Investigating high-frequency sige hbts : assessment of characterization and new architecture exploration

Soumya Panda

université de BORDEAUX

THÈSE PRÉSENTÉE
POUR OBTENIR LE GRADE DE

DOCTEUR DE L'UNIVERSITÉ DE BORDEAUX

ÉCOLE DOCTORALE DES SCIENCES PHYSIQUES ET DE L'INGÉNIEUR

SPÉCIALITÉ ÉLECTRONIQUE

présentée et soutenue publiquement par

Soumya Ranjan PANDA

le défense de thèse: 19th September, 2022

INVESTIGATING HIGH-FREQUENCY SiGe HBTs: ASSESSMENT OF CHARACTERIZATION AND NEW ARCHITECTURE EXPLORATION

Directeur(rice) de thèse : **Thomas ZIMMER & Anjan CHAKRAVORTY**

Membres du jury :

Prof. Philippe FERRARI,	Université de Grenoble-Alpes	Rapporteur
Prof. Nihar Ranjan MOHAPATRA,	IIT Gandhinagar	Rapporteur
Dr. Pascal CHEVALIER,	STMicroelectronics	Examinateur
Prof. Fabien PASCAL,	Université de Montpellier	Examinateur
Prof. David KOILPILLAI,	IIT Madras	Examinateur
Prof. Anjan CHAKRAVORTY,	IIT Madras	Co-directeur de thèse
Dr. Sébastien FREGONESE,	CNRS	Co-encadrant
Prof. Thomas ZIMMER,	Université de Bordeaux	Directeur de thèse

Université de Bordeaux
Laboratoire de l'Intégration du Matériau au Système
UMR CNRS 5218, F-33400 Talence, France

Abstract - EN

Keywords:High frequency characterization, Maximum oscillation frequency, Nanowire SiGe HBT, SOI SiGe HBT, Substrate bias effect.

Silicon germanium heterojunction bipolar transistors (SiGe HBTs) are rapidly evolving due to current communication systems'(4G, 5G & upcoming 6G) increased functionality and speed. Since the evaluation of SiGe BiCMOS technology, it has been serving the continuous demand for higher functionality and front-end performance quite well at low cost and medium to low volumes. As the operating frequency of SiGe HBTs exceeds 300 GHz, thus it allows critical circuits to operate beyond 100 GHz, which is called the lower end of the THz gap. The upper limit of the THz gap extends up to 30 THz. Within this THz gap range, a lot of applications are envisioned like (i) THz imaging and sensing (ii) Radar applications (iii) in measurement equipment like in ultra-high bandwidth analog to digital converters. With the emerging mm-wave and THz market in sight, the precise characterization and modeling of the devices in the sub-THz frequency range is compulsory to optimize the circuit performance and to minimize the number of design to fabrication loops. However, as we continue to develop devices with increased frequency performance, one of the major problems is accurate characterization at high frequencies (> 100 GHz).

In this work, firstly, a systematic method for verifying high-frequency measurement (up to 500 GHz) of SiGe HBT is proposed. The procedure entails a precise calculation of the passive environment's effect on the entire measurement via a comprehensive electromagnetic (EM) simulation. This ensures that the entire measuring environment is precisely incorporated into the framework for EM modeling. In order to additionally include the active SiGe HBT device, a technology computer aided design (TCAD) tool is used to simulate the device *S*-parameters. TCAD simulation results are fed into an EM-plus-SPICE simulation framework to emulate a complete on-wafer measurement environment. The final simulation results show an appreciable correlation with the on-wafer measurement data up to 500 GHz. Further, the need for proper calibration and de-embedding in high-frequency characterization is emphasized by investigating the S-parameters corresponding to a narrow-band amplifier at 170 GHz suitable for *G*-band radar applications.

Alongside, to bridge the THz gap further research on BiCMOS compatible SiGe devices with increased speed and breakdown voltage is being continued by various research groups across the globe. To accomplish this, two different SiGe HBT device architectures have been proposed in this work: one based on a nanowire device architecture with less lateral parasitic, which predicts an f_{MAX} above 900 GHz, and the other one is an SOI-based lateral SiGe HBT device that demonstrates an f_{MAX} of above 2.7 THz.

The asymmetric lateral SiGe HBT has a lightly doped collector that can be electro-statically adjusted by tuning the substrate bias (V_{sub}). The light collector doping of the device is very sensitive to substrate bias and allows one to switch from a high-speed device to a high voltage device. The novelty of this device is, that it achieves an f_{MAX} of 2.7 THz at V_{sub}=2 V with a BV_{CEO}=2.2 V and can be switched to an f_{MAX} of 0.8 THz with a BV_{CEO}=3.6 V at a V_{sub}=-2 V.

Indeed, this lateral SiGe HBT device provides additional leverage in switching between high-speed and high-power modes in response to the applied bias at the substrate contact, which will be very much helpful for RF circuit design engineers.

Résumé - FR

Mots-clés: Caractérisation haute fréquence, fréquence d'oscillation maximale, HBT SiGe à nanofils, HBT SiGe SOI, effet de polarisation du substrat.

Les transistors bipolaires à hétérojonction au silicium-germanium (SiGe HBT) évoluent rapidement en raison de l'augmentation des fonctionnalités et de la vitesse des systèmes de communication actuels (4G, 5G et bientôt 6G). La fréquence maximale d'oscillation des HBT SiGe dépassant largement 300 GHz, ceci permet de concevoir des circuits fonctionnant au-delà de 100 GHz; on entre alors dans la gamme THz.. Dans cette plage, de nombreuses applications sont envisagées, comme (i) l'imagerie et la détection THz, (ii) les applications radar, (iii) les équipements de mesure, comme les convertisseurs analogiques-numériques à très large bande passante. Avec l'émergence du marché des ondes millimétriques et du THz, la caractérisation et la modélisation précises des dispositifs dans cette gamme de fréquences sont obligatoires pour optimiser les performances du circuit et minimiser le nombre de boucles de la conception à la fabrication. Cependant, alors que nous continuons à développer des dispositifs aux performances accrues en termes de fréquence, l'un des problèmes majeurs est la caractérisation précise à haute fréquence (> 100 GHz). Dans ce travail, tout d'abord, une méthode systématique de vérification de la mesure haute fréquence (jusqu'à 500 GHz) du HBT SiGe est proposée. La procédure implique un calcul précis de l'effet de l'environnement passif sur l'ensemble de la mesure via une simulation électromagnétique (EM) complète. Cela garantit que l'ensemble de l'environnement de mesure est précisément incorporé dans le cadre de la modélisation EM. Afin d'inclure en plus le dispositif actif SiGe HBT, un outil de conception assistée par ordinateur (TCAD) est utilisé pour simuler les paramètres S du dispositif. Les résultats de la simulation TCAD sont introduits dans un cadre de simulation EM-plus-SPICE pour émuler un environnement de mesure complet sur la plaquette. Les résultats finaux de la simulation montrent une très bonne corrélation avec les données de mesure on-wafer jusqu'à 500 GHz. De plus, la nécessité d'un calibrage et d'un épluchage approprié dans la caractérisation haute fréquence est soulignée par l'étude des paramètres S d' un amplificateur à bande étroite à 170 GHz adapté aux applications radar en bande G. Parallèlement, pour combler le fossé THz, des recherches sur les dispositifs SiGe compatibles BiCMOS avec une vitesse et une tension de claquage accrues sont menées. Pour ce faire, deux architectures différentes de dispositifs SiGe HBT ont été investiguées dans ce travail : l'une basée sur une architecture de dispositif nanofil ayant moins de parasites latéraux, qui prédit un f_{MAX} au-dessus de 900 GHz, et l'autre est un dispositif SiGe HBT latéral sur SOI qui démontre un f_{MAX} au-dessus de 2.7 THz. Le HBT SiGe latéral asymétrique possède un collecteur légèrement dopé qui peut être ajusté électro-statiquement en appliquant une polarisation au substrat (V_{sub}). Le léger dopage du collecteur du dispositif est très sensible à cette polarisation et permet de passer d'un dispositif à haute vitesse à un dispositif à haute tension. La nouveauté de ce dispositif est qu'il atteint un f_{MAX} de 2.7 THz à V_{sub}=2 V avec un BV_{CEO}=2.2 V et peut être commuté à un f_{MAX} de 0.8 THz avec un BV_{CEO}=3.6 V et un V_{sub}=-2 V. En effet, ce dispositif latéral SiGe HBT fournit un levier supplémentaire dans la commutation entre les modes haute vitesse et haute puissance, ce qui ouvrira des nouvelles pistes de conception de circuits RF.

Acknowledgments

I would like to express my most sincere gratitude to Prof. Thomas Zimmer, and Prof. Anjan Chakrovrty for their insightful guidance, encouragement, and assistance during the course of my Ph.D. studies. I would like to extend special gratitude to my mentor Dr. Sebastien Fregonese for helping whenever I face some technical issue or need a discussion while carrying out some work. This work would have never materialized without the support of wafers from STMicroelectronics; I greatly appreciate their support.

Thanks to the doctoral committee for reviewing the work progress in time and their suggestions. Also, I would also like to thank my colleagues and many of my personal friends. My most grateful thanks to Karthi, Chandan, Biswadeep, Marco, Marine, and Dr. Marina Deng, Dr. Chhandak Mukherjee for their technical discussions.

I can't miss acknowledging the Bordeaux cricket team, Odia group in France, and friends like Subhranshu, Hari, Karan, Biswajit, Somanath, and many others. The stay, particularly during the COVID time, wouldn't have been this much interesting without them.

Thanks to my friends at IITM, Shubham, Karthik, and Aakashdeep, for discussion related to various topics. Special thanks to Karthik, Amar, Mahalingam, Balireddy, Sandeep, Smruti, our small Odia group, and many others at IITM for making campus life better during the pandemic.

Thanks to my friends and seniors, Ananta, Razim, Kalpana, Sila, Chinmay, Shiva, Mandar, Somanatha, Biswajit, Mohit, Rajesh and other friends, for their encouragement.

Last but not least, I would like to thank my family for their support. Thanks to my parents, brother, and sister for motivating me to strive for better achievements in life.

Lastly, the financial support from the French Nouvelle-Aquitaine authorities, the EU under project Taranto & IIT Madras is greatly appreciated.

Contents

List of Figures

List of Tables

General Introduction

RF to THz Applications

Increasing demand by civilians to always stay connected in the age of smartphones, computers, smartwatches, automotive vehicles, and the internet of things (IoT) is the primary driving force behind continuous research in the millimeter-wave (mm-wave), and terahertz (THz) devices and circuits domain. The electromagnetic spectrum is shown in Fig. 1 primarily focusing on the microwave to THz waves. It is to be noted that during real applications, the range of different bands in the frequency spectrum is not as rigidly as shown in Fig. 1, but falls within close range.

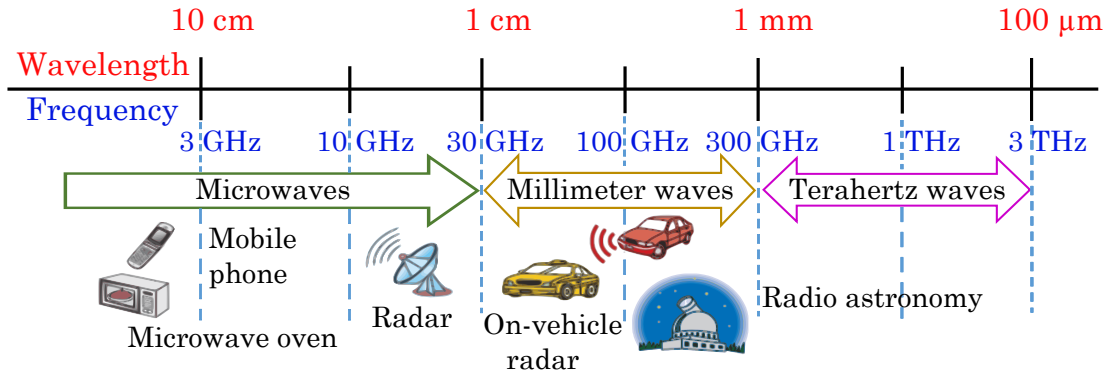


Figure 1: Frequency spectrum divided into different sub-bands and corresponding applications [1].

Generally, the microwave spectrum, which falls in the range of 3 MHz to 30 GHz, is hardly affected by environmental hazards, hence is used for long-distance communication applications like in television and radio broadcasting. However, since it has a high wavelength ($\lambda = 1/f$) compared to its high-frequency neighbors, signals from this spectrum deliver a low data transmission rate. Whereas, for a circuit to fall in the mm-wave category, the complex system must be able to broadcast between the frequency band of 30 GHz to 300 GHz. The millimeter-wave spectrum is widely used for high-speed communication and radar applications. Above 300 GHz up to 3 THz, it is called a sub-millimeter (sub-mm) wave or THz system [2]. Applications of millimeter-wave and terahertz research spread from the human body to the surface of the moon as presented in Fig. 2 [3].

Emerging THz applications can lead to economic growth in health care, safety, industrial control & environmental and security applications. One primary advantage of THz radiation is that it can penetrate several materials such as paper, clothes, and many plastic & ceramic materials. Industrial applications include pharmaceutical and bio-medical imaging, where the advantage is that it has no potential harmful effect on biological tissues since it has radiation has low photon energy.

Now, if we will consider one of the major driving forces behind radio frequency (RF) research, i.e., the typical electronic communication system, it consists of the following blocks as shown

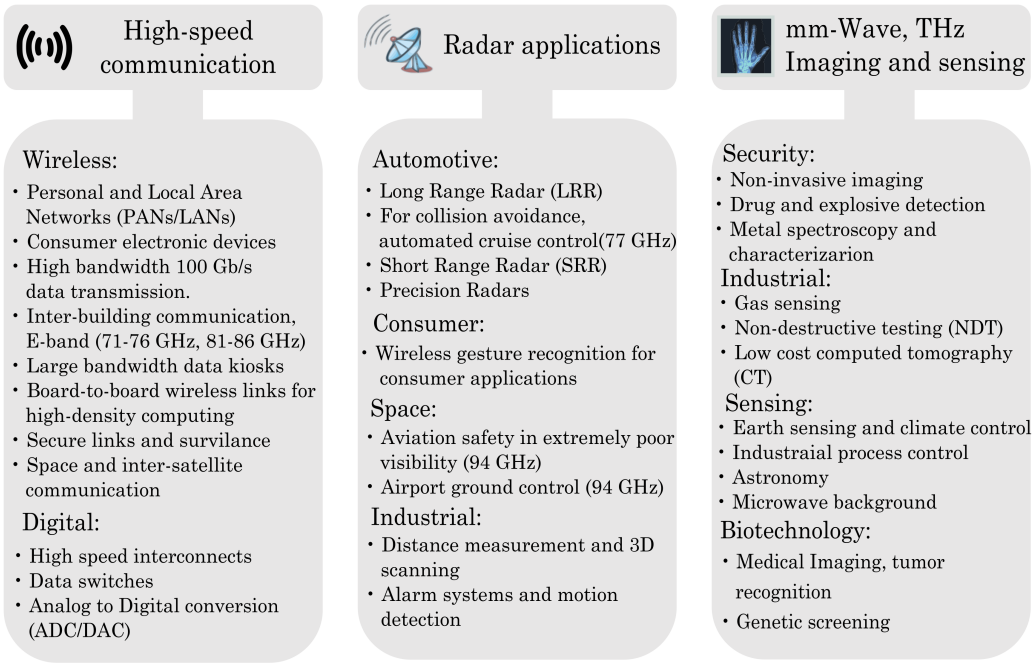


Figure 2: Potential applications of mm-wave and THz research [4].

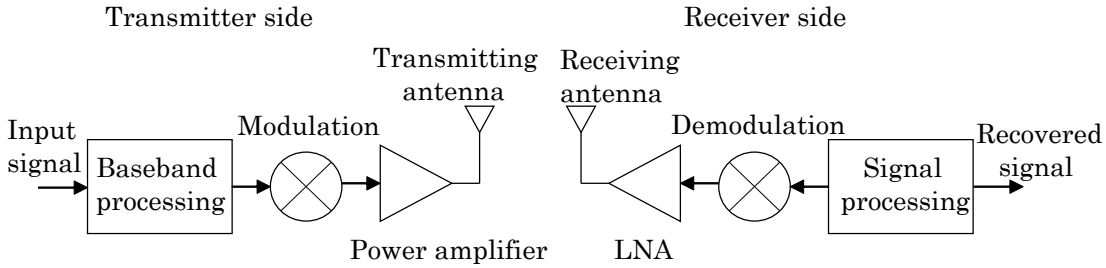


Figure 3: A typical electronic communication system.

in Fig.3. The transmission section consists of circuits for processing the base-band signal, translating it to a higher frequency, then increasing the power to a level sufficient for transmission. Similarly, the receiver section consists of a low noise amplifier (LNA) circuit whose job is to amplify the received weak signal, then the amplified signal is down-converted, and processed to the desired frequency range for the end-user. To execute a stable transmission and reception, increasingly complex topologies are required to build the discussed blocks. So to process the signals at different stages in the radio frequency spectrum, the complete system must be adequate to handle it. Transistors drive the circuit inside these modules. So for the system to work, those transistors must have a sufficiently high figure of merits (FoMs), mainly transit frequency (f_T) and oscillation frequency (f_{MAX}), and breakdown voltage. Those devices (transistors) are chosen considering various factors, most importantly, the economic and performance factors.

Why SiGe?

Complementary metal-oxide-semiconductor (CMOS) is today's dominating technology, being used in numerous applications & most widely in computer processors. It is desirable to use transistors with high power gain, high cut-off frequency, and low noise for RF applications.

But only moderate values of f_T and f_{MAX} can be achieved by CMOS technologies. On the other hand, basic bipolar transistors (Si-BJT) are considered to be unfit for modern radio frequency (RF) applications. So traditionally, III-V compound semiconductor devices are used for RF applications. However, silicon germanium (SiGe) heterojunction bipolar transistors (HBTs) offer the most effective and competing performance compared with CMOS and III-V technology. One of the reasons behind low cost is that it can be fabricated in the same CMOS technology; hence, low fabrication cost. A range of competitive advantages of SiGe over other devices is tabulated in 1.

Table 1: Performance metric comparison of various device technologies for radio frequency IC design. (E: excellent, V: very good, G:good, F: fair, P: poor), Source: [5]

Performance Metric	**SiGe HBT**	**Si BJT**	**Si CMOS**	**III-V MESFET**	**III-V HBT**	**III-V HEMT**
Frequency response	V	G	G	V	E	E
1/f and phase noise	E	V	F	P	G	P
Broadband noise	V	G	G	V	V	E
Linearity	V	V	V	E	V	E
Output conductance	E	V	F	F	E	F
Transconductance	E	E	P	F	E	F
Power dissipation	E	V	F	F	V	G
CMOS integration	E	E	–	P	P	P
IC cost	G	G	V	F	F	P

Increased functionality of modern communication systems drives the evaluation in high-frequency device technologies like CMOS, III-V HBT, III-V HEMT, and BiCMOS (SiGe-CMOS). Each of these technologies has a unique advantage over the others. For example, CMOS has the advantage of large functional density due to state-of-the-art advantageous lithography. However, the RF performance is at the lower end compared to its competitors. Due to increased functional density, interconnects get increased, creating a performance drop; in addition, due to increased parasitic gate capacitance (C_{gg}) in FinFET like structures, the RF performance can further deteriorate. III-V HBTs & HEMTs offer a faster device with f_{MAX} in the THz range, but the problem is that since these technologies are not silicon-based, hence they cannot be integrated into CMOS. BiCMOS provides the advantage of both CMOS in terms of functional density and RF performance of SiGe HBTs.

The consideration of technology in any application depends on cost, design, and volume of the circuit. For applications demanding high-speed devices like radar and communication, device performance is more important than cost. As a reminder f_T, f_{MAX}, and breakdown voltage are the most valuable characteristics for choosing a device for the above discussed RF applications. f_T and f_{MAX} are related by the standard equation in (1), and the exact calculation of f_T, f_{MAX}, and breakdown voltage will be discussed in the later part of this thesis.

$$f_{MAX} \simeq \sqrt{\frac{f_T}{8\pi R_B C_{BC}}} \tag{1}$$

In Fig. 4 a literature study is presented for CMOS, SiGe, and InP devices. The figure of merits is generalized w.r.t a characteristics size w_{char}, which is different for different technology. For

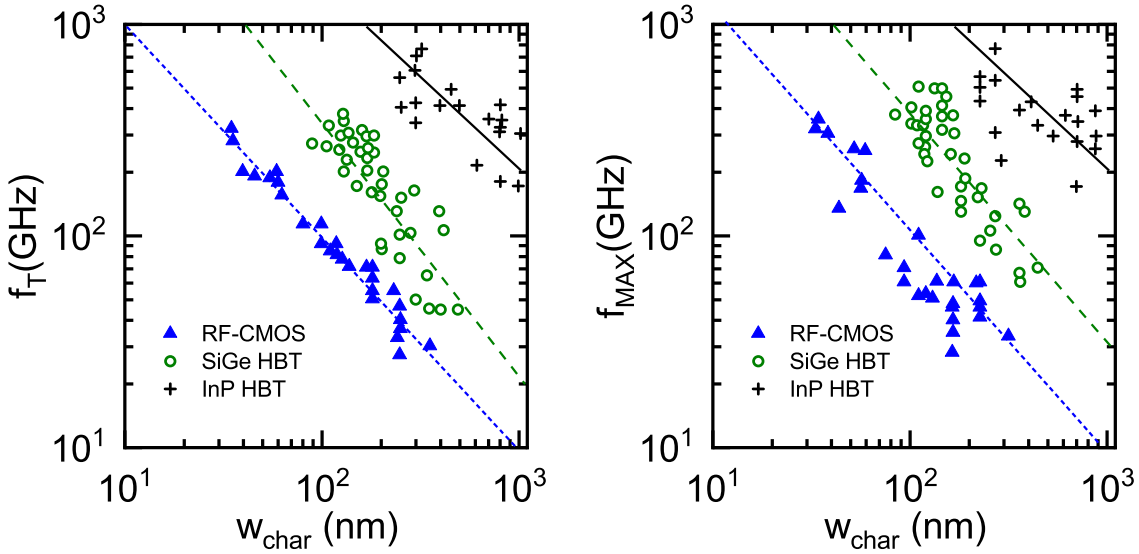


Figure 4: Prediction of f_T and f_{MAX} in different technologies from already published literature [6] on RF-CMOS, InP and SiGe HBT w.r.t characteristics size (w_{char}).

