlin, "Thermal resistance measurements for AlGaAs/GaAs heterojunction bipolar transistors," *IEEE Trans. Electron Devices*, vol. 38, pp. 1533-1554, 1991.

[23] D. C. Ahlgren et al., "Manufacturability demonstration of an integrated SiGe HBT technology for the analog and wireless marketplace," *Tech. Dig. Int. Electron Devices Meeting*, pp. 859-862, 1996.

[24] DIVA 225HR, Accent Optical Technologies, 2003.

[25] J. Scott, J. G. Rathmell, A. Parker and M. Sayed, "Pulsed device measurements and applications," *IEEE Tran. Microwave Theory and Techniques*, vol. 44, no. 12, pp. 2718-2723, Dec. 1996.

[26] J. D. Cressler and G. Niu, *Silicon-Germanium Heterojunction Bipolar Transistors,* Artech House, MA (2003).

[27] Operating and Programming Manual For the HP8510C Network Analyzer, Hewlett Packard, Inc., Santa Rosa, CA.

[28] J. Fitzpatrick, "Error models for systems measurement," *Microwave J.*, May 1978.

[29] R. Lane, "De-embedding device scattering parameters," *Microwave J.*, Aug. 1984.

[30] J. Staudinger, A Two-tier Method of De-Embedding Device Scattering Parameters Using Novel Techniques, Master Thesis, Arizona State University, May 1987.

[31] D. Metzger, "Improving TRL calibrations of vector network analyzers," *Microwave J.*, May 1995.

[32] M. Golio et al, *The RF and Microwave Handbook*, CRC Press, FL (2001).

[33] M. C. A. M. Koolen et al., "An improved de-embedding technique for on-wafer high-frequency characterization," *Proc. IEEE Bipolar/BiCMOS Circuits and Technology Meeting*, Sept. 1991, pp. 188C191.

[34] Q. Liang et al., "A simple four-port parasitic de-embedding methodology for high-frequency scattering parameter and noise characterization of SiGe HBTs," *IEEE Tran. Microwave Theory and Techniques*, vol. 51, no. 11, pp. 2165-2174, November 2003.

[35] P. Wambacq and Willy Sansen, *Distortion Analysis of Analog Integrated Circuits,* Kluwer Academic, MA (1998).

[36] *Operating and Programming Manual For the HP8563EC Spectrum Analyzer,* Hewlett Packard, Inc., Santa Rosa, CA.

[37] G. Niu, J. Pan, X. Wei, S. S. Taylor, and D. Sheridan, "Intermodulation linearity characteristics of CMOS transistors in a 0.12 μm process," Accepted by RFIC 2005.

[38] J. Pan, G. Niu, and D. C. Sheridan, "An experimental inverstigation of intermodulation linearity in SiGe HBTs," Submitted to BCTM 2005.

[39] W. E. Ansley, Microwave Characterization and Modeling of Silicon-Gemanium Heterojunction Bipolar Transistors, Ph.D dissertation, Auburn University, June 1998.

[40] G. Niu, S. Zhang, J. D. Cressler, A. J. Joseph, J. S. Fairbanks, L. E. Larson, C. S. Webster, W. E. Ansley, and D. L. Harame, "Noise modeling and SiGe profile design tradeoffs for RF applications," *IEEE Trans. Electron Devices*, vol. 47, pp. 2037-2044, 2000.

[41] G. Niu, Z. Jin, J. D. Cressler, R. Rapeta, A. J. Joseph, and D. L. Harame, "Transistor noise in SiGe HBT RF technology," *IEEE J. Solid-State Circuits*, vol. 47, pp. 2037-2044, 2000.

[42] J. Tang, G. Niu, Z. Jin, J. D. Cressler, S. Zhang, A. J. Joseph, and D. L. Harame, "Modeling and characterization of SiGe HBT low-frequency noise figures-of-merit for RFIC applications," *IEEE Trans. Microwave Theory and Tech.*, vol. 50, pp. 2467-2473, 2002.