CMOS, w_{char} is the gate length (L_g) but for HBT choosing an appropriate measure is difficult as the layer thickness is confidential, so inaccessible from publications. Therefore, here w_{char} for HBTs is referred to the width of emitter w_E. Its a cautionary note that other than w_{char} there are two other factors also exist, which can influence the accuracy of the data plotted (i) in the referred publications; the drawn emitter width is listed instead of the actual width (ii) the vertical profile can change the performance instead of w_{char}. The data are separated between different technologies fairly well, where the extrapolated lines show the future predictions. We can observe that to achieve a f_T of 1 THz; we need 10 nm in CMOS, 40 nm in SiGe HBT and 120 nm in InP HBTs. Similarly, for achieving a f_{MAX} of 1 THz we need a characteristics width of 12 nm in CMOS, 44 nm in SiGe HBTs, and 130 nm in InP HBTs. From Fig. 5 a simple extrapolation predicts that f_T, f_{MAX} of 1 THz can be achieved in near future with advanced process technology.

In addition to RF performance, the permissible operating voltage is also important in circuit applications. Suitable FoMs for assessing the above parameter is open base collector-emitter breakdown voltage (BV_{CEO}) and open emitter collector-base voltage (BV_{CBO}) in HBTs and zero gate bias drain to source voltage BV_{DSS} in MOSFETs. A fair comparison can be made between both the technologies with the help of BV_{CBO} and BV_{DSS} as both are measured completely in the same way, but unfortunately, since CMOS is not focused on high power applications, so the available data analyzed for BV_{DSS} is very less in literature; hence only SiGe and InP are compared in Fig. 6.

In this section, the overall discussion concludes that SiGe HBT can be chosen as a better choice over other options to lead the RF applications. The above-discussed improvements in SiGe performance are due to (i) continuous optimization of the germanium profile in the base (ii) scaling of both lateral & vertical dimensions, and (iii) development in material growth. In the next section, the physics of the operation of the SiGe HBT device is discussed in detail.

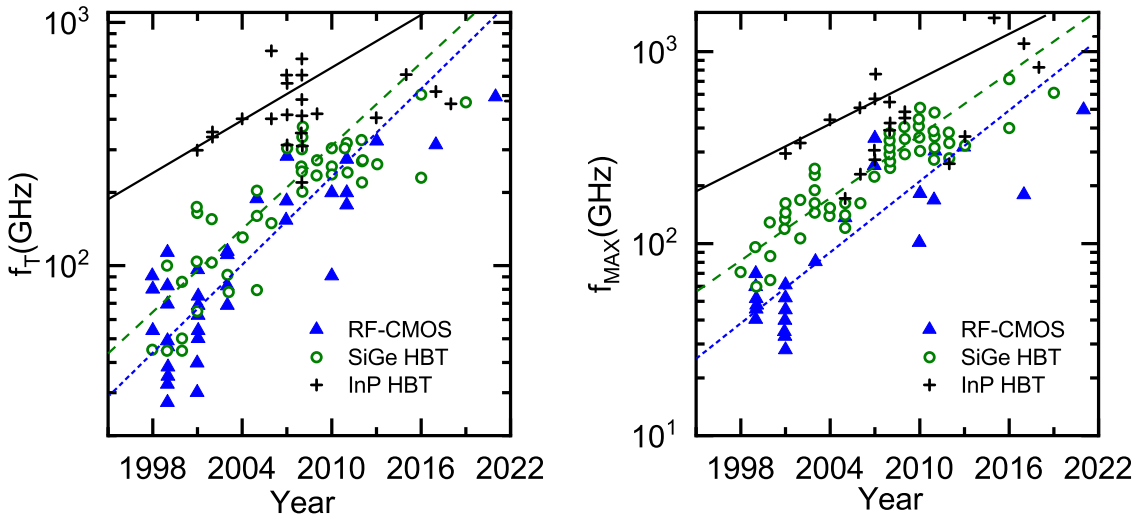


Figure 5: Prediction of f_T and f_{MAX} in different technologies from already published literature [6], [7], [8], [9] [10], [11] on RF-CMOS, InP and SiGe HBT w.r.t time.

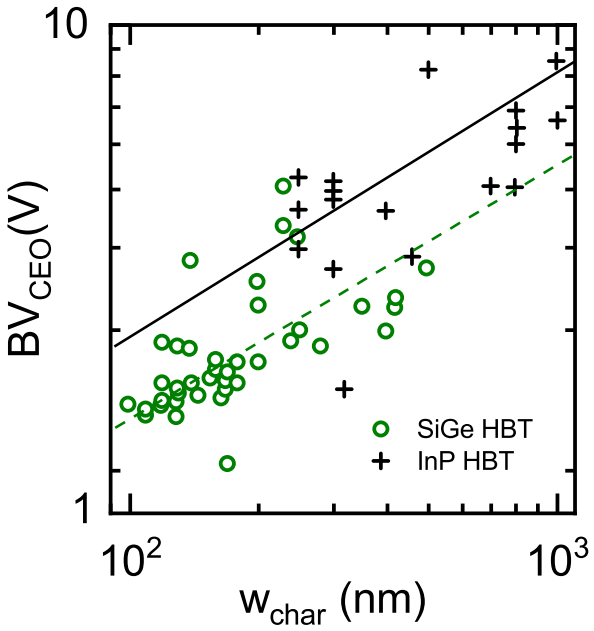


Figure 6: Results of BV_{CEO} in different technologies from already published literature [6] on RF-CMOS, InP and SiGe HBT w.r.t characteristics size (w_{char}).

SiGe HBT Device Physics

This section is dedicated to the operation & principles of SiGe HBT that make the device special for RF applications. The analysis begins with the concept of bandgap engineering, understanding the Gummel characteristics, high-frequency device figures of merits & limitations. The secondary effects are also discussed along with introductory device physics, which plays a critical role in actual device operation.

Kroemer introduced the idea of HBT; nine years later, the initial patent for the bipolar transistor was filed in 1948 [12]. In his work "Theory of a wide-gap emitter for transistors" Kroemer implemented the idea of decoupling the DC current gain (β_{DC}) from the ratio of doping in the base to its emitter, and it earned him a Nobel prize in physics in the year 2000.

Physics of Bipolar Junction Transistor

This section provides an overview of the bipolar junction transistor operating principle. A schematic of the device is given in Fig. 7(a). The emitter is heavily doped with n^{++}, the base is moderately doped with p^{+}, and the collector is lightly doped with n. The bias applied at the junction keeps the device in forward active mode of operation.

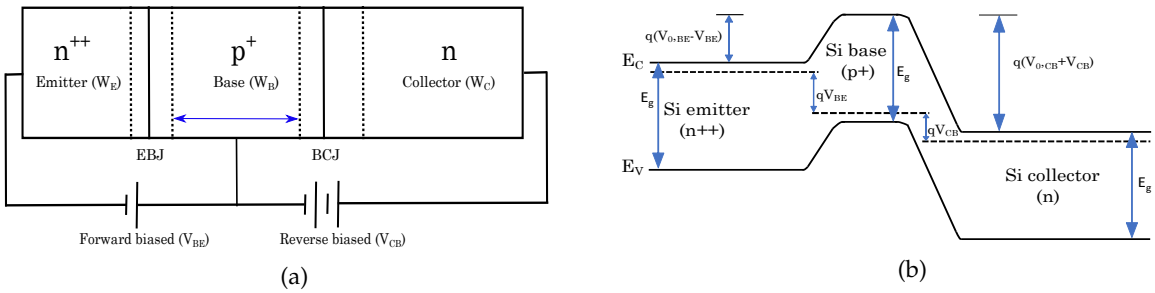


Figure 7: (a) Basic bipolar device structure, EBJ, and BCJ are the base-emitter and base-collector depletion regions, respectively (b) Energy band diagram of the corresponding device under forward bias condition, the dotted lines are to indicate the Fermi levels and we can observe the energy band gap E_g is uniform because of homo-structure.

Terminal Currents

The operation of a bipolar transistor can be comprehended by drawing its energy band diagram as depicted in Fig. 7(b). As the emitter is heavily doped, a large number of ionized dopants will result in a large number of free carriers (electrons) in the conduction band. Consequently, there will be a current component with a magnitude equal to the number of electrons with sufficient energy to cross the base-emitter barrier, given by $q(V_{0,BE} - V_{BE})$, where $V_{0,BE}$ is given by 2. Note: In N_{XY} notation, X is for the impurity type, either donor (D) or acceptor (A), and the Y is for the device region like for base (B), collector (C) and emitter (E).

$$V_{0,BE} = \frac{kT_a}{q} \ln\left(\frac{N_{AB}^{-} N_{DE}^{+}}{n_{i0}^{2}}\right) \tag{2}$$

where N_{AB}^{-} is the ionized acceptor impurity concentration in the base. N_{DE}^{+} is the ionized donor concentration in the emitter & n_{i0} is the intrinsic carrier concentration, q is the charge of electron [13]. If the base is sufficiently thin and recombination in the base is disregarded, then the electric field across the base-collector junction will sweep all the electrons diffused from the emitter into the base to the collector. As the electron distribution in the conduction band is approximately Boltzmann, the collector current density (J_C) is given by (3).

$$J_C \approx \frac{kT_a \mu_{nb} n_{i0}^{2}}{W_B N_{AB}^{-}} e^{qV_{BE}/kT_a} = N_{DE}^{+} \frac{kT_a \mu_{nb}}{W_B} e^{-q(V_{0,BE} - V_{BE})/kT_a} \tag{3}$$

Where μ_{nb} is the minority carrier mobility in the base, the ionized acceptor impurities in the base valence band will generate a hole diffusion current from the base to the emitter. Again assuming negligible recombination, the base current density (J_B) can be expressed as follows: (4).

$$J_B \approx \frac{kT_a\mu_{pe}n_{i0}^2}{L_{pe}N_{DE}^+}e^{qV_{BE}/kT_a} = N_{AB}^-\frac{kT_a\mu_{pe}}{L_{PE}}e^{-q(V_{0,BE}-V_{BE})/kT_a} \tag{4}$$

where μ_{pe} is the hole mobility, and L_{pe} is the diffusion length in the emitter. So the standard DC current gain (β_{DC}) formulae is

$$\beta_{DC} = \frac{J_C}{J_B} \approx \frac{\mu_{nb}L_{PE}N_{DE}^+}{\mu_{pe}W_BN_{AB}^-} \tag{5}$$

where W_B is the effective base width.

Transit Times

Bipolar devices are also called as charge controlled device. This is due to the fact that the current in the device is controlled by the minority charge carrier distribution. To the applied ac signal, the device responds with modulation in internal stored charge. This change in stored charge causes a delay that limits the transistor gain at high frequencies.

The total forward transit time i.e the delay of a carrier to travel from the emitter to collector is given by

$$\frac{1}{2\pi f_T} = \tau_{ec} = \tau_e + \tau_b + \tau_{cbd} + \frac{kT}{qI_C}(C_{jbe} + C_{jbc}) + r_cC_{jbc} \tag{6}$$

where τ_e is the charge storage time in the emitter, τ_b is the base transit time, τ_{cbd} is the base-collector depletion region transit time, C_{jbe} & C_{jbc} are the junction capacitance associated with base-emitter and base-collector depletion regions, respectively & r_c represents the collector resistance [14]. For a silicon BJT τ_e, τ_b and τ_{cbd} are given by

$$\tau_e \approx \frac{q}{2kT_a}\frac{{W_E}^2}{\mu_{pe}\beta_{DC}}, \quad \tau_b \approx \frac{q{W_B}^2}{\mu_{nb}kT_a}, \quad \tau_{cbd} \approx \frac{W_{CBD}}{2v_{sat}} \tag{7}$$

where, W_{CBD} is the width of the depletion region at base collector junction and v_{sat} is the saturation velocity. Transit frequency or unity current gain cut-off frequency is the frequency at which the short circuit ac current gain is unity. The aim of a device engineer is always to optimize f_T while maintaining breakdown voltage. Minimization of τ_{ec} requires (i) minimized W_E, W_B and W_{CBD} (ii) minimized paracitic components like C_{jbe} & C_{jbc} and collector resistance.

SiGe Hetero-Junction Bipolar Transistors

The conclusions from the last section is, to increase the f_T (or to reduce base transit time τ_b), base width (W_B) has to be decreased. But decreasing W_B, increases base resistance (r_b), and thus decreases f_{MAX}. This is why the base doping needs to be increased, resulting in a decrease in DC current gain.

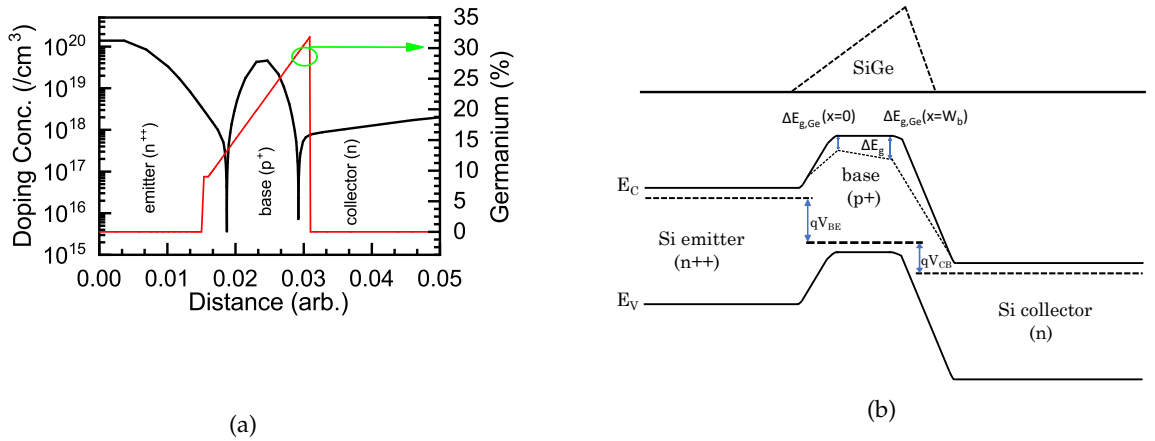
To elaborate since both material in base and emitter are same both see a potential barrier of same height. Therefore, for a given thermal excitation, both electron & holes see same push of $e^{(V_{BE}-V_0)/V_T}$ and to decrease τ_b, we need to reduce doping in the base (N_{AB}^-). However,

this process increases the base resistance. But this problem can be avoided by introducing germanium in the base region. Semiconductor materials exhibit an energy band structure, where forbidden energy band gap given by (E_g) = conduction band gap (E_c) - valence band gap (E_v). Due to introduction of Ge, the strained $Si_{1-X}Ge_X$ region formed will have smaller E_g than silicon. The reduction in energy gap (E_g) has been found experimentally [15], [5] to be

$$\Delta E_{g,Ge} \approx 0.96x - 0.43x^2 + 0.17x^3 \tag{8}$$

This bandgap reduction (ΔE_g) mostly occurs as an offset in the valence band. SiGe HBT devices are double hetero-junction type because when the bandgap in the base got narrowed due to the introduction of germanium, the emitter and collector are wide bandgap Si-regions. Hence, the hetero-junctions are formed close to base-emitter and base-collector junctions. These hetero-junctions will influence the carrier transport and will directly affect the terminal currents of bipolar transistors.

As discussed that $SiGe$ straining introduces offset in the valence band; hence, the carriers from the emitter to the base will see a low barrier height than the carriers back injected from the base into the emitter. It will result in an exponential increase in DC current gain (β_{DC}). If the difference between emitter and base bandgap is ΔE_g, then it will see a $e^{\Delta E_g/kT_a}$ factor increase compared to device without bandgap narrowing. Further, Kroemer postulated that if gradient-based bandgap engineering can be achieved where the maximum bandgap will be at the emitter side and minimum at the collector side it will introduce an additional quasi-electric field that will accelerate the carriers, thereby reducing the base transit time significantly. It took almost 30 years for the material engineers to develop a facility that could realize this theory into fabrication.



(a) (b)

Figure 8: (a) A standard $n^{++}p^{+}n$ doping profile with graded germanium in the base (b) (bottom) band diagram corresponding to the doping profile indicating the deviation from Si homo-structure band diagram, the graded Ge profile (top) giving rise to change in ΔE_g at different locations in the base. Note: Change in apparent band gap ($\Delta E_{g,app}$) is not shown here.

A device doping profile and corresponding band diagram is shown in Fig.8. We can observe that there is a bandgap reduction due to the introduction of the germanium. This reduction significantly helps to achieve the improved figures of merit like current gain, transit frequency, and oscillation frequency compared to only silicon-based device.

Terminal Currents

The collector current density (J_C) of a *SiGe* HBT can derived to be [5]

$$J_C \approx n_{i0,si}^2 \tilde{\gamma}\tilde{\eta}\frac{\mu_{nb,Si}}{N_{AB}^- W_B}\Delta E_g(grade) e^{\Delta E_{g,app}/kT_a} e^{\Delta E_{g,Ge(0)}/kT_a}\left(e^{qV_{BE}/kT_a}-1\right) \tag{9}$$

where $\tilde{\gamma} = (N_C N_V)_{SiGe}/(N_C N_V)_{Si} < 1$, $\tilde{\eta} = \mu_{nb,Si_{Ge}}/\mu_{nb,Si} > 1$, $n_{i0,Si}$ and $\mu_{nb,Si}$ are the intrinsic carrier density and electron mobility in silicon, $\Delta E_{g,app}$ is an apparent band gap reduction related to heavy doping effects, N_C, N_V are the effective carrier density in both conduction and valency band. Also J_C can be written as

$$J_C \approx J_{C0}(e^{V_{BE}/V_T}-1) \tag{10}$$

where J_{C0}, is the collector saturation current which got exponentially enhanced due to introduction of *Ge* in the base.

$$J_{C0} = n_{i0,si}^2 \tilde{\gamma}\tilde{\eta}\frac{\mu_{nb,Si}}{N_{AB}^- W_B}\Delta E_g(grade) e^{\Delta E_{g,app}/kT_a} e^{\Delta E_{g,Ge}(0)/kT_a} \tag{11}$$

Similarly the base current density (J_B) can be written as

$$J_B \approx \frac{kT_a \mu_{pe} n_{i0}^2}{L_{PE} N_{DE}^+} e^{qV_{BE}/kT_a} = \frac{q}{G_e} e^{qV_{BE}/kT_a} \tag{12}$$

where G_e is the emitter Gummel number given by $G_e = L_{PE}N_{DE}^+ \Big/ D_{PE} n_{i0,e}^2$. From above derivations the DC current gain of the SiGe HBT can be written as

$$\beta_{DC} \approx \frac{\mu_{nb,Si} L_{PE} N_{DE}^+}{\mu_{pe} N_{AB}^- W_B}\tilde{\gamma}\tilde{\eta}\frac{\Delta E_g(grade)}{kT_a} e^{\Delta E_{g,app}/kT_a} e^{\Delta E_g(0)/kT_a} \tag{13}$$

$$= \beta_{DC,Si}\left(\tilde{\gamma}\tilde{\eta}\frac{\Delta E_g(grade)}{kT_a} e^{\Delta E_{g,app}/kT_a} e^{\Delta E_g(0)/kT_a}\right) \tag{14}$$

where ΔE_{Ge} is the *Ge* induced bandgap narrowing estimated from the mole-fraction of germanium content same as in (8)

$$\Delta E_{g,Ge} \approx 0.96x - 0.43x^2 + 0.17x^3 \tag{15}$$

and the apparent induced bandgap narrowing

$$\Delta E_{g,app} \approx 18 \times 10^{-3} \ln\left\{\frac{N_{AB}^-}{N_{DE}^+}\right\} eV \tag{16}$$

The expression for $E_{g,app}$ is only valid when the doping concentration is $> 7 \times 10^{17} cm^{-3}$, if the base doping is below this level then the doping induced bandgap narrowing is negligible in the base. Then the modified expression is

$$\Delta E_{g,app} \approx -18 \times 10^{-3} \ln\left\{\frac{N_{DE}^+ \times 300k}{T_a \times 7 \times 10^{17} cm^{-3}}\right\} eV \tag{17}$$

For instance, if we assume $\tilde{\eta}$ and $\tilde{\gamma}$ are close to unity, then the effect of *Ge* on β can be estimated from eq(13). Let's consider if the ionized emitter doping $N_{DE}^{+} = 10^{20} cm^{-3}$ & base doping $N_{AB}^{-} = 5 \times 10^{18} cm^{-3}$ then for a graded *Ge* content of 10% with a total of 20% at the emitter side of the base can give an improvement of ~ 400 in DC current gain.

In addition, the presence of germanium content will improve the Early voltage (V_A), yielding higher V_A and thus higher output conductance. The V_A with and without germanium in the base are related by

$$\frac{V_{A,SiGe}}{V_A} \approx e^{\Delta E_{g,Ge}(grade)/kT} \left[\frac{1 - e^{\Delta E_{g,Ge}(grade)/kT}}{\Delta E_{g,Ge}(grade)/kT} \right] \tag{18}$$

Effect of Ge on Transit Time

Due to the introduction of germanium, there is a barrier lowering, along with that when the germanium profile is graded, it will add an additional quasi-electric field in the base region, so together will help the base transit time to reduce significantly. Also, due to barrier lowering between emitter and base, the emitter transit time will reduce. On the other hand, we can expect only a minor change in the collector base (CB) space charge region (SCR) transit time. The modified transit time equations due to the introduction of *Ge* are

$$\tau_{e,SiGe} \approx \frac{q}{2kT_a} \frac{W_E^2}{\mu_{pe,Si}.\beta_{DC,SiGe}} = \tau_{e,Si} \frac{\beta_{DC,Si}}{\beta_{DC,SiGe}} \tag{19}$$

$$\tau_{b,SiGe} \approx \frac{qW_B^2}{\tilde{\eta}\mu_{nb,Si}} \frac{1}{\Delta E_g(grade)} = \frac{\tau_{b,Si}}{\tilde{\eta}} \frac{kT_a}{\Delta E_g(grade)} \tag{20}$$

$$\tau_{cbd,SiGe} \approx \frac{W_{CBD}}{2v_{sat}} = \tau_{cbd,Si} \tag{21}$$

where we can observe the base and emitter transit time are greatly reduced when compared with homo-junction device.

Second Order Effects

All discussions made so far assumes that the device operates in the low injection mode of operation, but if we will refer to the transit time equation 6 then, it's apparent that to minimize the time constant, we need a higher collector current density (J_c). As circuit designers are aggressively working towards moving the operating frequency closer to peak f_T; hence most of the modern SiGe HBTs are being operated in the high injection regime. As a consequence of this, various secondary effects come into the picture. Although there are various secondary effects due to the complex underlying physics of operation, only two are discussed here.

Webster Effect

While deriving the DC terminal currents in case of Si BJT or SiGe HBT, it is assumed that the number of electrons injected into the base (n_b) are much lesser than the ionized hole concentration ($n_b << N_A^{-}$). Thus the hole concentration in the base region will be $p_b = N_A^{-}$. However, when the device enters to high injection regime of operation, this condition looses its validity and the total concentration in the base becomes $p_b = N_A^{-} + n_b = N_A^{-} + J_C/qv_d$. Where J_C is the

current density A/cm^2 and v_d is the drift velocity of electrons. Now considering the current gain (β) equation of a Si transistor, it can be re-written as

$$\beta_{DC,Si} \approx \frac{\mu_{nb} L_{PE} N_{DE}^{+}}{\mu_{pe} p_b W_B} \approx \frac{\mu_{nb} L_{PE} N_{DE}^{+}}{\mu_{pe} N_{AB}^{-} W_B} \frac{1}{1 + J_C / q N_{AB}^{-} v_d} = \beta_{DC0,Si} \frac{1}{1 + J_C / q N_{AB}^{-} v_d}. \tag{22}$$

Thus, if J_C becomes $> qN_{AB}^{-}v_d$ then the DC current gain will start rolling-off with increase in collector current. This effect was first demonstrated by Webster in 1954, so called Webster effect [16].

Kirk Effect (Base Push-out)

When the topic of discussion is about high speed bipolar devices, concept of Kirk effect [17] cannot be ignored.

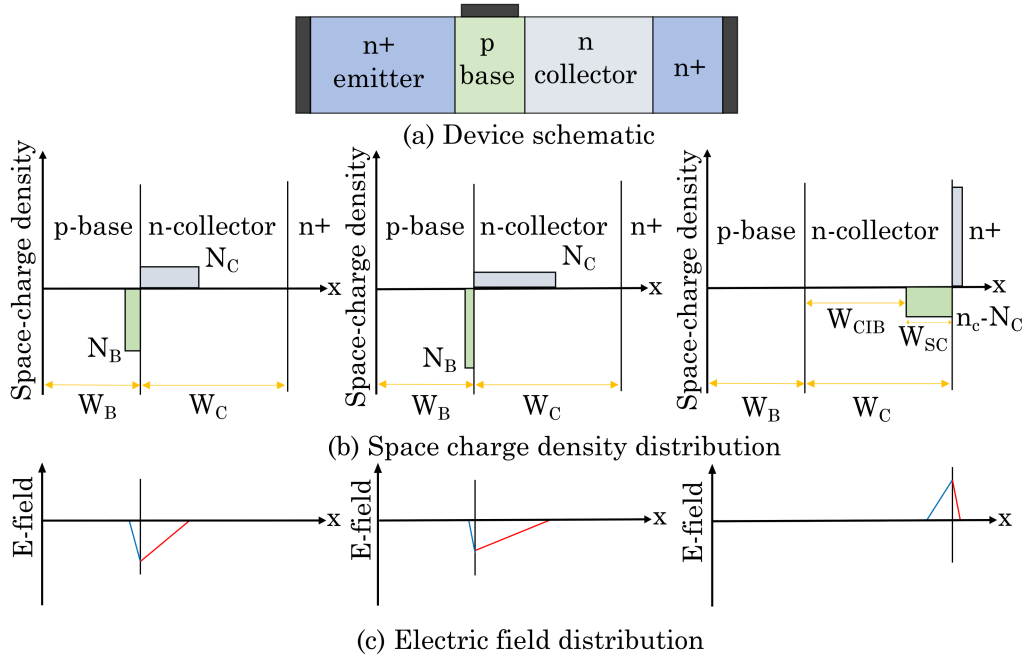


Figure 9: (a) Schematic of the structure depicting different regions (b) Space charge density distribution w.r.t different regions (c) Electric field at varation w.r.t space charge variation [18].

While doing the calculation of transport equations, it is generally assumed that the background concentration of at the collector-base space charge region (CB SCR) is greater than the injected carriers (electrons) in npn device, i.e $J_C/v_{sat} << qN_{DC}^{+}$. Now if we will write the Poission's equation

$$\frac{dE}{dx} = \frac{q}{\varepsilon_0 \varepsilon_r} \left(p - n + N_{DC}^{+}\right) \approx \frac{1}{\varepsilon_0 \varepsilon_r} \left(qN_{DC}^{+} - qn\right) \approx \frac{1}{\varepsilon_0 \varepsilon_r} \left(qN_{DC}^{+} - J_C/v_{sat}\right) \tag{23}$$

where v_{sat} is the saturation velocity and p, n are the mobile hole and electron density in the CB SCR. Referring to the Poisson equation, it is evident that the derivative of electric field (dE/dx) in the CB-SCR depends on the net current in this region.

Now for charge neutrality, the net charge on the base side of CB SCR must be equal to the net charge on the collector side of CB SCR. The magnitude of the electric field is maximum

at the metallurgical junction, and it is zero outside the SCR. Under low injection condition the penetration of the depletion regions into corresponding neutral regions is decided by the ionized impurity in the corresponding regions and applied voltage that modulates dE/dx. So under low injection, the depletion region is independent of the current in the device.

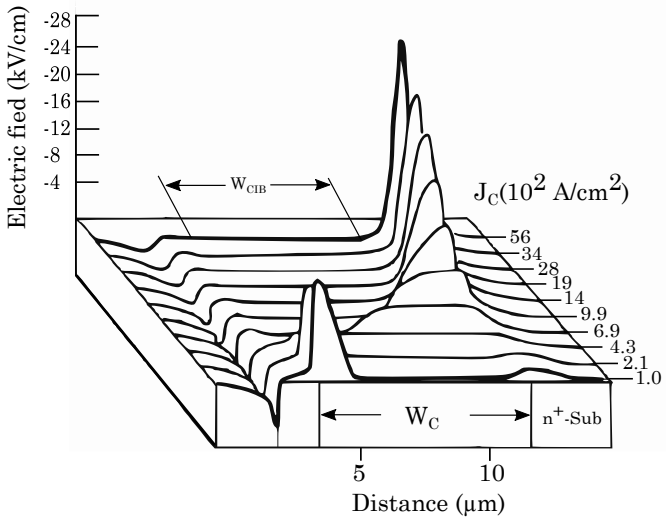


Figure 10: Shifting of the electric field from the collector base space charge region to the edge of buried sub-collector region with increase in collector current density [19].

However, when it enters high injection, the mobile charge density in CB SCR is no longer negligible compared to the background concentration since, at the collector side of the CB SCR, the immobile carriers are positive, it will get reduced by J_C/v_{sat} and to balance it accordingly the immobile carriers of the base side of CB SCR will increase. In fact, this will widen the SCR region to neutral collector region until the collector current density (J_C) becomes $qv_{sat}N_{DC}^{+}$ this will continue. Once collector current density crosses $qv_{sat}N_{DC}^{+}$ the intrinsic collector region cannot support the electric field, and the collector side of the space charge region moves to the buried sub-collector region, and without further increase in the collector current density, the net change in the intrinsic collector region becomes positive, and the base side of the CB SCR will move to the interface of the edge of the intrinsic and buried collector layer (refer Fig. 9 & 10). Thus there will be a widening of the base with an additional width of the intrinsic collector.

This effect was discovered by Kirk in 1962 and published titled as "A theory of transistor cutoff frequency (f_T) fall-off at high current densities" [17]. This effect is important because, as we know, base transit time $\tau_b \propto W_B^2$, so with the increase in effective base width, the base transit time will increase, and the transit frequency will fall.

Breakdown Voltage

The breakdown voltage determines the maximum operating voltage of a bipolar transistor in the particular transistor configuration for different modes of operation. For bipolar devices, the three significant sources of breakdown are avalanche, punch-through, and reach-through breakdown. The overall breakdown voltage of the transistor is dominated by the mechanism that occurs at a lower voltage.

Traditionally, bipolar devices have wide base width, and their operation is limited by avalanche breakdown. This breakdown phenomenon is due to the mechanism of avalanche breakdown in a pn junction that limits the maximum reverse bias voltage applied to that junction. In the case

of open emitter configuration, the breakdown would occur across the collector-base junction and is denoted as BV_{CBO}. For a planner junction, it can be approximated by the following expression [20]:

$$BV_{CBO(plane)} = \frac{\varepsilon_s E_{br}^2}{2qN_C} \tag{24}$$

where E_{br} is the critical electric field (which is approximately 2×10^5 V/cm), ε_s is the dielectric constant of silicon and N_C is the collector doping concentration. However in reality, the junction curvature effect shifts the electric filed in the curved part of the depletion region. therefore, the actual measured breakdown voltage BV_{CBO} is lower than the $BV_{CEO(plane)}$ value. The emitter-collector breakdown voltage BV_{CEO} with base opened is approximated by the empirical expressions [20]:

$$BV_{CEO} = \frac{BV_{CBO}}{\beta^{\frac{1}{n}}} \tag{25}$$

where n is a constant ranging between 3 and 6.

The second source of breakdown mechanism in bipolar transistors is the punch-through breakdown, which occurs when emitter-base (EB) and collector-base (CB) depletion regions touch inside the base. Once punch-through has occurred, the EB and CB junctions become electrostatically coupled and cause a large current flowing from the collector to the base regions. The base-collector voltage V_{PT} causing punch-through breakdown voltage is given by [20]:

$$V_{PT} = \frac{qN_B W_B^2}{2\varepsilon_s} \tag{26}$$

Therefore, the thickness and doping concentration of the base region should be chosen carefully to avoid V_{PT} below the supply voltage [21].

The third source of breakdown for the BJTs is the reach-through breakdown. This phenomenon occurs when the edge of the depletion layer is in the thin epi-layer collector side contacts the buried layer edge and triggers the avalanche breakdown. The reach-though voltage V_{RT} is given by [20]:

$$V_{RT} = \frac{qN_C W_C^2}{2\varepsilon_s} \tag{27}$$

As shown in (27), the reach-through breakdown is dominated by W_C (width of collector and collector doping N_C). Therefore, the thickness and doping concentration of the epi-collector should be chosen carefully to avoid V_{RT} below the supply voltage. In modern RF bipolar transistors, due to the use of narrow lightly-doped collectors to reduce collector and collector-base junction transit times, the BV_{CBO} as well as the BV_{CEO} are dominated by V_{RT} [20].

The improvement in transistor f_T and f_{MAX} have not been occurred without some significant compromises, primarily involving transistor BV_{CEO} [22]. The reduction of breakdown voltage is related to the well-known Johnson's limit on bipolar transistor performance, where the product of breakdown voltage and transit frequency is roughly constant (approximately 340 GHz.V in the case of silicon devices and higher than 500 GHz.V for SiGe-HBT [23]). This limit is material-related and consequently is not subject to device design and process technology improvements. Overall, the Si-bipolars and SiGe-HBTs have penetrated into many RF and high-speed mixed-signal applications, where breakdown, speed and gain are more favorable than

CMOS. However, the process for making high-performance bipolar devices has become more complicated than ever, with increasing process steps. The next section discusses the history of SiGe development with some advanced technologies used today.

History of SiGe Fabrication

IBM demonstrated the first working SiGe HBT device in 1987. It was fabricated using molecular beam epitaxy (MBE) with different germanium contents processed at low-temperature [24]. After the successful demonstration by IBM, the graded *Ge* SiGe devices were tested, resulting in a frequency response of 75 GHz [25]. The first commercial production of SiGe on a 200 nm wafer started in 1994 [26]. In 2003 by integrating *SiGe* with 130 nm CMOS node (schematic in Fig. 11) for the first time a record high f_T, f_{MAX} of 200 GHz and 280 GHz achieved respectively [27]. Again after a year, a new record was achieved by same IBM demonstrating a f_T of 300 GHz, and f_{MAX} of 350 GHz [28]. The world's first 90 nm SiGe BiCMOS was also announced by IBM in the year 2014, featuring f_T of 300 GHz, and f_{MAX} of 360 GHz [29]. However, later the evaluation of SiGe HBT performance moved to Europe.

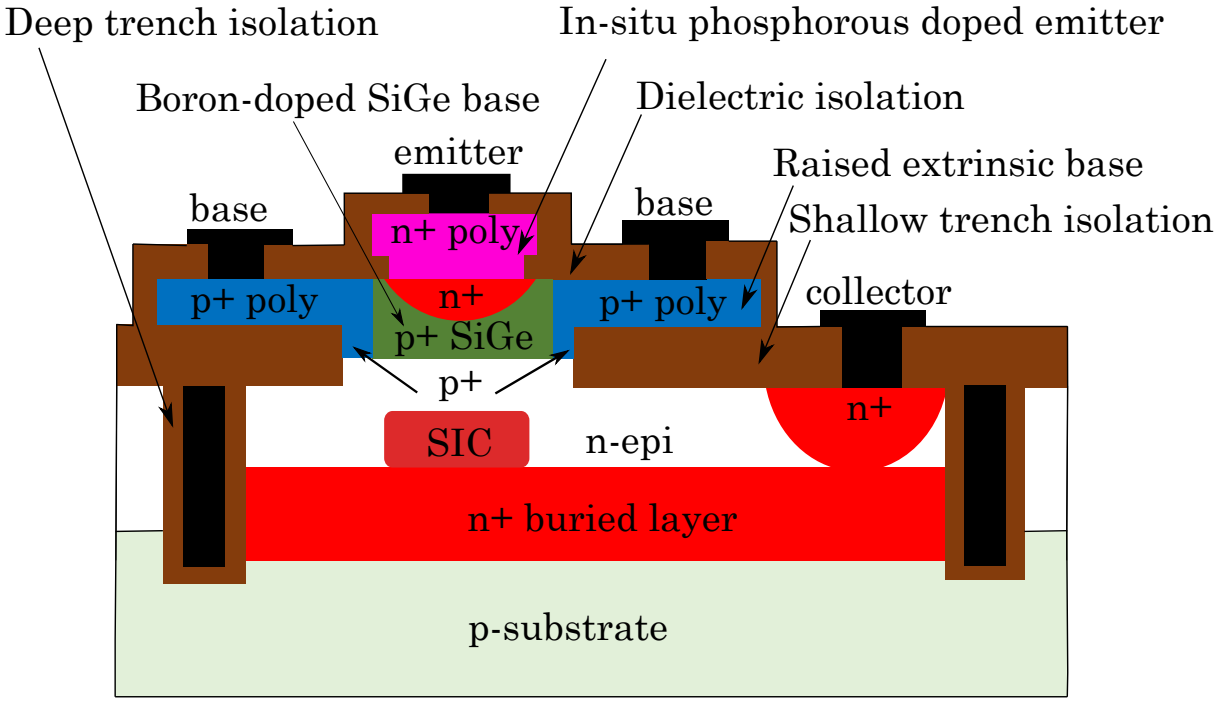


Figure 11: Schematic cross-sectional view of IBMs' 130 nm SiGe HBT [30].

To lead the area of SiGe HBT and take it to a cutting edge frequency performance, the European semiconductor industry received massive support from the European Union (EU) through 4 projects DOTFIVE, DOTSEVEN, TARANTO, and RF2THZ. In 2010 and 2011, IHP first achieved f_{MAX} of 500 GHz in the DOTFIVE [31], this milestone achievement became possible by reducing the extrinsic base resistance (R_{Bx}) by elevated extrinsic base (EEB) and base link architecture. Again by IHP in DOTSEVEN, with modification in EEB architecture itself 570 GHz f_{MAX} was achieved [32]. Since IHP's device tops in RF performance to date, the performance improvement achieved by IHP's device and a brief about the kind of architectural modification done is discussed in section 1.5.1. In RF2THz project, for the first time using a Double-Poly-Silicon-Self-Aligned (DPSA) architecture with selective epitaxial growth (SEG) of the base, the

device was integrated with 55 nm (300 nm wafer) STM's BiCMOS technology [33]. The f_T achieved by the resulting technology is 320 GHz.

Structural Modifications in IHP/IFX Device

In DOTSEVEN project Infineon (IFX) and IHP jointly demonstrated the potential of epitaxial base link (EBL) architecture [34]. This device was implemented in Infineon's 130 nm BiCMOS environment. In this case, the collector module, buried layer, STI and DTI were fabricated at Infineon, and then the EBL module was fabricated at IHP; the 6 layer Cu metalization process, was completed at Infineon. This device is one of the advanced architectures that resulted into f_T of 300 GHz and f_{MAX} of 500 GHz [35].

The fastest device by IHP was achieved in a process with non-selective epitaxial growth (NSEG) for the SiGe base and elevated extrinsic base (EEB) using a separate selective epitaxial process. The optimized version of the device in a bipolar-only process flow resulted in a peak f_T/f_{MAX} value of 505 GHz/720 GHz, which is the best to date [36]. These are bipolar only because, there are some challenges while integrating with CMOS which are as follows:

- Adaptation of thermal constraints with CMOS process flow.
- Nickel silicidation is compatible with 130 nm baseline CMOS process flow.

And intermediate result of f_T/f_{MAX} 470 GHz/610 GHz has also been demonstrated in the BiCMOS process flow.

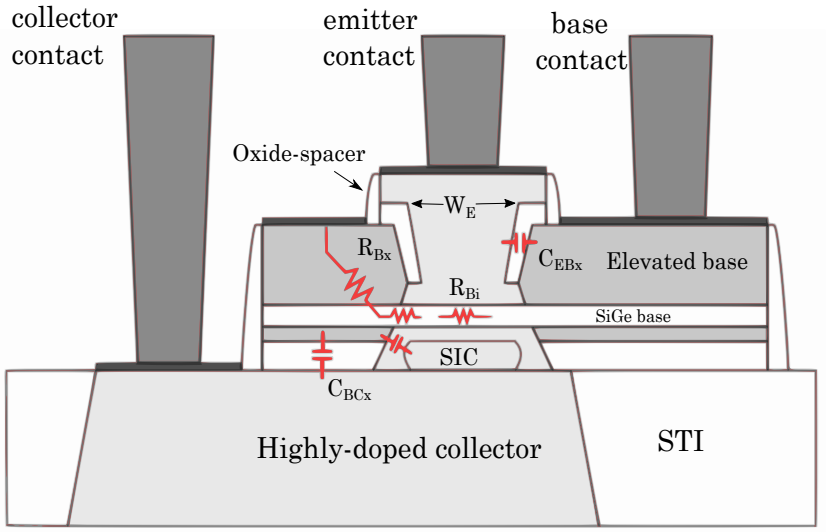


Figure 12: Schematic diagram of IHP's 130 nm SG13G2 high performance SiGe HBT highlighting the elevated extrinsic base (EEB) [37].

With this, IHP being a leader in SiGe production with the highest f_{MAX} so far, the novelty in IHPs technology is discussed here.

Elevated Extrinsic Base (EEB) Architecture

The key features of the technology are listed below:

- The elevated extrinsic base (EEB) regions are self-aligned to the emitter (refer Fig.12). It resulted in a low R_{Bx} compared to double-polysilicon-self-aligned (DPSA) architecture.

- Due to the formation of HBT in a single active area without STI between collector and emitter contact regions resulted in a low collector resistance (R_c) and small collector to substrate junction areas. Based on this modifications the state-of-the-art 570 GHz f_{MAX} HBT is obtained [38], [32].