[43] B. G. Malm and M. Ostling, "Ge-profile design for improved linearity of SiGe double HBTs," *IEEE Electron Device Letters*, vol. 23, pp. 19-21, 2002.

[44] *8510C Network Analyzer System Operating and Programming Manual*, Agilent Technologies, 2001.

[45] G. Niu, J. D. Cressler, S. Zhang, U. Gogineni, D. C. Ahlgren, "Measurement of collector-base junction avalanche multiplication effects in advanced UHV/CVD SiGe HBT's," *IEEE Trans. Electron Devices*, vol.46, pp. 1007-1015, May 1999.

[46] A. J. Joseph, J. D. Cressler, D. M. Richey and G. F. Niu, "Optimization of SiGe HBT's for operation at high current densities," *IEEE Trans. Elect. Dev.*, vol. 47, pp. 2037-2044, 2000.

[47] G. F. Niu, and J. D. Cressler, " The impact of bandgap offset distribution between conduction and valence bands in Si-based graded bandgap HBT's," *Solid-State Elect.*, vol. 43, pp. 2225-2230, 1999.

[48] G. Niu, W. E. Ansley, S. Zhang, J. D. Cressler, C. S. Webster, and R. A. Groves, "Noise parameter optimization of UHV/CVD SiGe HBT's for RF and microwave applications," *IEEE Trans. Electron Devices*, vol. 46, pp. 1589-1598, 1999.

[49] X. Wang, D. Wang, C. Masse, P. Bacon, "Low phase noise SiGe voltage-controlled oscillators for wireless applications," *Microwave Journal*, vol. 45, pp. 84-99, February 2002.

[50] D. A. Hitko, *Circuit Design and Technological Limitations of Silicon RFICs for Wireless Applications*, Ph.D thesis, MIT, 2002.

[51] J. Reynolds, "Nonlinear distortions and their cancellation in transistors," *IEEE Trans. On Electronic Devices*, no. 11, pp. 595-599, 1965

[52] S. A. Maas, B. L. Nelson, and D. L. Tait, "Intermodulation in Heterojunction Bipolar Transistors," *IEEE Trans. Microwave Theory and Tech.*, vol. 40, no. 3, pp. 442-448, 1992.

[53] G. Niu, Q. Liang, J. D. Cressler, C.S. Webster, and D. L. Harame, "RF Linearity Characteristics of SiGe HBTs," *IEEE Trans. Microwave Theory and Tech.*, vol. 49, no. 9, pp. 1558-1565, 2001.

[54] M.P. van der Heijden, H. C. de Graaff, and Leo C. N. de Vrreede, "A novel frequency-independent third-Order intermodulation distortion cancellation technique for BJT amplifier," *IEEE Journal of Solid State Circuits*, vol. 37, no. 9, pp. 1176-1183, 2002.

[55] M. Iwamoto, P. M. Asbeck, T. S. Low, C. P. Hutchinson, J. B. Scott, A. Cogntata, X. Qin, L. H. Camnitz and D. C. D'Avanzo, "Linearity characteristics of GaAs HBTs and the influence of collector design," *IEEE Trans. Microwave Theory and Tech.*, vol. 48, no. 12, pp. 2377-2388, 2000.

[56] M. Iwamoto, C. P. Hutchinson, J. B. Scott, T. S. Low, M. Vaidyanathan, P. M. Asbeck and D. C. D'Avanzo, "Optimum bias conditions for linear broad-band InGaP/GaAs HBT power amplifiers," *IEEE Trans. Microwave Theory and Tech.*, vol. 50, no. 12, pp. 2954-2962, 2002.

[57] 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 HBT with a non-self-aligned structure," *IEEE Electron Device Letters*, vol. 22, no. 11, pp. 542-544, 2001.

[58] *Spectre Circuit Simulator Device Model Equations*, Product Version 5.0

[59] L. H. Camnitz, S. Kofol, T. Low, and S. R. Bahl, "An accurate, large Signal, high frequency model for GaAs HBTs," *Tech. Dig. of Gallium Arsenide Integrated Circuit (GaAs IC) Symposium*, pp. 303-306, Nov. 1996.

[60] R. F. Pierret, *Semiconductor Device Fundamentals*, Addison Wesley, MA (1996).

[61] *Verilog-A Reference Manual*, Agilent Technologies.

APPENDIX A

S-PARAMETER MEASUREMENT PROCEDURE

1. Turn on the VNA. Let the VNA warm up for at least 90 mins. The order of turning on is from bottom (the source) to the top (the processor).

2. Load the instrument state previously stored in the VNA. If the state is not stored in the VNA, load the instrument state first.

3. Perform a power flatness calibration on port 1 using the power meter. Notice the power flatness should be done when the VNA is measuring S12 or S11. The power meter should connect to the end of the cable where the probe is supposed to be attached.

4. Connect cables to the probes. Adjust probe station and positioner to make probes ready for the devices about to be measured.

5. Place the calibration substrate on the chuck and perform the OSLT full two port calibration.

6. Take the calibration substrate off the chuck and place the wafer on the chuck. Find the device and the corresponding open and short structure.

7. Open the program called "GetYopenZshort" to record the open and short structure's S-parameters. Fig. A.1 shows the panel of the "GetYopenZshort" program.

8. Probe the device needed to be measured. Open the program called "fTtestHBT". Input the device size, start stop current, DC bias voltage, open short data file locations, and data saving directory. Press the "Start" button, the program will control the VNA and DC

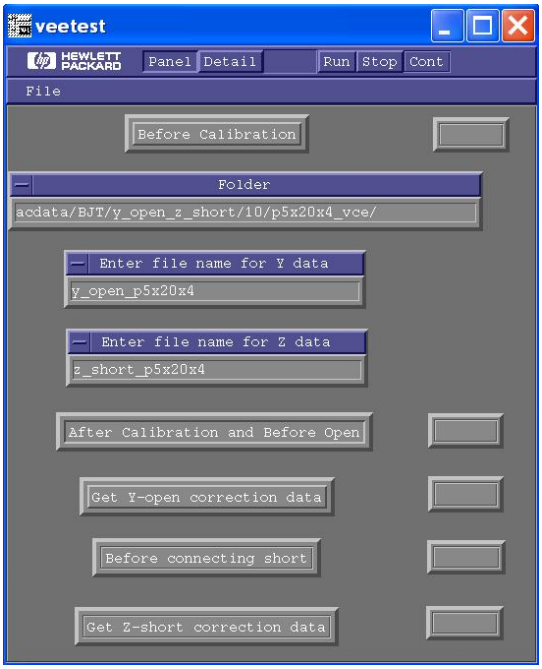


Figure A.1: Program panel for measuring open short structure and recording the results.

power supply to finish the measurement circle. During the measurement, f_T and H_{21} will be shown on the screen. Fig. A.2 shows the panel of the "fTtestHBT" program.