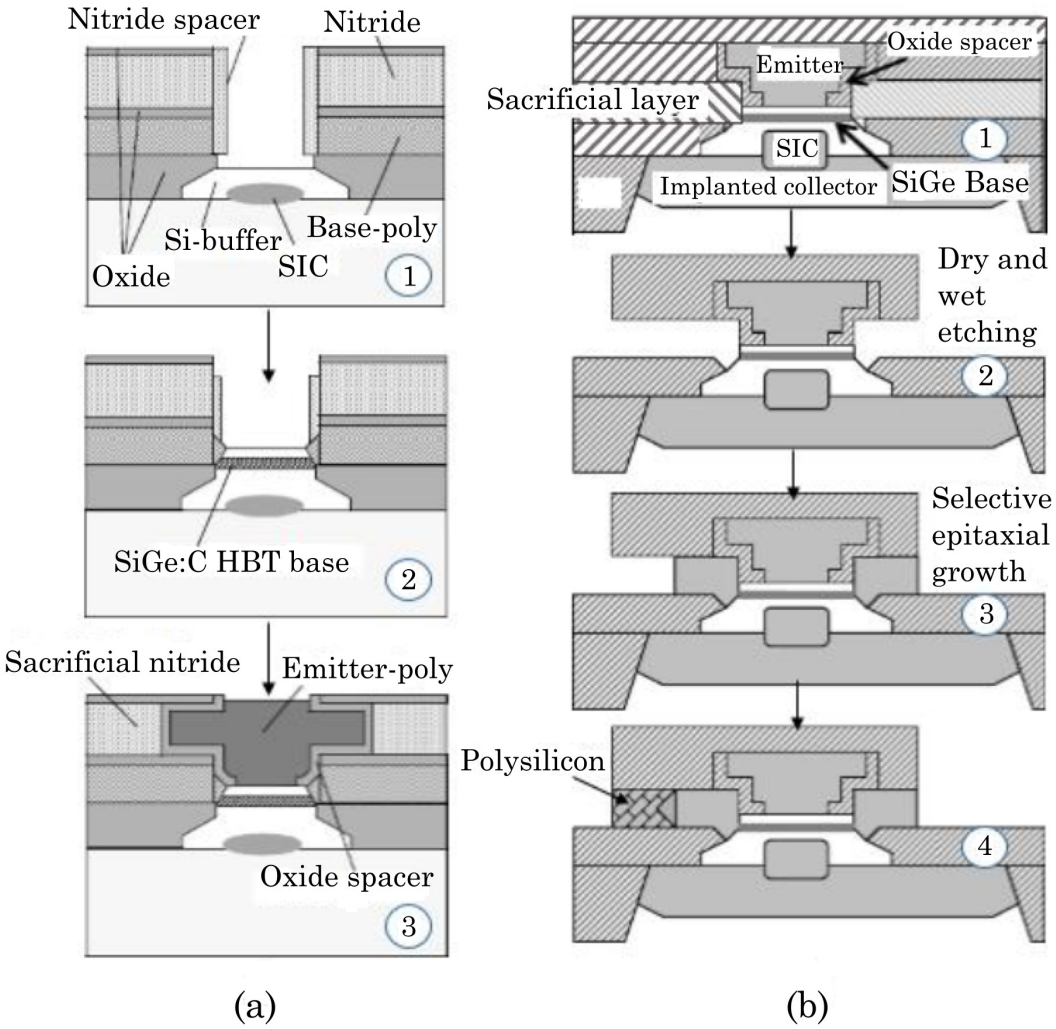


Figure 13: Schematic diagram of base link formation [34] (a) before the in-situ arsenic doped emitter, (b) after the arsenic doped emitter [39].

Lateral Base Link Architecture

Here, a lateral base link module is employed to enhance RF performance. Through lateral epitaxial overgrowth, the baseline module connects both the intrinsic and extrinsic base. This is of two types: (a) the lateral base link is formed before the in-situ arsenic doped emitter, and (b) the epitaxial base link is formed after the in-situ arsenic doped emitter (refer Fig.13).

In scenario (a), three lithographic processes are used to define the collector well, emitter window, and emitter patterning. The process begins with the development of implanted collectors, followed by pedestal-oxide, in-situ boron-doped extrinsic poly-base, and deposited nitride stack layers. The emitter window is subsequently opened by etching through these stacks.

Then, nitride sidewalls are deposited to safeguard the poly-base during the removal of the pedestal-oxide layer using wet etching. Then, a silicon buffer layer is formed selectively, followed by collector implantation that is self-aligned. To build a connection between the extrinsic and intrinsic bases, a portion of the nitride sidewalls are removed. During the process of selectively growing the SiGe:C base layer, the intrinsic and extrinsic bases are joined. Next, the sacrificial nitride layer is etched laterally to create the T-shaped emitter layer overlap. After the formation of the spacers and the in-situ As-doped emitter layer, this layer is removed using CMP outside the HBT areas. The architecture introduces an innovative concept that employs the lateral base connection between the extrinsic and intrinsic bases. However, the architecture overcomes the fabrication difficulties described in [40].

- The shielded nitrides are partially removed in order to open the lateral link areas between poly-bases and the intrinsic collector. This critical step is governed by the etching time; as a result, this step is constrained under industrial fabrication conditions.
- Due to the complexity of base link development during the B-doped SiGe:C layer, the base link is faceted (see Fig.13(a)).This yields a large R_B [34]. Even if the architecture was improved in [39], this architecture's normalized R_B was not superior to a standard DPSA-SEG architecture.
- Using chemical mechanical processing (CMP) after emitter formation results further complexity in the fabrication process.

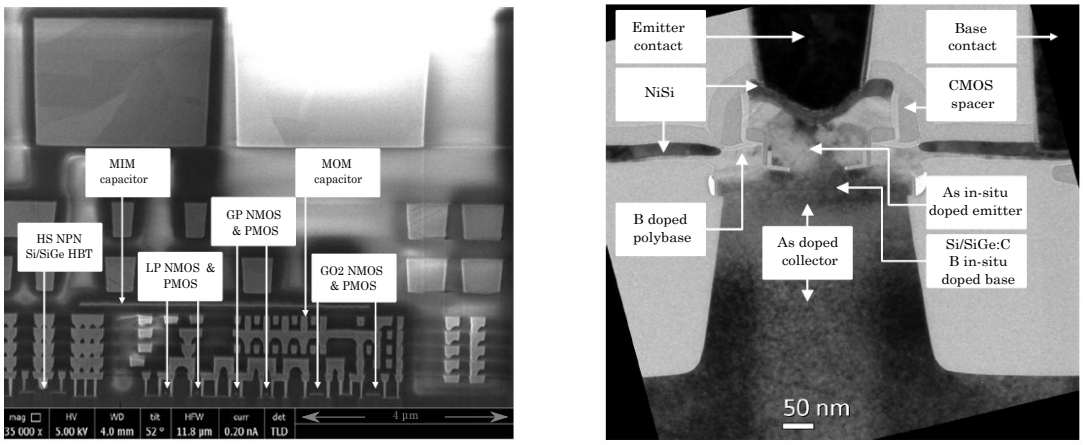
To overcome these limitations, IHP created the EBL architecture in 2011. In the EBL structure, the poly-base layer is replaced by a sacrifice layer (see Fig.13(b)). As a result, the development of B-doped SiGe:C epitaxy and emitter module is simplified. The extrinsic base is formed by the B-doped epitaxy that is performed in-situ. The base link permits the combination of a very low R_B and a lower C_{BC} in comparison to typical DPSA-SEG systems. This architecture achieves f_T and $f_{MAX}$of 310 GHz and 480 GHz, respectively.

As most of the works discussed here will be related to ST Microelectronics, complete developments of SiGe BiCMOS technology at STM are discussed in the next section.

Developments at ST Microelectronics

The well-known DPSA-SEG architecture [39] is the most popular architecture among the devices obtained at companies like STM, IFX, and Hitachi for its record electrical performance.

BiCMOS9MW, the first BiCMOS technology of STM using the DPSA-SEG, resulted in a f_T of 230 GHz and f_{MAX} of 280 GHz. Even more, the state-of-the-art architecture recorded f_T, f_{MAX} of 320 GHz and 370 GHz mapped in 55 nm SiGe BiCMOS technology. The SEM image of the device is shown in Fig.14 [33]. Further a record f_T of 410 GHz and at cryogenic temperature f_T of 640 GHz demonstrated at STM by optimizing process thermal budget [41]. The evaluation of device technology and corresponding f_T, f_{MAX} is shown in Fig.15. As a leading semiconductor company at ST SiGe HBTs are integrated into faster and denser CMOS nodes. The detail of the impact of CMOS nano-scale constraints are discussed in [40]. A DPSA-SEG architecture without STI between the collector and base contact and implanted collector structure resulted in 330 GHz f_T but with a f_{MAX} lower than targeted. The lower f_{MAX} obtained is due to large R_{BX} which is main limitation of DPSA-SEG architecture. Therefore several architectures were explored, and fully self-aligned (FSA) selective epitaxial growth (SEG) featuring epitaxial extrinsic base-isolated from collector (EXBIC) is one of the best to be integrated to 28 nm FDSOI CMOS technology targeted to feature a f_T of 400 GHz and f_{MAX} of 600 GHz in this node [40].



(a) (b)

Figure 14: (a) SEM cross-section of BiCMOS055 showing active devices (both MOS and SiGe) and BEOL up-to Metal layer 8 (b) TEM cross-section of a high speed SiGe HBT with dimension $0.1 \times 4.9\ \mu m^2$ [33].

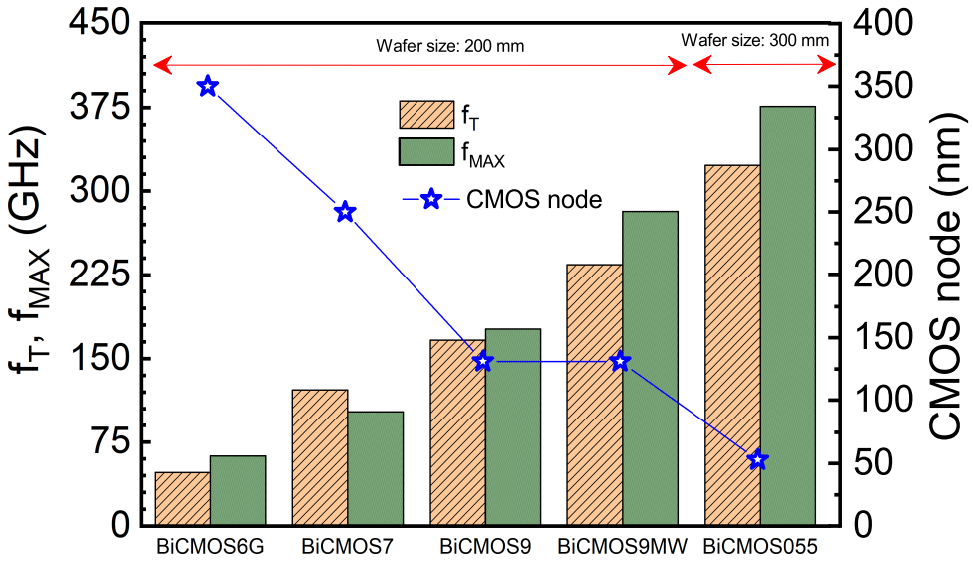


Figure 15: Evaluation of frequency performance at STM's BiCMOS technology [40].

SiGe on SOI Substrate

All devices discussed so far are conventional vertical devices on a bulk substrate. SiGe HBTs on SOI are very much suitable for mixed-signal RF applications. The use of SOI substrates reduces parasitic capacitance, protects the device from latch-up, and has the potential to reduce crosstalk. SiGe HBTs fabricated at IBM, STM, IHP, and TowerJazz are built on thick SOI, and the architectures can be classified into two categories, in one case at STM [42], IBM [43],[44], and TowerJazz [45] the SiGe HBT module is floated on the buried oxide (BOX), or the collector is on top of the BOX, whereas in the second case at IHP [46] the collector is a high-dose implantation substrate below the BOX (refer Fig.16).

As shown in Fig.16(a) at IBM, the concept of a fully depleted collector was used, keeping the SiGe base a vertical profile [43]. The poly-emitter SiGe device doesn't need deep trench isolation (DTI), and the sub-collector and the whole operational region of the device are floated

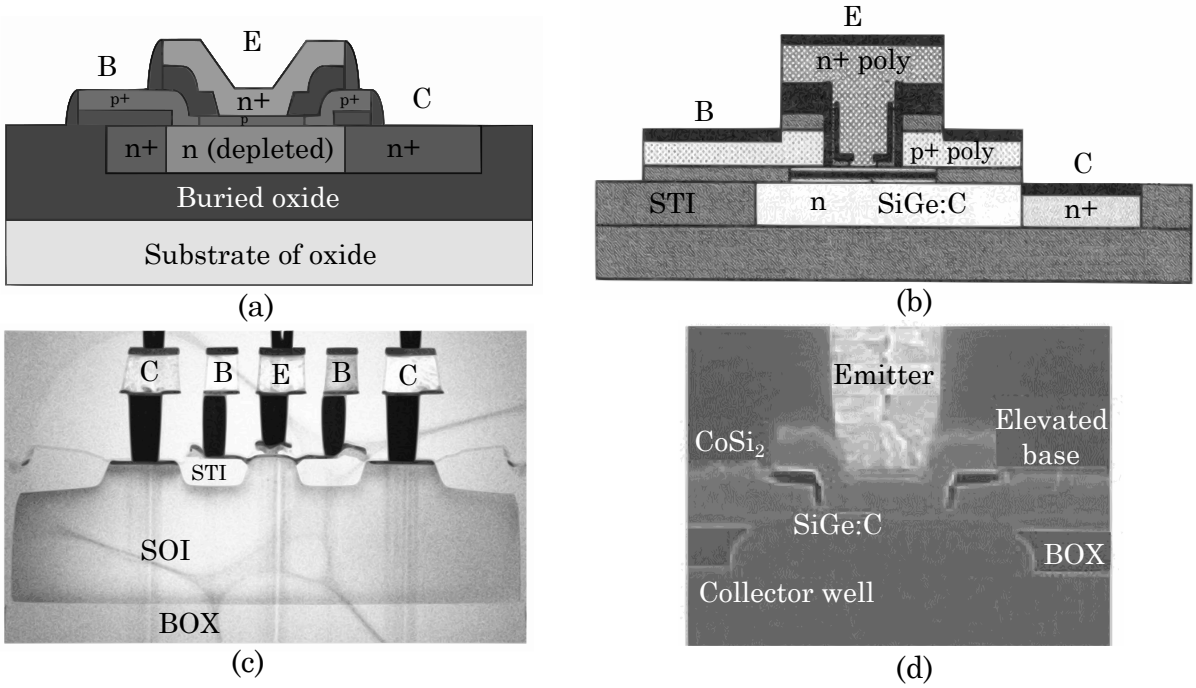


Figure 16: The cross-section of SOI SiGe HBT architecture at (a) IBM [43] (b) STM [42] (c) Tower [45] (d) IHP [46].

on an SOI. Following the same approach, a compact self-aligned base-emitter structure with SEG (refer Fig.16(b)) was developed at STM. No longer the STI is used to isolate the extrinsic base and collector. The maximum performance achieved by this device are f_T, f_{MAX} of 86 GHz and 149 GHz [42]. The performance was not satisfactory as the collector resistance was quite high due to thin-film structure (< 120 nm). Using a thick SOI FSA-NSEG architecture (refer Fig.16(c)) f_T, f_{MAX} of 79 GHz and 280 GHz are achieved at TowerJazz [45]. However, as already introduced in previous sub-section at IHP, the EEB architecture was integrated into SOI using a novel concept. As shown in Fig.16(d) an opening in the BOX is made, and a heavily doped collector is formed, resulting in reduced collector resistance, but due to the opening in BOX low collector substrate capacitance (C_{CS}) offered by BOX architecture can't be achieved. However, this approach opened a new way of integrating SOI with CMOS leading BiCMOS technology, the performance achieved by the device is 220 GHz in f_T and 230 GHz in f_{MAX} [46].

SiGe Device Performances at Different Companies and Future Prediction

To meet the requirements of mm-wave applications, the research on SiGe was carried out extensively at various companies and research labs, a state-of-the-art f_T/f_{MAX} by TowerJazz, Infineon, ST, NXP, Imec, Global foundry (GF), IHP shown in Fig. 17 [47], [8].

The international roadmap for devices and systems (IRDS) predicated that the RF performance will improve significantly with industries' aggressive scaling of the device dimensions [48]. The futuristic values of the performance matrices are shown in Fig. 18. It can be observed that based on the progress in technology, SiGe can overcome the THz limit in near future.

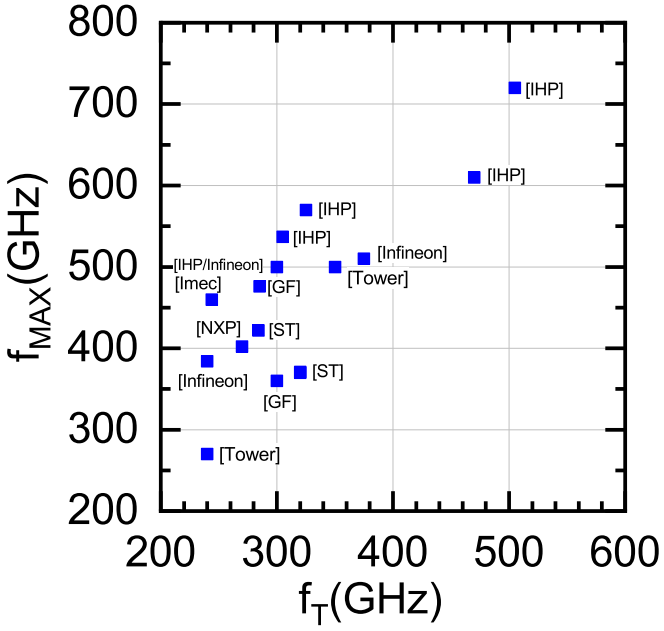
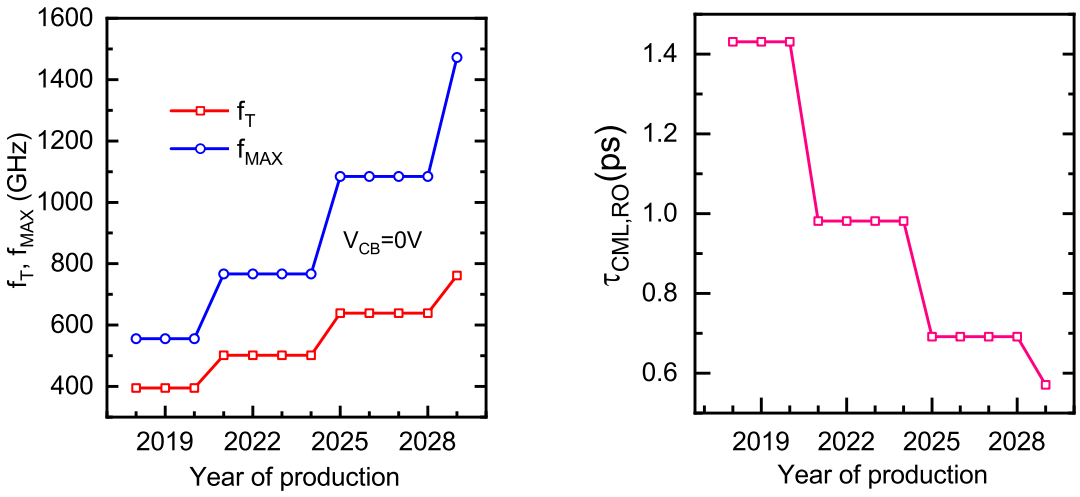


Figure 17: Peak f_T and f_{MAX} demonstrated by various semiconductor companies [47], [8].



(a) (b)

Figure 18: Roadmap of (a) f_T, f_{MAX} and (b) ring oscillator delay $\tau_{CML,RO}$ prediction of SiGe HBTs [48].

Thesis Organization

This dissertation is about the verification of high frequency (HF) characterization results through TCAD & EM simulations. Proposals of new device architectures to achieve THz frequency performance, and analysis of their physics [49].

Basically, this thesis is concerned with the characterization, design, and optimization of high-speed, low-power SiGe HBTs. As we keep on achieving higher and higher frequency performance devices, one of the challenges is the accurate characterization up to high frequency (> 100 GHz). In chapter 2, a systematic method for verifying high-frequency measurement (up to

500 GHz) of silicon germanium heterojunction bipolar transistor (SiGe HBT) is proposed. First of all, the method involves an accurate estimation of the effects of the passive environment on the overall measurement by a detailed electromagnetic (EM) simulation. This ensures that the complete measurement environment like probes, pads, and access lines, and the appropriate layouts are precisely included in the EM simulation framework. In order to additionally include the active device like *SiGe* HBTs, a technology computer aided design (TCAD) tool is used to simulate the device *S*-parameters. TCAD simulation results are fed into an EM-plus-SPICE simulation framework to emulate a complete on-wafer measurement environment. The final simulation results show an appreciable correlation with the on-wafer measurement data up to 500 GHz. Further, the need for proper calibration and de-embedding in high-frequency characterization are emphasized by investigating the *S*-parameters corresponding to a narrow-band amplifier at 170 GHz suitable for G-band radar applications.

In chapter 3, a novel 3D nano-wire-based SiGe HBT device is proposed for the first time. The overall purpose is to predict RF performance of a nano-wire device when the lateral parasitics of the STMs B55 device discussed in chapter 2 are minimized, keeping the vertical doping profile the same. The challenge with device fabrication and the solution to overcome the issues are discussed elaborately. Effect of variation in collector, emitter, and base dimension on figure of merits like f_T, and f_{MAX} are demonstrated. A single nano-wire device can't deliver sufficient current for microwave applications, so using the device a nano-wire array is also computed. The effect of device-to-device separation gap on RF performance is discussed. The proposed device predicts a f_{MAX} above 900 GHz.

In chapter 4, a new asymmetric lateral SiGe HBT on SOI is proposed. This asymmetric structure allows one to modulate the carrier densities in the collector region by the application of substrate bias that causes significant improvements in the device performance. The open base breakdown voltage can be tuned from 2.2 V to 3.6 V, and the transit frequency (f_T) is improved up to 90% by varying the substrate bias. The bias-dependent variation in transit time is elaborately discussed using regional analysis. This work, for the first time, demonstrates an f_{MAX} of 2.7 THz achievable by tuning the substrate bias in an asymmetric SOI lateral SiGe HBT.

In chapter 5, the conclusion and discussion of the work presented in this dissertation are briefed. Further ideas to model the SOI-based device using the existing standard HiCUM model are discussed.

Chapter 1

TCAD and EM Co-Simulation Method to Verify SiGe HBT Measurements up to 500 GHz

1.1 Introduction

The transistor plays a crucial role as an active component in most high frequency electronics circuits. Hence, to realize a successful working RF circuit, the device must be carefully fabricated, characterized, and modeled.

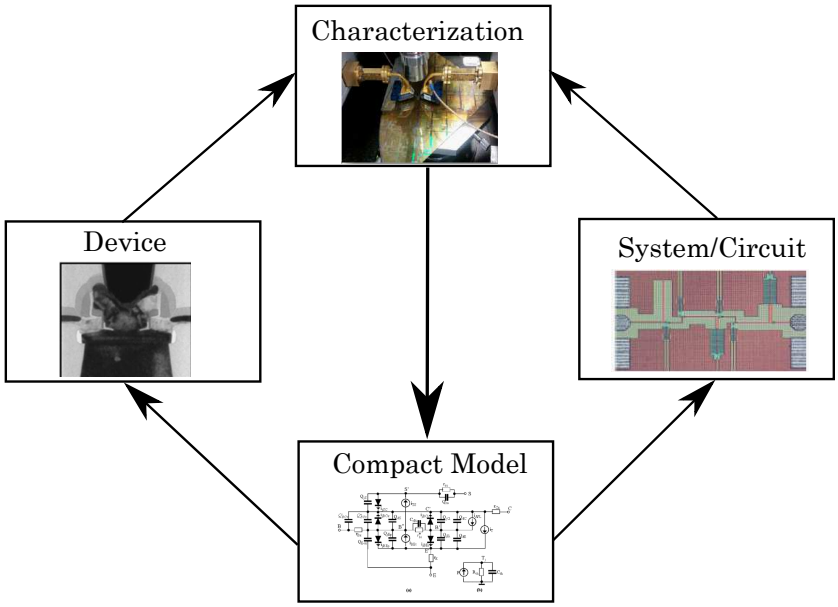


Figure 1.1: Importance of accurate characterization

The circuit design at high frequency for sub-THz or automotive radar applications requires accurate model cards for both passive and active devices [50]. Hence, there is always a compulsory demand for precise characterization and modeling of devices in the sub-THz frequency range to optimize the desired circuit performance and minimize the number of design-to-fabrication loops as shown in Fig.1.1. A complete compact model validation strategy presented by Ardouin et al. [51] have demonstrated that wrong or non-extracted non-quasi-static (NQS) model parameters can induce a variation of more than 5 dB for the estimation of a 220-GHz LNA (Low Noise Amplifier) power gain. Indeed, this NQS parameter extraction is still not

correctly performed by manufacturers since measurements are limited in frequency range as well as in accuracy. Hence it is essential to understand the *S*-parameter characterization in high frequency range and verify the models in CAD tools in that range [52].

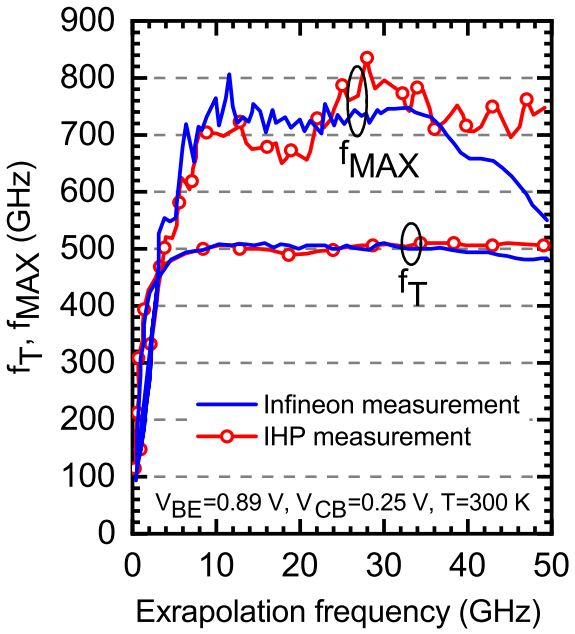


Figure 1.2: Dependence of f_T and f_{MAX} on extrapolation frequency for the same HBT measured at IHP and Infineon. SOLT calibration with an impedance standard substrate (ISS) - adapted from [36].

Although the cutoff frequency (f_T) of a transistor is the most cited performance metric because it is easy to characterize, it is the maximum oscillation frequency (f_{MAX}) that best describes the potential of a transistor in RF, high speed, and mm-wave applications [53], [54]. For advanced SiGe HBTs, the f_{MAX} is typically estimated from the product of the square root of unilateral power gain ($\sqrt{U}$) $\times$ frequency [55]. According to standard definition unilateral power gain:

$$U = \frac{|Y_{21} - Y_{12}|^2}{4(\mathrm{Re}[Y_{11}]\mathrm{Re}[Y_{22}] - \mathrm{Re}[Y_{12}]\mathrm{Re}[Y_{21}])} \tag{1.1}$$

$$= \frac{|Z_{21} - Z_{12}|^2}{4[\mathrm{Re}(Z_{11})\mathrm{Re}(Z_{22}) - \mathrm{Re}(Z_{12})\mathrm{Re}(Z_{21})]} \tag{1.2}$$

In fact, the power gain U is very sensitive to any error in the real part of the Y or Z - parameters of the device, especially Y_{12} which is challenging to measure [53]. As an example shown in Fig. 1.2, in the framework of the semiconductor companies' competition, which is based on the figure of merit (FOM) measurement, is the measurement of silicon HBT technology having f_{MAX} world record of 720 GHz [36]. This one was measured in two different laboratories, giving similar results below 40 GHz and showing a large deviation above 40 GHz. This is due to the fact that, as the transistors get higher in frequency, weaker and weaker elements have to be measured. Thus, the impact of the probe couplings becomes quasi-dominant over the intrinsic device at a very high frequency. The next section presents a detailed discussion of published works and various factors affecting high-frequency characterization.

1.2 Literature Survey

Characterization above 110 GHz is still a challenge, and there is no established de-embedding method above this range. In fact, a very few research groups have published *S*-parameter measurement results above 110 GHz [56], [57], [50], [58], [59], and there are numerous practical reasons for the error, which are discussed here.

Raw measurement (particularly the RF ones) results are always affected by errors introduced by the measurement setup. These errors can be of different kinds like random errors, errors due to non-linearity, drift errors, and systematic errors [60].

- Random errors: are caused by lack of repeatability during characterization; they can be described statistically but not by systematic correction. The source of this type of error is the instrument noise, electro-magnetic (EM) interference due to different probe positioning during measurement, and electrical noise from different components of the vector network analyzer (VNA), in particular from the load oscillator. The electrical noise can be removed by increasing the power level, and other random errors can be corrected if their statistical average is zero, which can be learned from multiple measurements.
- Error due to non-linearity: While characterizing non-linear devices like bipolar transistors, spurious harmonics may get generated when input power is high. The linearity in power gain starts degrading with an increase in input power, and after certain point, output power goes into 1-dB compression, and gain flattens. The device response is no more linear and it produces distortions. Hence, there always needs to have a trade-off between random errors, and non-linearity that can be tolerated.
- Drift errors: These errors occur due to characteristics of the instrument that change with time. For example, change in contact resistance during characterization and thermal dilation of RF cable. A controlled temperature during measurement avoids such errors. This issue can be corrected to some extent by re-calibration.
- Systematic errors: These are produced due to imperfections in the instrument and connections; they are consistent and repeatable. Since these are predictable and do not change with time, a systematic correction can be applied by subtracting the error terms from the raw measurement vectors. This correction of errors is called the method of calibration and is regularly needed.

Hence, *calibration* is the method by which we can determine the error terms of an error model. The goal is to remove all these errors and to obtain the *S*-parameters at a well defined location called $reference\ plane$, which can be either at the probe tips during $off-wafer$ calibration or closer to the device under test (DUT) for $on-wafer$ calibration. In fact, close to approximately correct characteristics of the target DUT can be extracted with additional $open-short$ or $short-open$ de-embedding. More details about different calibration techniques are discussed further.

High frequency measurements require careful calibration procedures. Off-wafer calibration, as shown in Fig.1.3 & 1.4, is more popular in the production environment [61]. However, since the calibration substrate (impedance standard substrate) is different from that of the desired wafer to be characterized, the off-wafer calibration leads to systematic errors, which is correlated to different probe-to-substrate electromagnetic coupling [62], [63].

Similarly, as different probes (ref. Fig.1.8) are used to measure different frequency bands, discontinuity appears in the measured data due to change in RF probes if the probe's cross-talk is not fully corrected by the calibration procedure. A detailed explanation of the work has been reported in [63], [64].

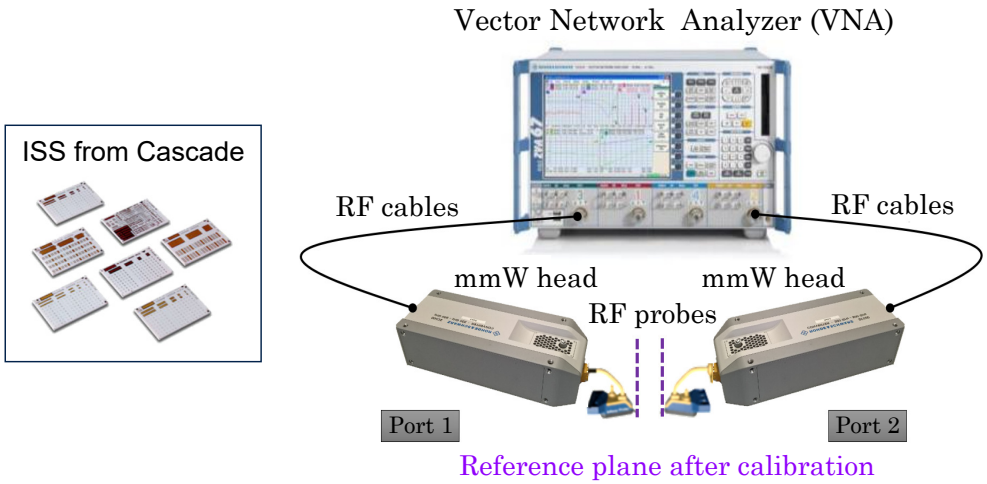


Figure 1.3: Off-wafer Calibration: depicting the reference plane, which is at the probe tips. In the left side the impedance standard substrates from cascade are shown.

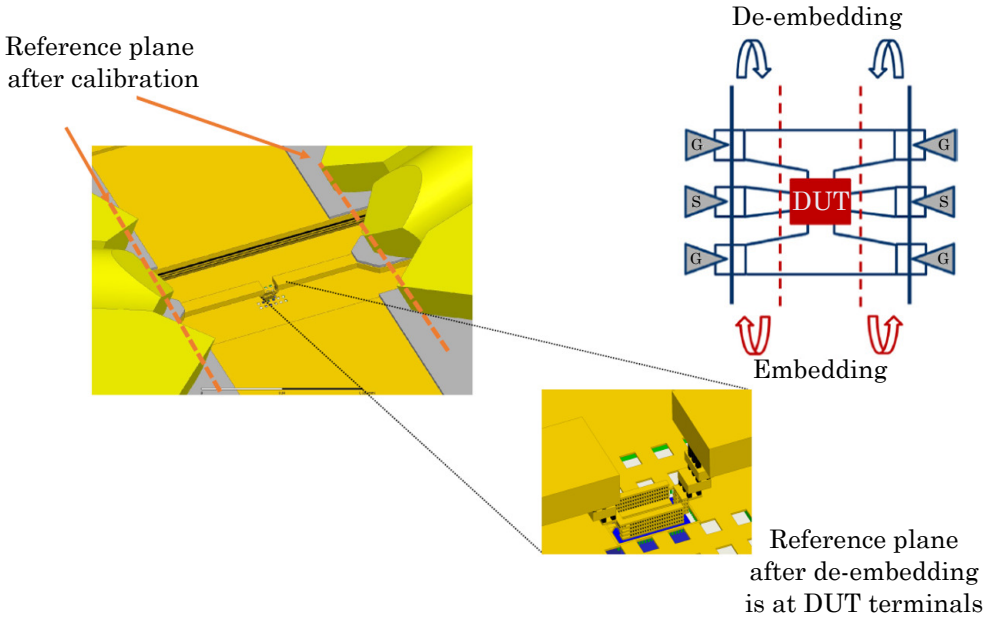


Figure 1.4: Off-wafer Calibration with additional de-embedding, depicting the reference plane shifted closer to the DUT.

On-wafer calibration, as shown Fig.1.5 is a better choice for high frequency (HF) measurement because it avoids some of the above mentioned limitations observed in off-wafer calibration. Thru-Reflect-Line (TRL) calibration algorithm is the method of choice [56] since it provides a well-defined reference plane (ref. Fig.1.5) and does only require a low knowledge of the standards (geometrical dimensions and characteristic impedance extraction), that indeed gives better accuracy in the high frequency domain. Nevertheless, this method is not a good choice for the lower frequency since it would require extremely long lines, and SOLT/LRM method are more suitable in this low range of frequency.

But during TRL calibration, while probing the Thru and Line, it is necessary to shift one probe, which may add errors in the measurement, particularly in the case of manual probe stations [50], [65]. The drawback of the on-wafer calibration is the need for a specific design of test structure which can be time-consuming. This design requires engineering effort and expertise.

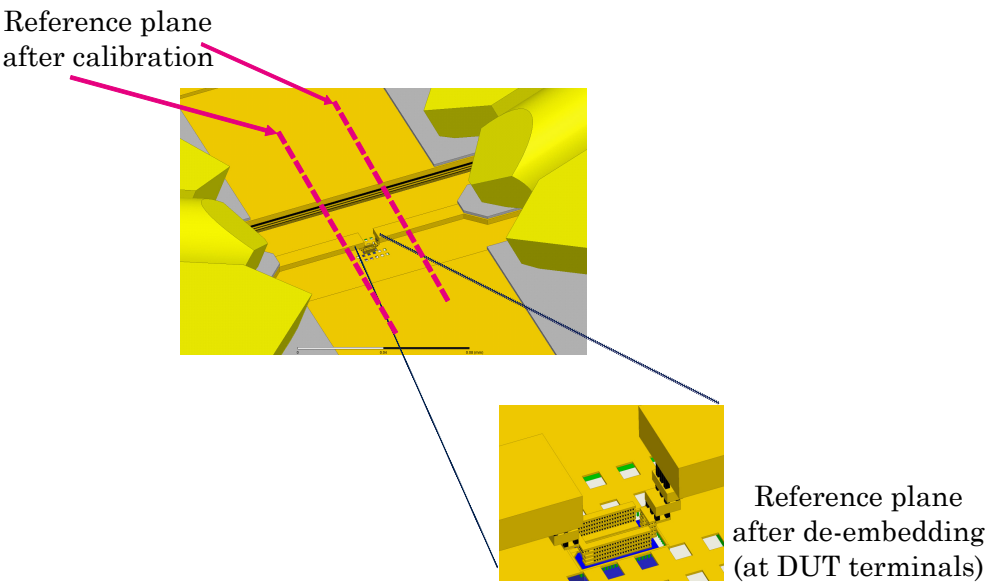


Figure 1.5: On-wafer calibration with additional de-embedding, depicting the reference plane shifted closer to the DUT.

For example, the line lengths need to be chosen properly to cover the required frequency band; the geometry of the line needs to minimize the loss and allow only one mode to propagate; the design of the pad itself needs to be optimized to limit the probe-to-substrate coupling; the pad-to-pad distance needs to be sufficiently long in order to ensure a reduced cross-talk, and finally one has to keep enough space between the structures to avoid coupling [63], [66], [67]. Advanced technologies have a high cost per square millimeters; hence the devices need to be placed close by to reduce the cost, but different laboratories have shown that the measurement results can get affected because of the above reason of close placing of the devices. Hence there needs to have a trade-off between cost due wafer area consumption, and error that can be tolerated [68].

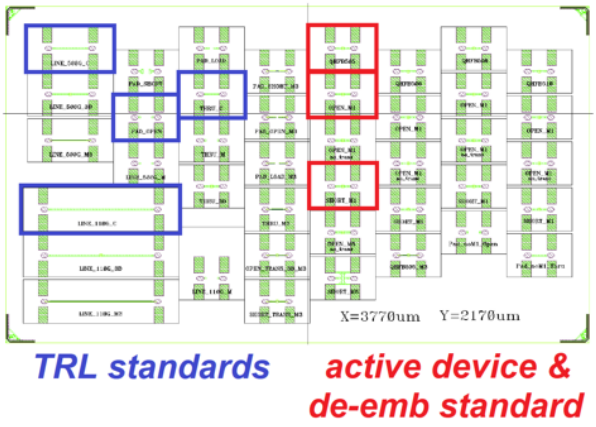


Figure 1.6: Floor plan of the active and passive DUTs placed on wafer.

The results presented above have been measured for wafers with the optimized placement of DUTs (shown in Fig.1.6). As explained in [69], the test structures were placed in staggered rows, and a continuous ground plane and pad shield was implemented in this chip's layout.

Another subject of inaccuracy is the probe coupling. From Fig.1.7, one can observe that each

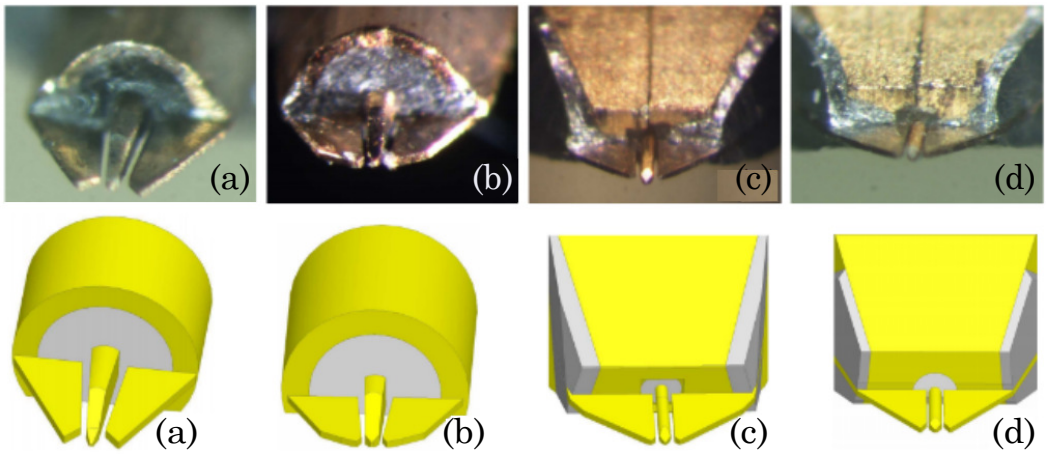
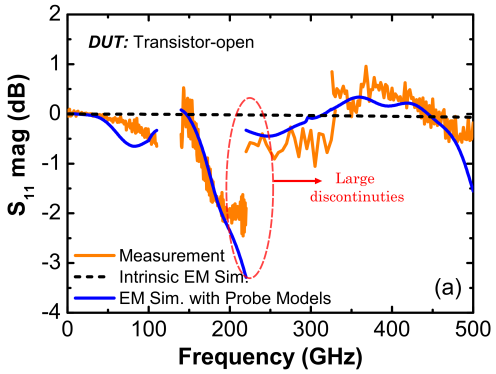
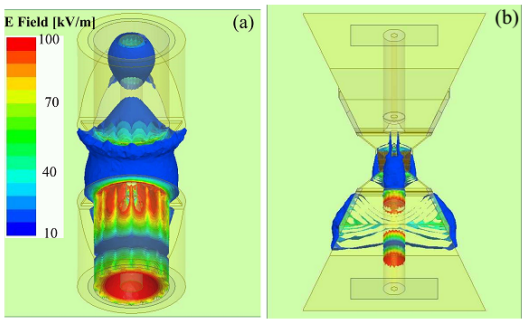


Figure 1.7: Probe geometry used for different frequency bands (a) 1-110 GHz (b) 140-220 GHz (c) 220-325 GHz and (d) 325-500 GHz both real and their corresponding mimicked EM-simulated ones respectively.

frequency band is measured with the probe having a different geometry. For example, the probes below 220 GHz have a much larger geometry than the one above. The reader should focus on the output of the micro-coax, which is much larger for WR5.1 (140-220 GHz) probe compared to WR3.4 (220-325 GHz) probe [60]. Hence, as shown in Fig.1.8 the EM wave coming out of the probe will spread much more in the case of the probe at low frequency compared to the probes at high frequency. From a measurement point of view, one can observe large discontinuity at each frequency band while the intrinsic simulation of the device shows continuous behavior [64].