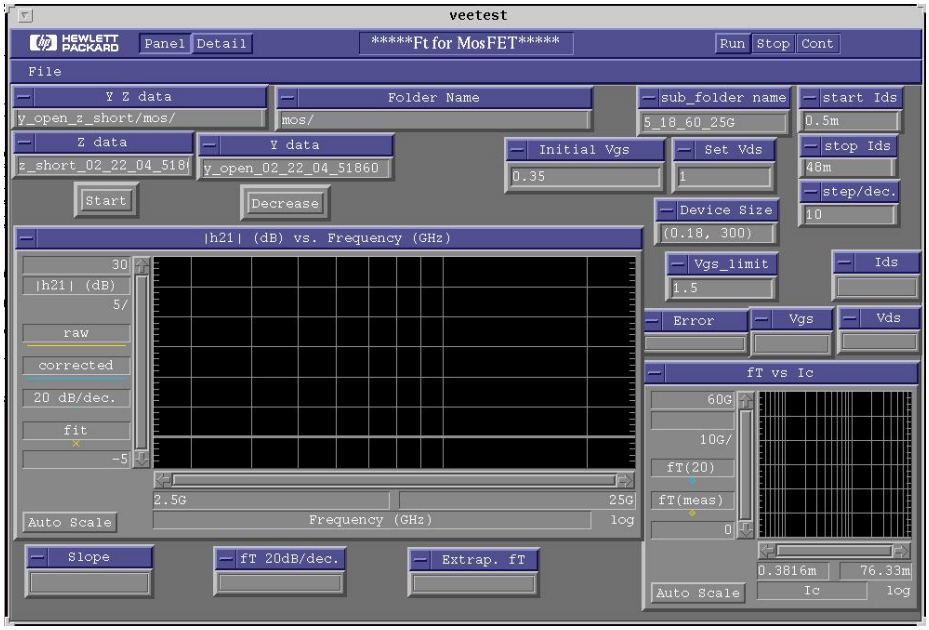


Figure A.2: Program panel for measuring and recording the S-parameters of the devices.

APPENDIX B

LARGE SIGNAL MEASUREMENT PROCEDURE

1. Turn on the VNA. Let the VNA warm up for at least 90 mins. The order of turning on is from bottom (the source) to the top (the processor).

2. Load the instrument state previously stored in the VNA. If the state is not stored in the VNA, load the instrument state first.

3. Perform a power flatness calibration on port 1 using the power meter. Notice the power flatness should be done when the VNA is measuring S12 or S11. The power meter should connect to the end of the cable, where the probe is supposed to be attached.

4. Connect cables to the probes. Adjust the probe station and positioner to make the probes ready for the devices about to be measured.

5. Place a calibration substrate on the chuck and perform the OSLT full two port calibration.

6. Perform the receiver calibration for power reading. After the calibration is done, change VNA to the power domain. The power domain is set at the frequency where the marker was in the frequency domain.

7. Open the VEE program “PoutPin”. Input the device size, start stop input power, DC bias voltage, and data saving directory. Press the “Start” button, program will control the VNA and DC power supply to finish the measurement circle. During the measurement, input power, output power and power added efficiency will be shown on the screen. Fig. B.1 shows the panel of the “PoutPin” program.

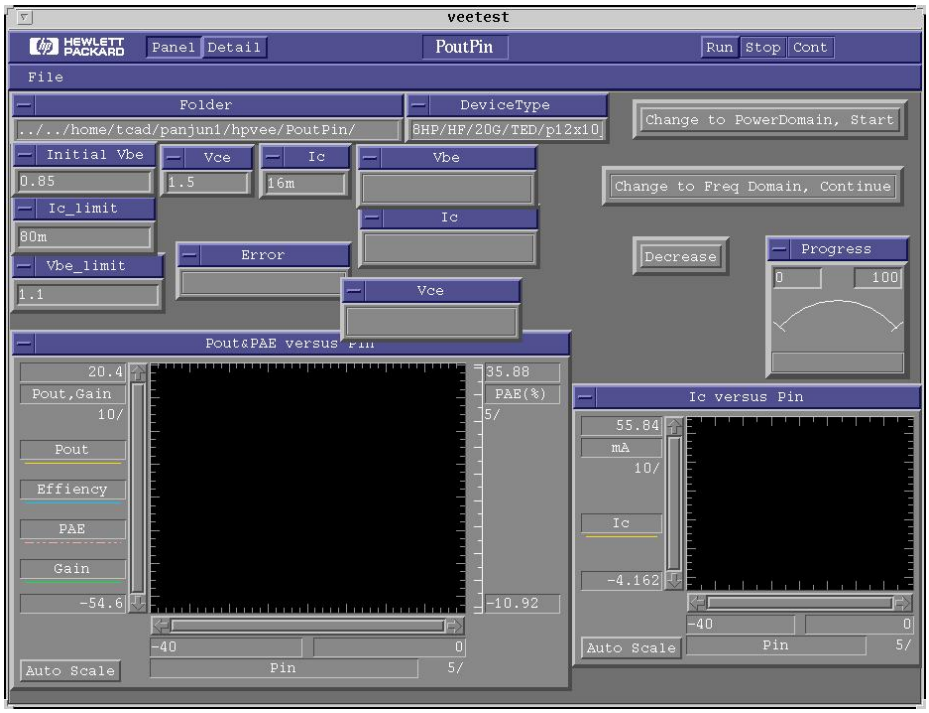


Figure B.1: Program panel for measuring and recording large signal performance of the devices.