(a)



(b)

Figure 1.8: (a) Discontinuity observed at band change for a transistor open DUT (b) electric field distribution (top views) in the same transistor-open DUT at 220 GHz using two probe models.

The effect of the neighboring structures/DUTs for passive devices is shown in Fig.1.9 with the help of EM simulation. By considering the neighboring structures' effects in Fig.1.9(b) instead of only the DUT as in Fig.1.9(a), the simulated fluctuations are followed as in measurement, which hints towards the need for careful design of adjacent DUTs. The procedure to verify the above discussed anomalous deflections during characterization is studied in this chapter.

In order to check the trend, consistency, and accuracy of the high-frequency measurement results, one has to compare the data with those obtained from calibrated simulation. One can think of options like TCAD simulation or electrical compact model simulation for this purpose.

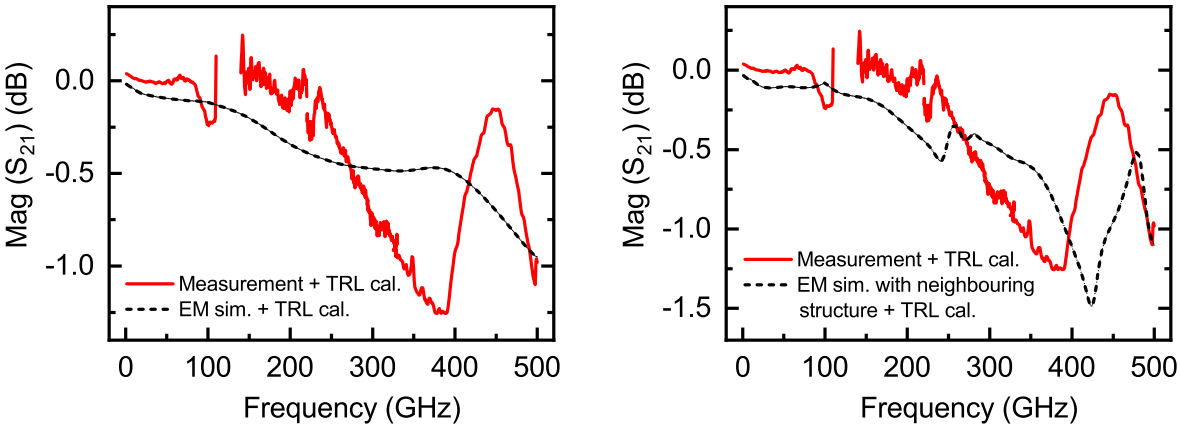


Figure 1.9: Effect of neighboring DUTs on the measurement characteristics: (a) EM simulation of an isolated DUT, (b) EM simulation of DUT with its neighboring structures

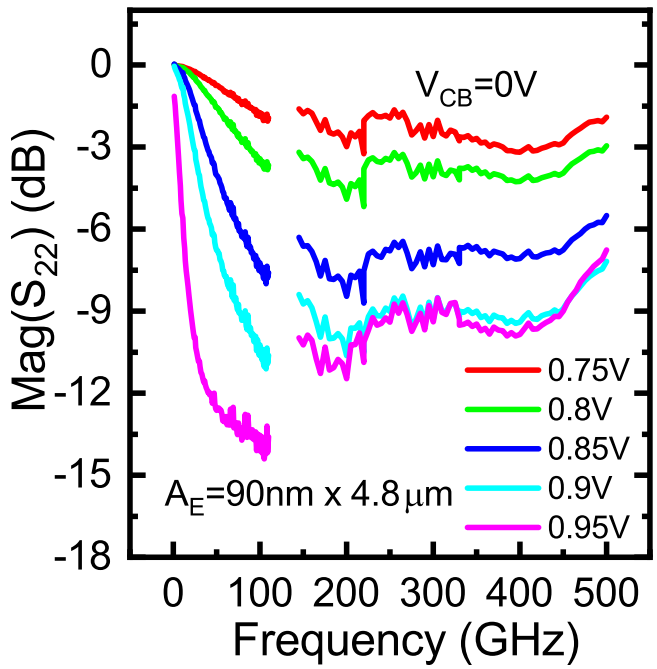


Figure 1.10: Frequency dependent magnitude of S_{22} with V_{BE} variation at V_{CB}=0 V.

Since the compact model is not verified against the measurement at such high frequency (for example, non-quasi-static effects (NQS), etc.) TCAD simulation has been chosen as a reference in this work [70], [49], [71].

In this work, using Sentaurus TCAD and Ansys-HFSS electromagnetic (EM) simulations, we explore and verify the high frequency measurements of SiGe HBTs. As a motivation for such an investigation, Fig.1.10 compares the measurement and calibrated TCAD simulation of frequency-dependent S_{22}-parameter corresponding to a SiGe HBT from ST Microelectronics B55 technology at different bias in the main operating regime. It is observed that in the lower frequency range (below 110 GHz), the measurement and simulation are in good agreement. However,

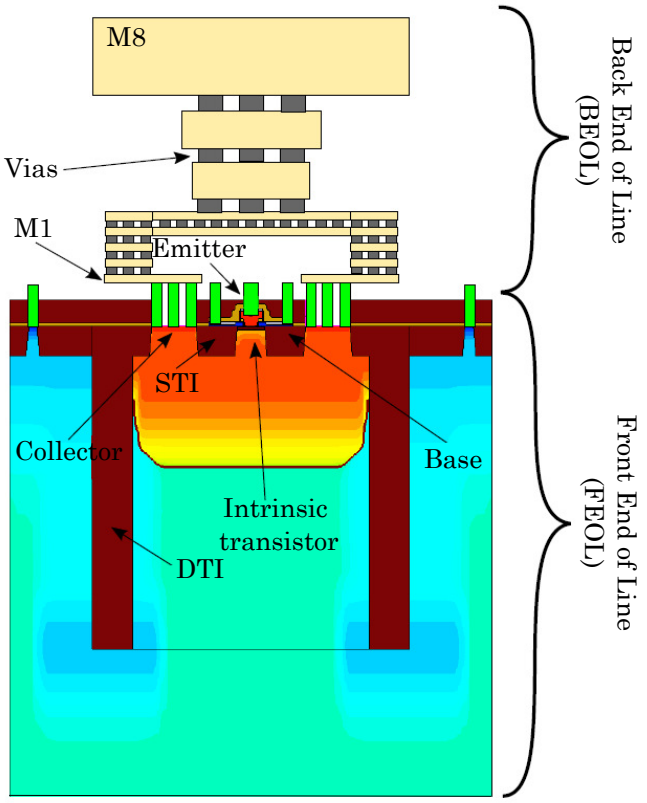


Figure 1.11: Full structure of the device FOEL and BEOL together making the complete device.

some discontinuities are observed in the measurement data at the cross-over points (110 GHz, 220 GHz, and 325 GHz) of different frequency bands; which can be correlated to couplings involving probes, layout of the pad and adjacent structures as explained in previous paragraph. Also, the characteristics in the upper frequency range tend to deflect from the TCAD simulations. This motivates us to emulate a more realistic measurement environment. To do so, we combine our TCAD simulation results with EM simulation that includes the back-end-of-line (BEOL) along with the probes. A reference for the front-end-of-line (FEOL) along with BEOL is shown in Fig.1.11 and the detailed explanation about how the probe included into the simulation set-up is discussed in section 2.5 of this chapter. This TCAD plus EM simulation data together will be henceforth called as *Virtual measurement* throughout the chapter. After that, we follow all the on-wafer calibration and de-embedding procedures on this virtual measurement data. Finally, for verification, both TCAD and de-embedded virtual measurement data are compared with actual measurement results. Also, a circuit-level verification is carried out by designing a G-band single-stage amplifier based on TCAD as well as de-embedded virtual measurement data.

In section 2.3, a brief description of the measurement setup and procedures are provided. Section 2.4 of this chapter provides a comprehensive discussion on the calibration of TCAD simulation. The methodology of combining TCAD and EM simulation to produce virtual measurement results is explained in section 2.5. In section 2.6, the assessment of virtual measurement data is carried out in terms of some figures-of-merit (FoMs) of a narrow-band 170 GHz amplifier circuit. Finally, the work is concluded in section 2.7.

1.3 Actual Measurement Procedure

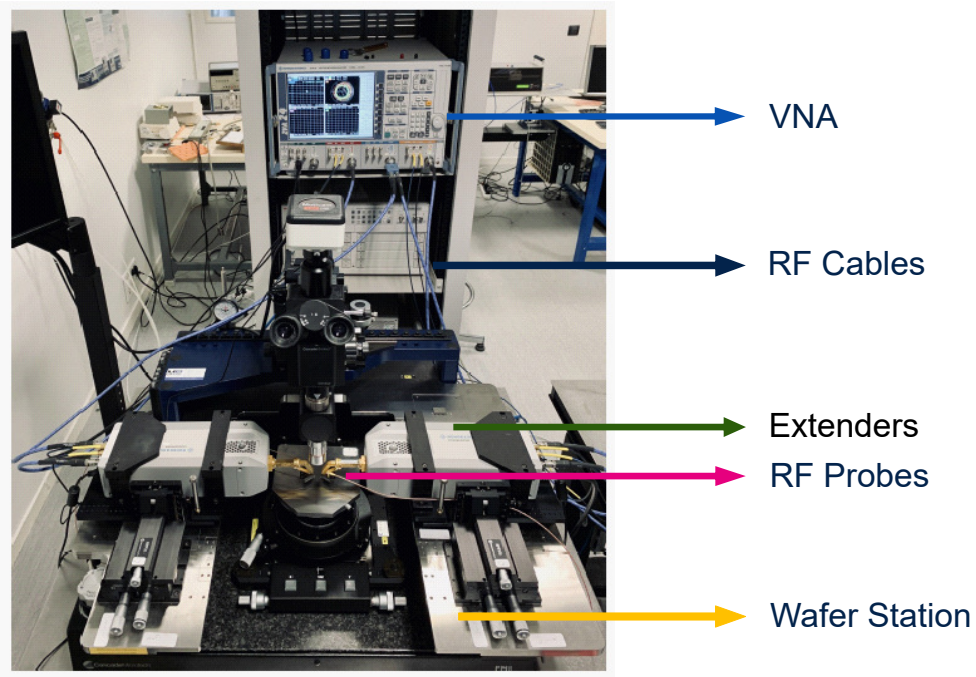


Figure 1.12: Measurement bench set-up for 220-500 GHz bands.

The *S*-parameter measurements are carried out up-to 500 GHz for base-emitter voltage in common emitter configuration varying from 0.7 V to 1 V at V_{BC}=0 V. In order to measure from 1 GHz to 500 GHz, four different measurement benches such as (i) Agilent's E8361A VNA up to 110 GHz using extenders (N5260-60003) above 67 GHz (ii) 140-220 GHz (WR5) (iii) 220-330 GHz (WR3) and (iv) 325-500 GHz (WR2.2) bands with a four-port Rohde & Schwarz ZVA24 VNA coupled with extenders (ZC220-ZC330-ZC500) have been used. A picture for measurement setup above 110 GHz frequency bands is shown in Fig.1.12. The power level is set to approximately -32 dBm using a calibration table provided by the extender supplier in the four bands for measurement of both active and passive elements (the latter are used for de-embedding purposes). Picoprobe RF probes with 50 μm pitch are used for covering the bands above 110 GHz, and 100 μm pitch probes are used below 110 GHz. On-wafer TRL calibration and standard Short-Open de-embedding have been performed. The reference plane of the TRL is set at the top metal and at sufficiently close proximity of the *vias*. The characterized frequency dependent *S*-parameters along with f_T and f_{max} results are presented in Figs.1.13, 1.14 and 1.15.

1.4 TCAD Simulation Setup and Calibration

To verify the measurements discussed in section 2.4, in commercially available Sentaurus TCAD simulator, a SiGe-HBT structure corresponding to ST Microelectronics' s B55 process is developed as shown in Fig. 1.16 and Fig. 1.17. All the dimensions, including the shallow and deep trench isolation, are ensured to be closely identical with those of the B55 process during the device design.

The doping profiles have been reproduced using analytic mathematical models as shown in Fig.1.18 [40], [72]. All necessary physical effects are taken into account in the simulation using

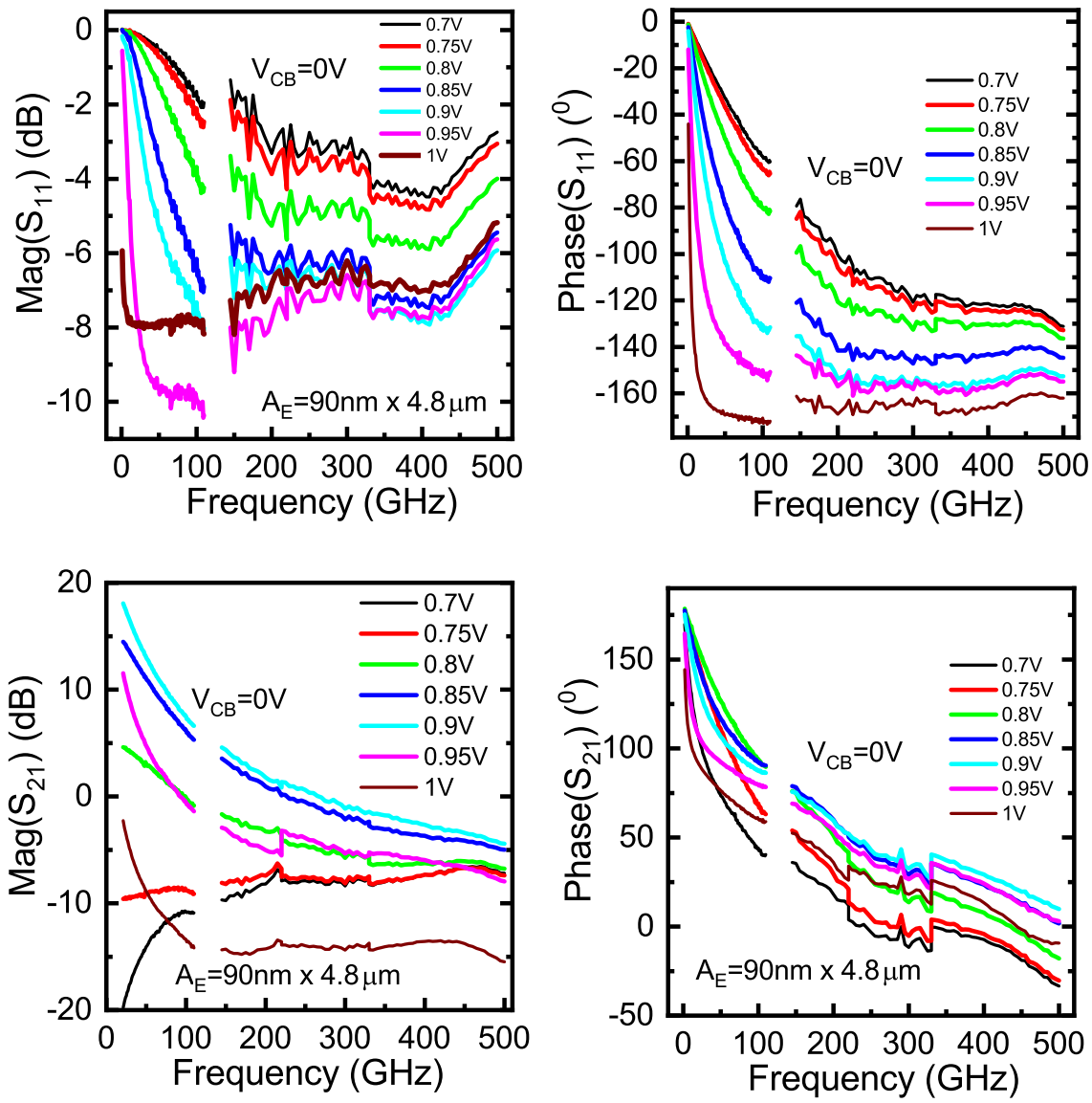


Figure 1.13: Frequency dependent magnitude and phase of S_{11} & S_{12} -parameters at V_{CB}=0 V and V_{BE}= 0.8 V and 0.9 V for the SiGe HBT up to metal-1.

appropriate models, and parameter sets are appropriately tuned from Monte Carlo simulation as discussed in [73], [70].

1.4.1 Device Models Used During Calibration

The following physics based models are included into TCAD $s-device$ deck during simulations:

- Hydrodynamic model
- Band gap Narrowing (BNG)
- Mobility models (like doping dependence and high field saturation)
- SRH and Augur models are also included to account recombination

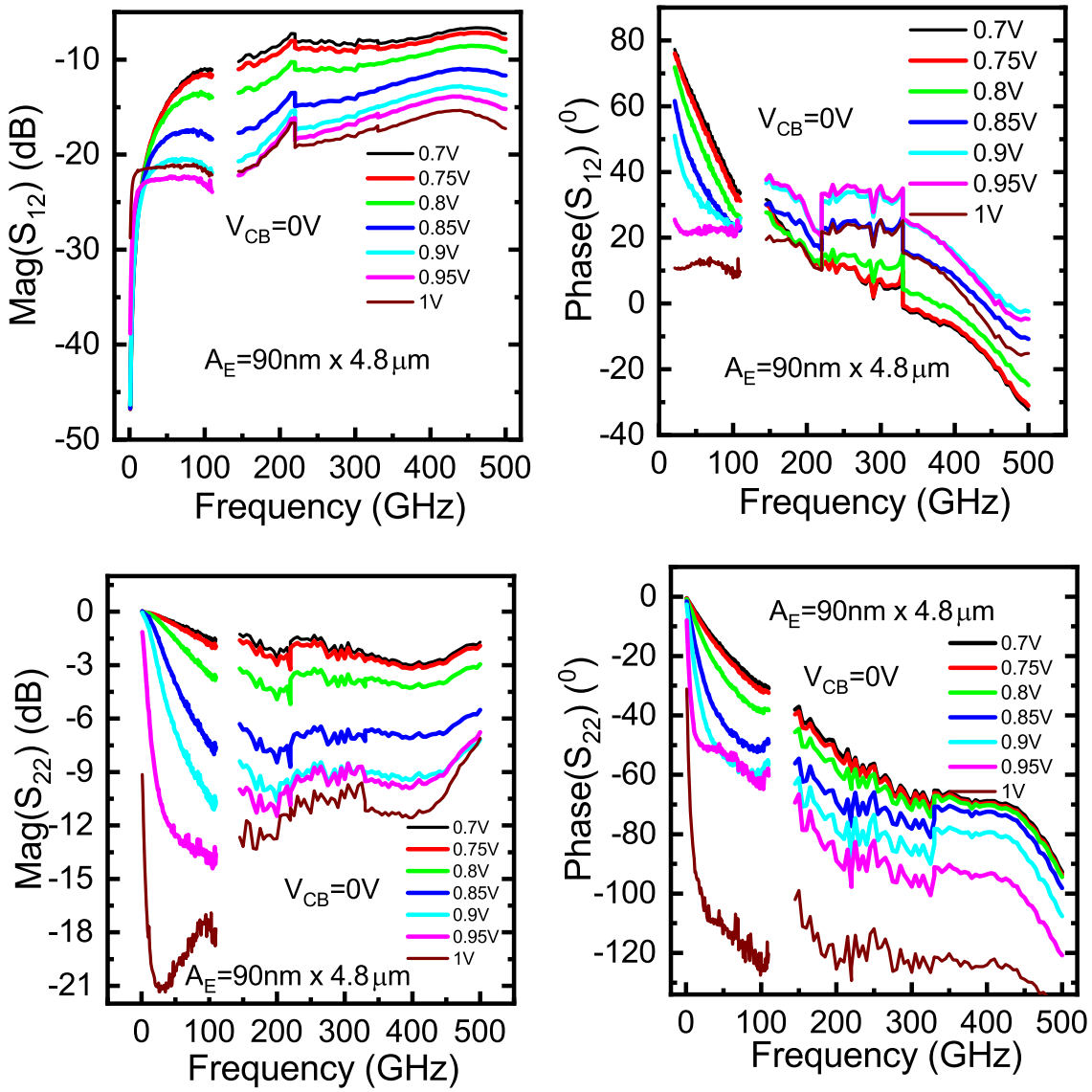


Figure 1.14: Frequency dependent magnitude and phase of S_{12} & S_{22} -parameters at V_{CB}=0 V and V_{BE}= 0.8 V and 0.9 V for the SiGe HBT up to metal-1.

- Lackner's avalanche model for accounting breakdown effects

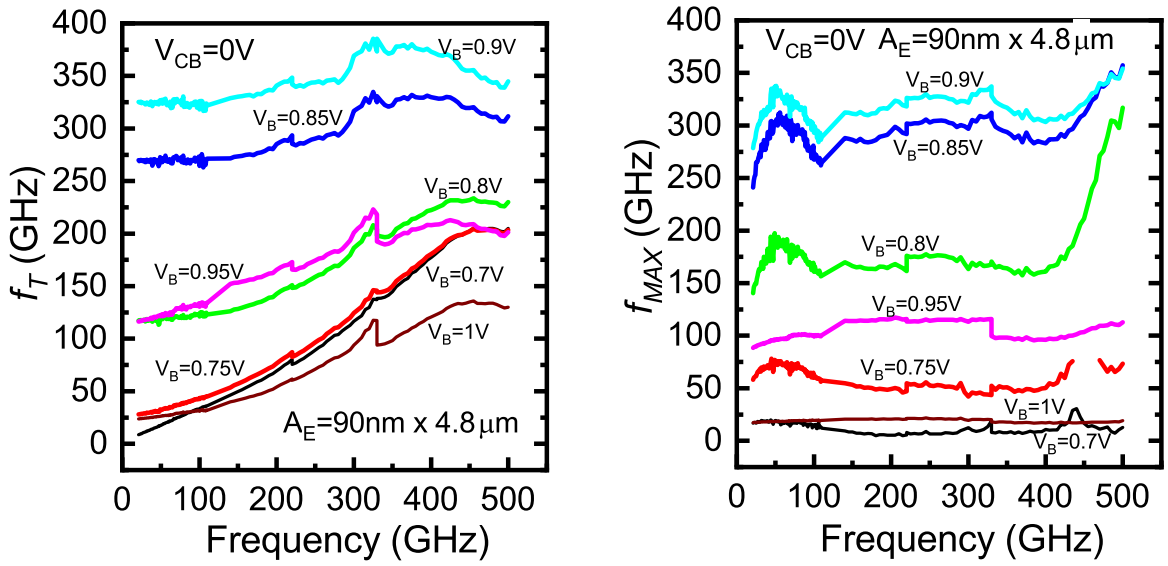
Hydrodynamic Model

Drift-Diffusion model (DD) is considered as the default transport model, wherein the electron and hole current densities (J_n, J_p) respectively, are given by:

$$\overrightarrow{J_n} = \mu_n(n\nabla E_C - 1.5nkT\nabla \ln m_n) + D_n\left(\nabla n - n\nabla \ln \gamma_n\right) \tag{1.3}$$

$$\overrightarrow{J_p} = \mu_p(p\nabla E_V + 1.5pkT\nabla \ln m_p) - D_p\left(\nabla p - p\nabla \ln \gamma_p\right) \tag{1.4}$$

Where μ_n and μ_p are electron and hole mobility, n and p represent the electron and hole densities, E_C and E_V are conduction and valency energy bands, D_n and D_p signify the diffusion



(a) (b)

Figure 1.15: Frequency-dependent transit frequency (f_T) and maximum oscillation frequency (f_{MAX}) at V_{CB}=0 V and V_{BE}=0.8 V and 0.9 V for the SiGe HBT up to metal-1.

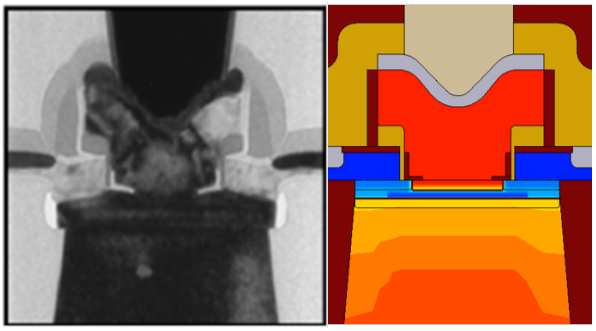


Figure 1.16: TEM picture (left) of *SiGe HBT* and corresponding simulated TCAD structure (right), only the crucial parts of the device is shown in this figure.

constants, k, and T are Boltzmann's constant and temperature respectively with γ_n, and γ_p being the Fermi statistic constants. The first term of both current density(J_n,J_p) equations takes the spatial variation of electrostatic potential, bandgap and electron affinity into account. In the same way the second term includes a gradient of carrier densities and spatial variation of effective masses for both electrons and holes.

In the similar manner J_n, J_p for Hydrodynamic (HD) model is given by

$$\overrightarrow{J_n} = \mu_n(n\nabla E_c + kT_n\nabla_n - nkT_n\nabla \ln \gamma_n + \lambda_n f_n^{td} kn\nabla T_n - 1.5nkT_n\nabla \ln m_n) \tag{1.5}$$

$$\overrightarrow{J_p} = \mu_p(p\nabla E_v - kT_p\nabla_p + pkT_p\nabla \ln \gamma_p - \lambda_p f_p^{td} kp\nabla T_p + 1.5pkT_p\nabla \ln m_p) \tag{1.6}$$

Where most of the symbols has same meaning as described in drift-diffusion mode, additional parameters and their importance is discussed below.

During calibration, hydrodynamic model is chosen over the drift-diffusion model for conduction current density calculation due to its higher degrees of freedom. So in this work for the

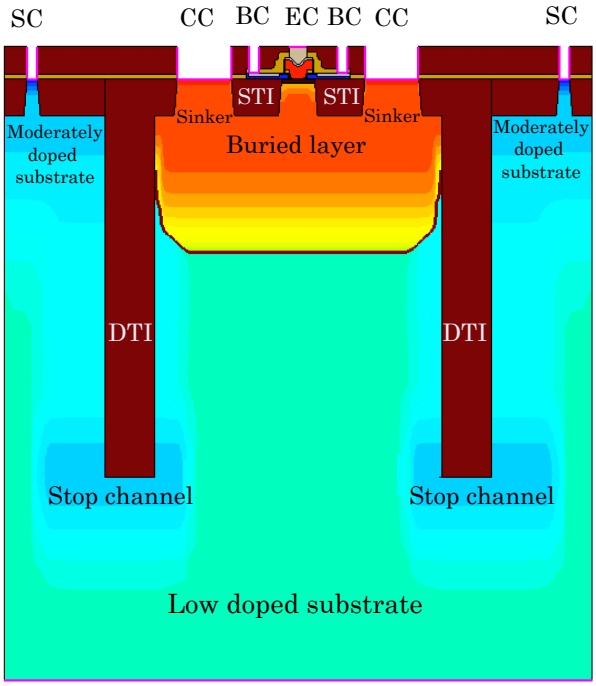


Figure 1.17: Full structure of TCAD simulated *SiGe HBT* corresponding to full B55 device including STI, DTIs.

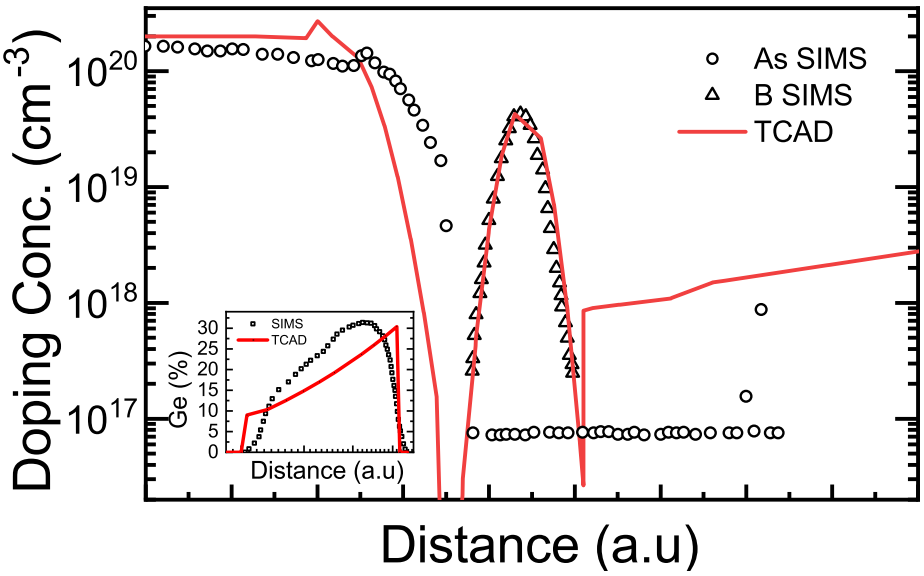


Figure 1.18: Vertical doping profiles: Estimated SIMS profiles of *As* and *B* doping (shown in different symbols) taken from [40] and corresponding analytic profile (line) incorporated in TCAD (arbitrary unit of distance). Inset shows the Ge mole-fraction measured by EDX (symbol) and one used in TCAD (line).

carrier transport, we have used hydrodynamic model that solves the basic drift-diffusion transport equations along with the additional energy balance equations for minority carriers as discussed. It is to be noted that performing device simulation, including standard models with their default parameters, can yield anomalous and un-physical effects. The last terms of (1.5) & (1.6) account for the additional driving forces due to change in effective masses in heterostructure devices so that the force related to the change in band edge energies is included in the valence band and conduction energy gradients. In hydrodynamic model the parameters

such as energy flux coefficient (r), thermal diffusion parameter (f^{td}) and heat flux diffusion factor (f^{hf}) have a significant impact on the results. Considering the suggestion from [74] and [75], the energy flux coefficient and the thermal diffusion parameter are set to 1. The heat flux diffusion factor has to be sufficiently small to avoid negative Early voltage. In addition to hydrodynamic equations, the lattice temperature increase due to current flow of carriers (i.e Self-Heating), is obtained by solving the lattice heat equations [76], [77].

Bandgap Narrowing

It has been seen that the overall effect of the high dopant concentration is to reduce the bandgap. The Slootboom bandgap narrowing model in eq.1.7 with appropriately tuned N_{ref} value is used to account for the effect of carrier concentration at heavy doping [78].

$$\Delta E_{g,app} = \Delta E_{g,o} \times \left[\ln\left(\frac{N}{N_{ref}}\right) + \sqrt{\left(n\frac{N}{N_{ref}}\right)^2 + C} \right] \tag{1.7}$$

default value $\Delta E_{g,o}$= 6.92 meV, N is the impurity concentration, N_{ref}=1.3x10^{17} cm^{-3} and C= 0.5. The origin of $\Delta E_{g,app}$ reduction depends on apparent $\Delta E_{g,app}$ and true band gap Narrowing.

Mobility Models

Mobility (μ) is a function of germanium content, doping, carrier types, and electric field, etc. The core model consists of the standard Caughey-Thomas equation for low field mobility.

$$\mu_{1F} = \mu_{\min} + \frac{\mu_{\max} - \mu_{\min}}{1 + \left(G/G_{ref}\right)^{\alpha}} \tag{1.8}$$

where $G = N_A^- + N_D^+$, for low doping G/G_{ref} tends to zero and μ_{1F} becomes $\mu_{\max}$. For high doping G/G_{ref} infinite and μ_{1F} becomes $\mu_{\min}$.

$$\mu_{High}(F) = \frac{\mu_{Low}}{\left[1 + \left(\frac{\mu_{Low}}{V_{sat}}\right)^{\beta}\right]^{1/\beta}} \tag{1.9}$$

The high field mobility model by Caughey-Thomas which is used here is a function of low field mobility (μ_{Low}), saturation velocity (v_{sat}) and fitting parameter β. As suggested in [78] the β is the only parameter that can be tuned to match the f_T and f_{MAX} performance considering a given structure. The default value of β are 1.384 and our chosen value is 2.8 [6].

Recombination and Lifetime Models

Shockley-Read-Hall and Auger recombination models with their default parameters are used to take care of the recombination effects in different regions of the device.

$$\tau_{dop}(N_{A,0} + N_{D,0}) = \tau_{\min} + \frac{\tau_{\max} - \tau_{\min}}{1 + \left(\frac{N_{A,0}+N_{D,0}}{N_{ref}}\right)^{\gamma}} \tag{1.10}$$

where the symbols have their usual meaning.

Impact Ionization

Lackner's model is used for accounting for the breakdown voltage. As we know, when there is a reverse bias at the junction, breakdown may occur, as mentioned in [79] Lackner formulated the model modifying the already existing Chynoweth law. The temperature-dependent factor γ introduced to:

$$\alpha_v(F_{ava}) = \frac{\gamma a_v}{z} \exp\left(-\frac{\gamma b_v}{F_{ava}}\right) \tag{1.11}$$

where $v = n, p$ with

$$z = 1 + \frac{\gamma b_n}{F_{ava}} \exp\left(-\frac{\gamma b_n}{F_{ava}}\right) + \frac{\gamma b_p}{F_{ava}} \exp\left(-\frac{\gamma b_p}{F_{ava}}\right) \tag{1.12}$$

and

$$\gamma = \frac{\tanh\left(\frac{\hbar\omega_{op}}{2kT_0}\right)}{\tanh\left(\frac{\hbar\omega_{op}}{2kT}\right)} \tag{1.13}$$

The default parameter are tabulated and are applicable withing $10^5\ Vcm^{-1}$ to $10^6\ Vcm^{-1}$ electric filed range. Lackner's model with default parameter set has been considered to include the effect of breakdown at the high electric field.

Table 1.1: The default parameters and their corresponding modified values used during B55 device calibration are tabulated in the following table.

Model name	Parameter	Default value	Our value
HD	r, f^{td}, f^{hf}	1, 1, 1	1, 1, 0.2
Slotboom	N_{ref}	1.30×10^{17}	1.5×10^{18}
Caughey-Thomas	β	1.384	2.8
Kappa	$NiSi$	–	2
Poly-silicon	τ_n, τ_p	3×10^{-8}, 1×10^{-8}	3×10^{-8}, 3×10^{-8}

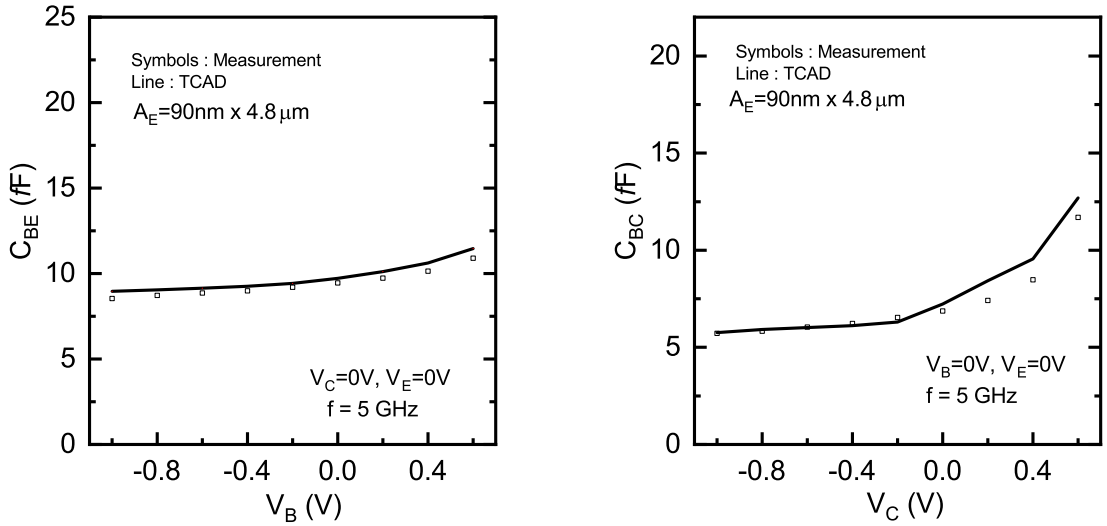
1.4.2 Calibrated Device Characteristics and Their Matching With Corresponding Measurement Data

During the measurement, the junction capacitance results are obtained from *cold* S-parameter measurement. *cold* measurement is the bias condition at which transistor operates at a reverse bias to a very low forward bias, at which the carrier injection across the junctions is negligible [80]. The base-emitter and base-collector junction capacitances are shown in Figs. 1.19 (a) and (b) respectively. The capacitances and their extraction from TCAD Y-parameters are expressed below.

Base emitter capacitance (C_{BE})

$$C_{BE} = \frac{img\{Y_{11}\} + img\{Y_{12}\}}{2\pi f}. \tag{1.14}$$

The base collector capacitance (C_{BC})



(a) (b)

Figure 1.19: Base/collector-emitter voltage dependent (a) base-emitter capacitance (C_{BE}) and (b) base-collector capacitance (C_{BC}): comparison between actual measurement and TCAD simulation.

$$C_{BC} = -\frac{img\{Y_{12}\}}{2\pi f}. \tag{1.15}$$

The capacitances are obtained from the simulation at low-frequency (5 GHz) and are matched with the corresponding measured data to validate the approximation used to emulate the doping profile in TCAD [37].

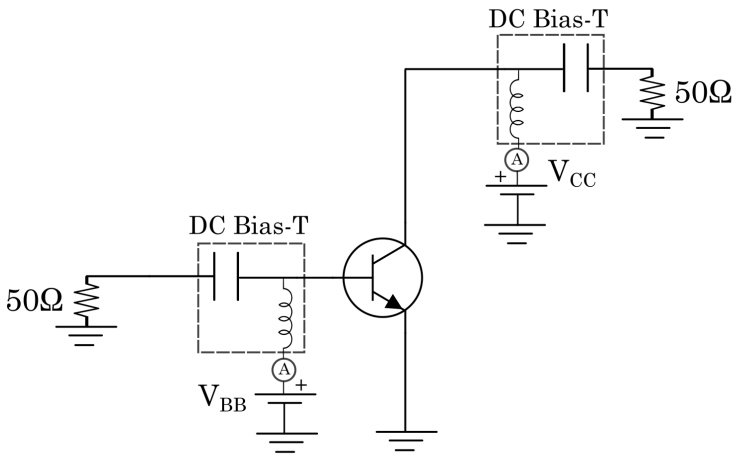
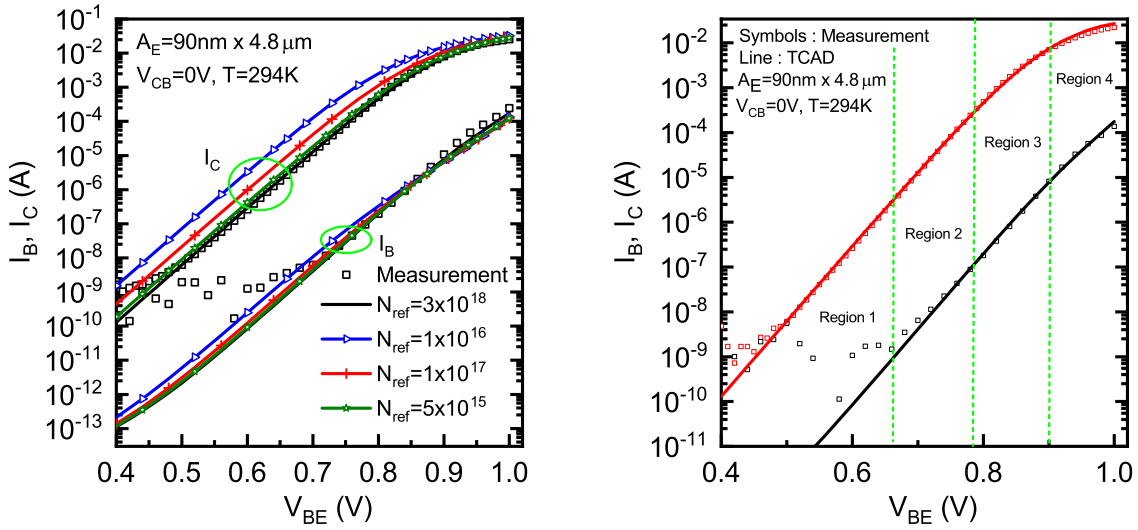


Figure 1.20: A typical measurement set-up used to characterize Gummel curves.

Concerning the DC characterization, a typical test setup is shown in Fig.1.20, it is important to note that if a high-quality RF termination is not ensured, then oscillation can occur at high bias levels; thus, the DC bias is fed through DC bias tees, while the RF ports terminated with 50 Ω load impedance. The input voltage is swept in a linear fashion, resulting in quasi-logarithm

currents. Once the results are collected, they are plotted against base voltage (V_B) on a semi-logarithm scale.



(a) (b)

Figure 1.21: Measured and TCAD simulated Gummel characteristics for 90 $nm \times 4.8\ \mu m$ SiGe HBT at V_{CB}=0V: (a) The impact of N_{ref} parameter is demonstrated considering different intermediate values from default to used value (b) Gummel characteristics depicting different regions of the characteristics.

The impact of N_{ref} parameter from Slotboom model as discussed in last subsection is demonstrated in Fig.1.21(a).

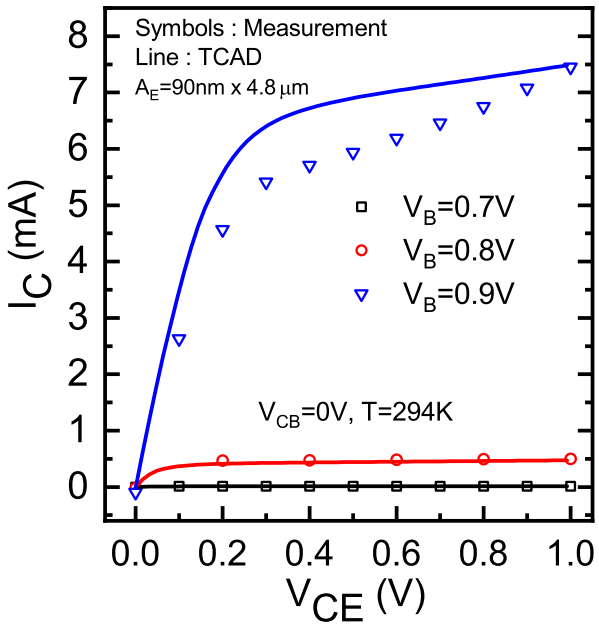


Figure 1.22: Measured and TCAD simulated output characteristics for 90 $nm \times 4.8\ \mu m$ SiGe HBT at different V_B.

The Gummel currents shown Fig.1.21(b) can be divided into different sections based on physical factors affecting it.