APPENDIX C

INTERMODULATION MEASUREMENT PROCEDURE

1. Turn on Spectrum Analyzer, RF sources, DC power supply. Wait 30 mins for the system to warm up.

2. Turn the ALC function on RF sources off. Connect the power meter to the end of the cable where the probes are supposed to be attached. Get the actual power reading from the power meter to get the attenuation in the cable path.

3. Probe on a through. Connect the power meter to the end of the cable where the spectrum analyzer is supposed to be connected to. Measure the actual power output to calibrate the spectrum analyzer.

4. Probe the device. Open the program "IIP3test". Input the device size, two tone frequency, two tone spacing, power range, power attenuation, DC bias, and data saving directory. Press "Start". The program will run the measurement and record the fundamental tone power, third tone power, input power, and DC bias. Fig. C.1 shows the panel of the program.

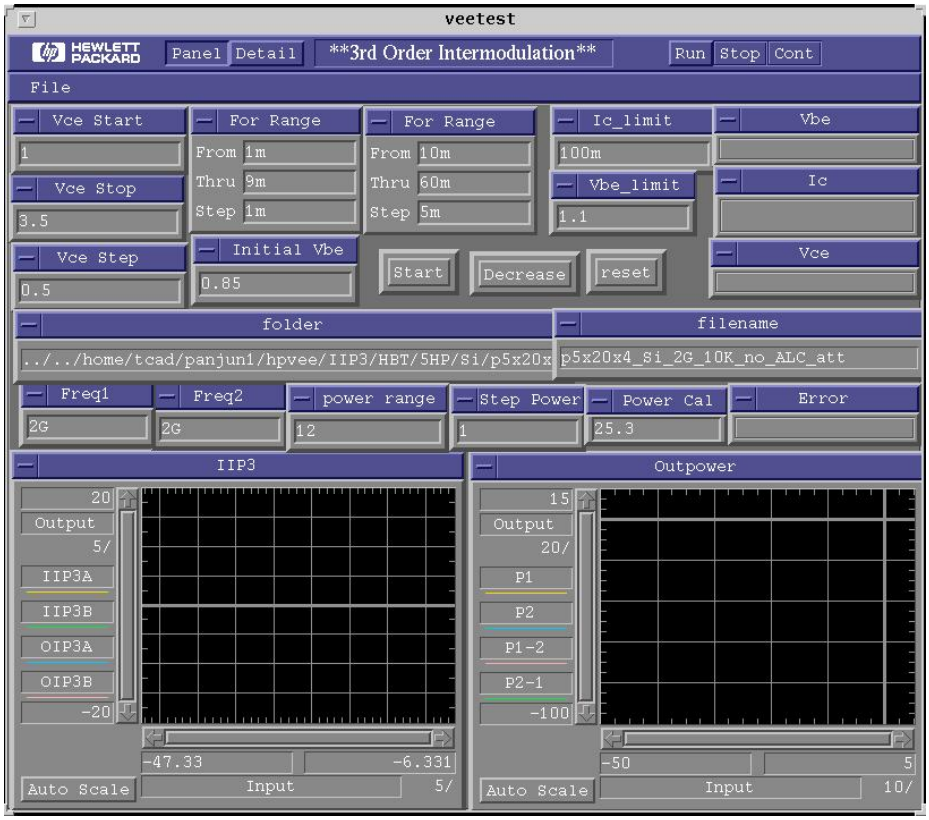


Figure C.1: Program panel for measuring and recording IIP3 of the devices.