- In region 1, the base current (I_B) is mainly attributed to the recombination and the collector current (I_C) is mainly due to the ideal transport current. The dominant component of the base current density (J_B) can be written $J_{B,RC} = J_{B0,RC} \exp\{V_{BE}/\eta_{b,rc}V_T\}$, where $J_{B0,RC}$ is the base recombination saturation current density and $\eta_{b,rc}$ is the low injection base recombination current ideality factor. The default $\eta_{b,rc}$ is always $\geq$, with 1 is the value in ideal case.
- In region 2, the collector current is still ideal but the ideal component of the base current begins to dominate. In this region, the current density (J_B) can be ideally described by two ideal current components are $J_B = J_{B1} + J_{B,RC} = J_{B0} \exp\{V_{BE}/\eta_b V_T\} + J_{B0,RC} \exp\{V_{BE}/\eta_{b,rc}V_T\}$, where as mathematical equation for collector current density (J_C) remains same as in region 1.
- In region 3, the log of both base and collector current are no longer linear w.r.t base voltage due to parasitic base and emitter resistances, which results in a reduction of voltage (V_{BE}) for a given value of base voltage (V_B).
- In region 4, the high injection effects described in chapter 1 come into the picture resulting in the deviation of the terminal currents, and resistance from their ideal values.

From the detailed Gummel characteristics discussed here, it will be easy to look into concerned model parameters and calibrate the device. The recombination parameter (refer Tab:1.1) in the poly-silicon emitter has been tuned to match the base current. An ideal germanium profile as shown in Fig.1.18 is considered keeping the maximum mole fraction within the permissible range. This effectively helped us in capturing the collector current in the Gummel plot. Similarly, the output characteristics at different base voltages is shown in Fig.1.22.

Once the Gummel plot has been adjusted, the next step is to calibrate transit frequency (f_T) to evaluate the accuracy of the transit time and high injection effects. The frequency at which the magnitude of the small-signal short circuit current gain becomes unity is referred to as transit frequency (f_T). The analytic expression derived from the small-signal equivalent model is expressed in equation 1.16.

$$f_T = \frac{1}{2\pi \left(\tau_f + \frac{C_{BE}+C_{CE}}{g_m}\right)} \tag{1.16}$$

where τ_f is the forward transit time of minority carriers and g_m is the transconductance of the device, C_{BE} and C_{CE} are the depletion capacitances at the corresponding junctions. The total forward transit time can be separated into different components based on regional analysis [81].

Determination of f_T from measurement is not so straight forward as expressed in last equation, it is extracted from the h_{21} parameter as below equation 1.17.

$$f_T = \left| f_{extrap}.h_{21}(f_{extrap}) \right| \tag{1.17}$$

where f_{extrap} is the extrapolation/spot frequency and h_{21} is small signal current gain in common-emitter configuration approximated to equation 1.18 .

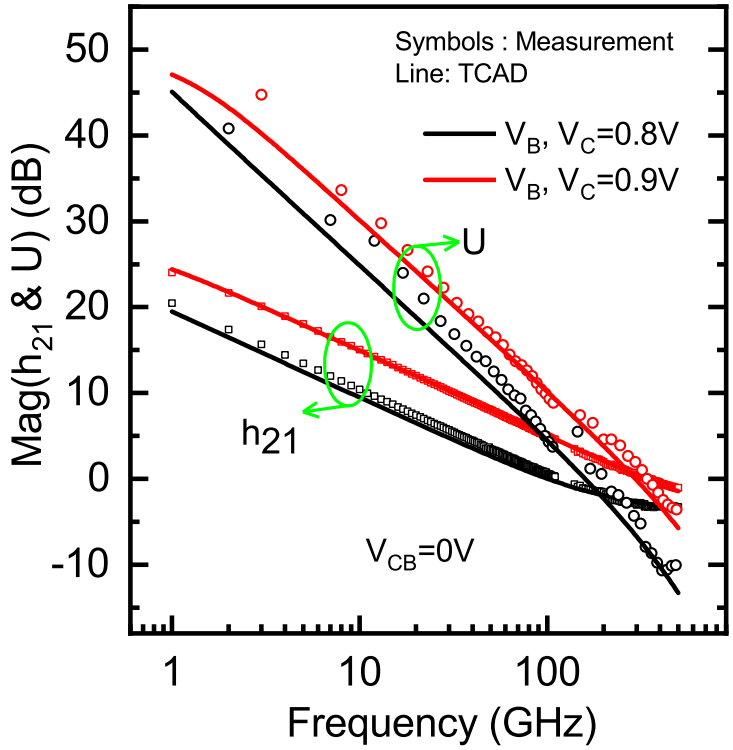


Figure 1.23: Small signal current gain (h_{21}) & Manson gain (U) vs. frequency at V_{CB}=0V: comparison between actual measurement and TCAD simulation.

$$h_{21} = \frac{g_m}{j\omega(C_{BE} + C_{BC})}. \tag{1.18}$$

The frequency at which the power gain is unity is considered as the maximum oscillation frequency (f_{MAX}). The expression to calculate it is as in the equation 1.19.

$$f_{MAX} = \sqrt{\frac{f_T}{8\pi C_{BC} R_B}}. \tag{1.19}$$

Here we can see f_{MAX} benefits from increase in transit frequency, base resistance(R_B) and base collector junction capacitance (C_{BC}). Similarly like f_T, f_{MAX} can be extracted from different power gain formulations. Here we used Mason's unilateral power gain (U).

$$U = \frac{|Y_{21} - Y_{12}|^2}{4(\mathrm{Re}[Y_{11}]\mathrm{Re}[Y_{22}] - \mathrm{Re}[Y_{12}]\mathrm{Re}[Y_{21}])}. \tag{1.20}$$

The small-signal current gain (h_{21}) and Mason gain (U) are shown in Fig. 1.23 at two different bias in the main operating range, which confirms a good level of calibration to proceeds for f_T and f_{MAX} extraction.

It is to be noted here that, in order to calibrate small-signal current gain h_{21} or the transit frequency f_T, the vertical doping profile of the device must be carefully matched. The base width

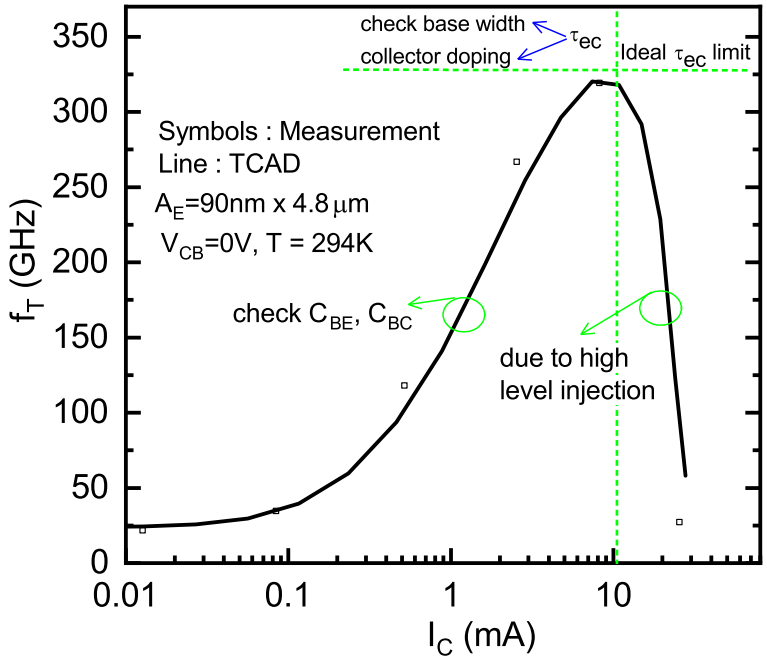


Figure 1.24: Collector current dependent (a) transit frequency (f_T) at V_{CB}=0 V: comparison between actual measurement and TCAD simulation.

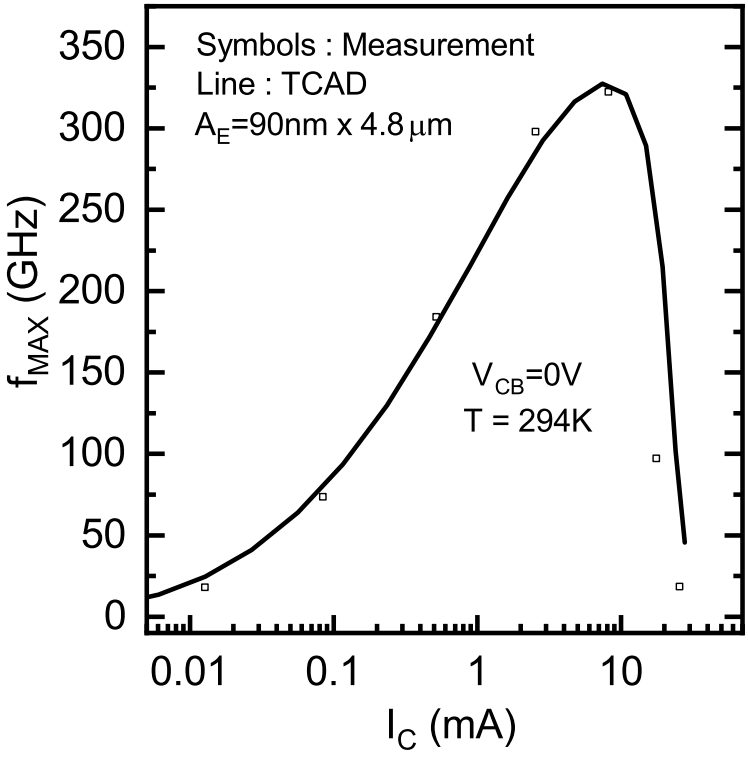


Figure 1.25: Collector current dependent maximum oscillation frequency (f_{MAX}) at V_{CB}=0 V: comparison between actual measurement and TCAD simulation.

and emitter junction depth have to be optimized to control the overall emitter to collector transit time (τ_{ec}). The matching in the junction capacitances C_{BE} and C_{BC} helps in calibrating the steeper rising slope of the transit frequency curve as shown in Fig.1.24. It is also to be noted a device having a good f_T at lower bias helps in low power consumption applications. In order to match the high current region of the f_T vs. I_C characteristics, the collector doping needs to be carefully matched. Increasing the collector doping can decrease the collector junction depletion width and collector resistance. This counteracts the high-level injection and forthcoming Kirk effect. In fact, this increases the C_{BC}, hence, to have both low and high current f_T matching needs careful calibration of doping in the collector.

To summarize, the small-signal current gain (h_{21}) and Manson gain (U) are shown in Fig.1.23 further, the TCAD simulated transit (f_T) and maximum oscillation frequencies (f_{MAX}) are compared with the measured data at V_{CB}=0 V and T=294 k in Figs.1.24 and Figs.1.25 , respectively, prior to carry out the S-parameter comparison. It is to be noted that for the f_T and f_{MAX} extraction, a spot frequency at 63 GHz has been chosen [82]. Since the Gummel (Fig.1.21) and f_T characteristics (Fig.1.24 (a)) show a high level of correlation between the TCAD and measured data and since these two characteristics are highly dependent on the intrinsic part of the HBT, one can conclude that a reasonably accurate calibration of the 1D profile of the device has been obtained in TCAD [83]. Since f_{MAX} is dependent on f_T and other external parameters such as base resistance and base-collector and collector-substrate capacitances, the high level of agreement between the TCAD and measured f_{MAX} characteristics ensures that the profile parameters related to the extrinsic base, buried layer in the collector region and the substrate doping at its contact and around the deep trenches are closely identical with those of the B55 process. The same is also confirmed from S-parameter shown in Figs.1.26, 1.27 and frequency dependent f_T, f_{MAX} in Fig.1.28 in the main operating regime (V_{BE}=0.75 V to 0.9 V). This is also validated after looking into the doping profiles received from ST Microelectronics (not disclosed here due to intellectual properties agreements).

1.5 Virtual Measurements, Results and Discussion

As discussed in section 2.5, after observing a mismatch between the actual measurement data for S-parameters with calibrated TCAD simulation, we decided to include the non-ideal environment factor related to BEOL, pad-parasitics, probes, etc. into the TCAD simulation results via additional EM simulation. For this purpose, we performed three-dimensional (3D) EM simulations using the commercially available simulator, *Ansys*-HFSS. The simulation layout is imported on a silicon substrate, and the probes are placed the way they appear during the measurement [64]. Then, the whole structure along with the probes, is covered within an air box as shown in Fig.1.29. During the setup, proper care has been taken in choosing the meshing and assigning the appropriate boundary conditions for absorbing the radiation at the faces of the air-box in order to emulate an infinite free space environment. Four separate probe models for each frequency band, as shown in Fig. 1.7 are chosen, and finally, simulations are performed by giving input excitation at each wave-port. Together, the TCAD simulation of SiGe HBT up to Metal-1 and the EM simulation to account for the remaining non-ideal environmental factors, as described above and shown in Fig.1.29, Fig.1.30 and Fig.1.31 generate the virtual measurement data for our further investigations.

Virtual measurement data are achieved following the flow graph described in Fig.1.31. First, the RF probes, the pads, and the BEOL down to metal-1 where the transistor is connected are simulated with HFSS (4-port EM simulation). Next, the TCAD (2-port TCAD) and EM simulation results are concatenated. This data set contains both the access (probes, pads, BEOL) as

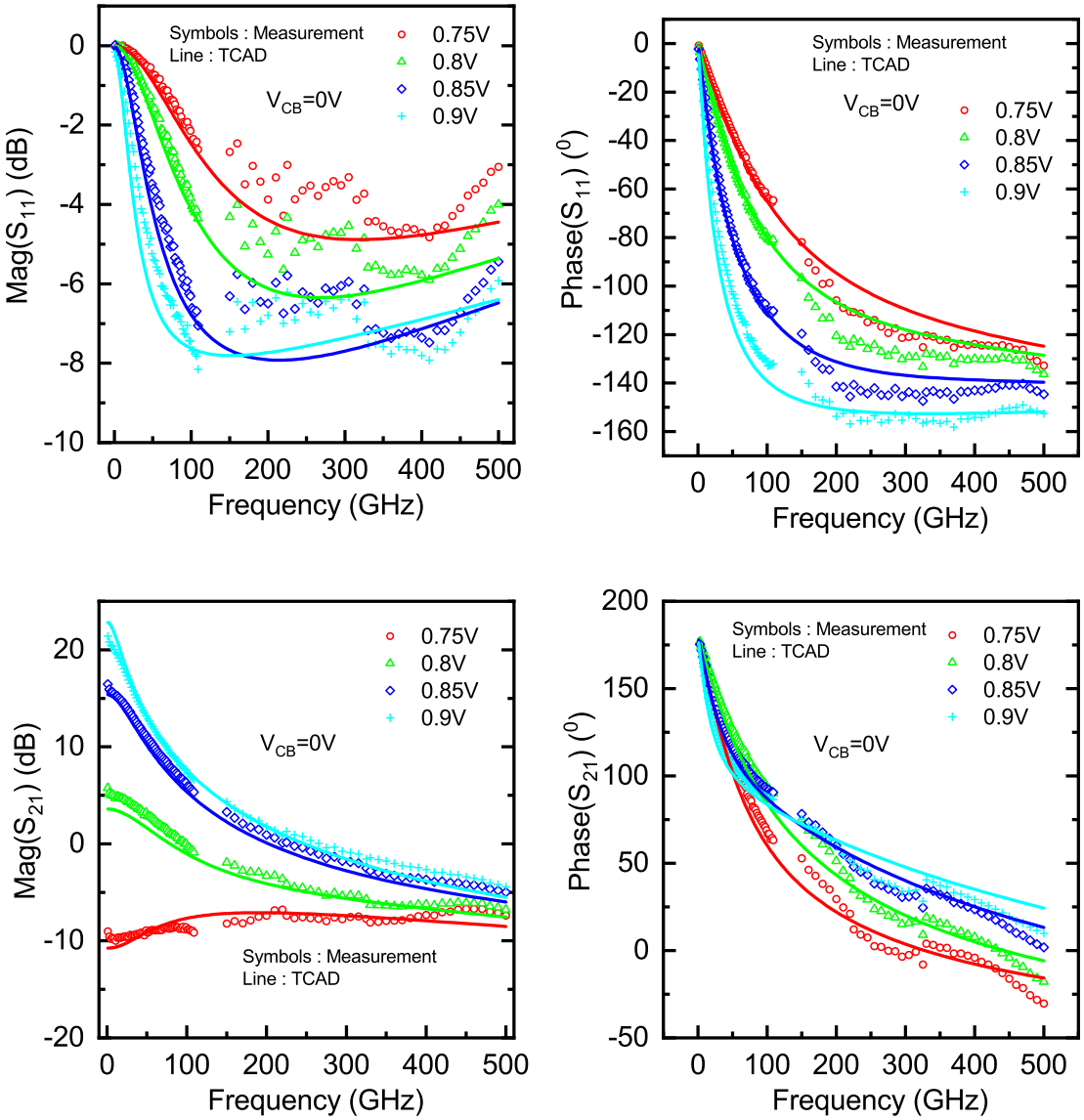


Figure 1.26: Frequency dependent magnitude and phase of S_{11} & S_{12} -parameters at V_{CB}=0 V and V_{BE}= 0.8 V and 0.9 V for the SiGe HBT up to metal-1: measurement and TCAD comparison.

well as the actual transistor itself. The next step is to perform a calibration and de-embedding as done in real measurement. For that purpose, all TRL calibration standards are simulated with HFSS, and the calibration is performed as shown in Fig.1.31 [69]. Finally, the Short-Open de-embedding structures are simulated as well, and the de-embedding is completed leading to the virtual measurement data. It is to be noted that the detailed EM-simulation are performed by [64], [69] and [71] and are used in this work, so more detail about that can be found in corresponding articles. If calibration and de-embedding had been ideal, the de-embedded virtual measurement data would perfectly match with the TCAD simulation of the actual transistor. The results for the three data sets (actual measurements, TCAD simulation and virtual measurements) are compared in Figs.1.32, 1.33 and 1.36 for the magnitude and phase of S_{11}, S_{12}, S_{21}, S_{22} and f_T, f_{MAX} up to 500 GHz.

From the comparison plots following observations are made.

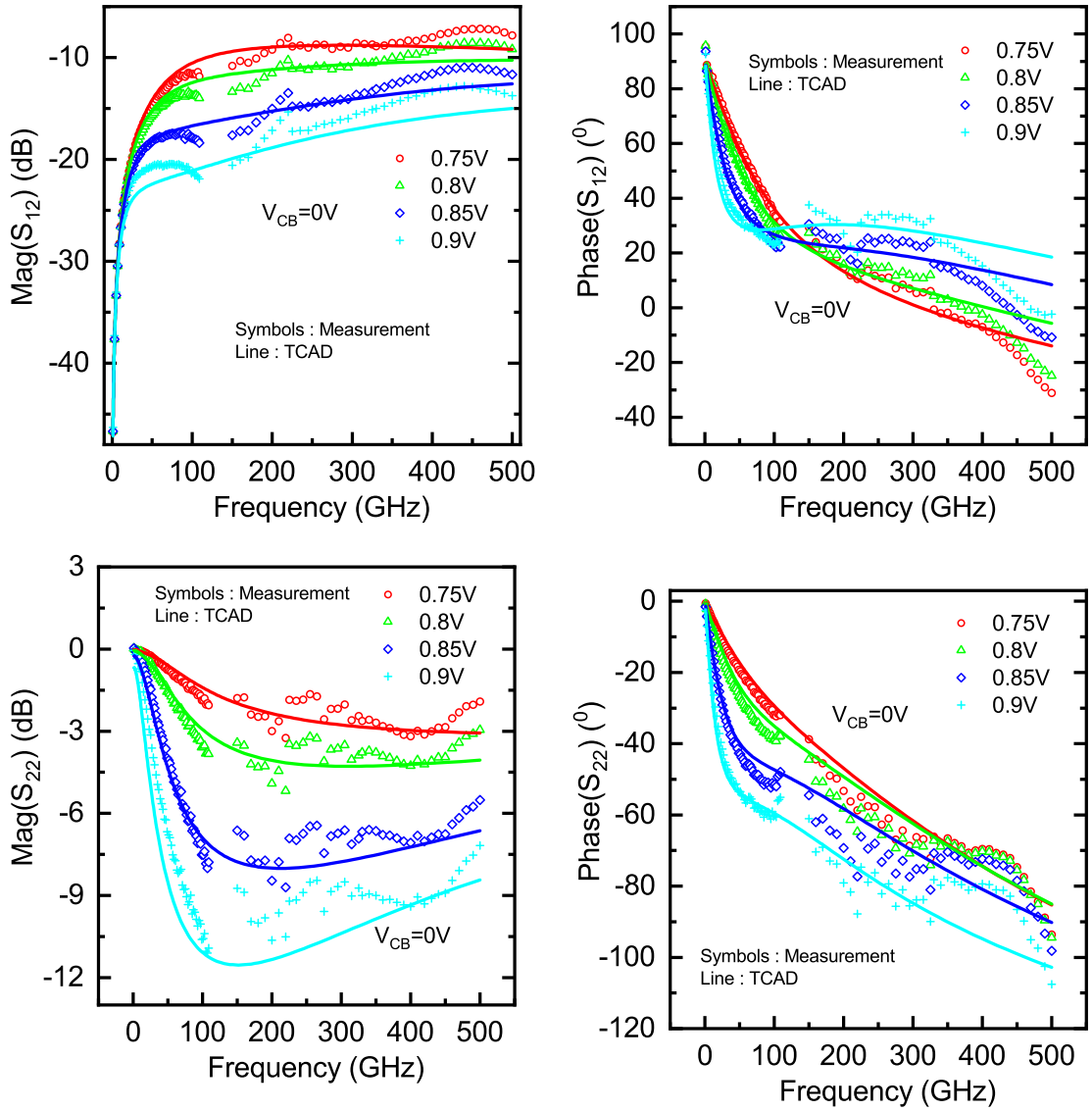
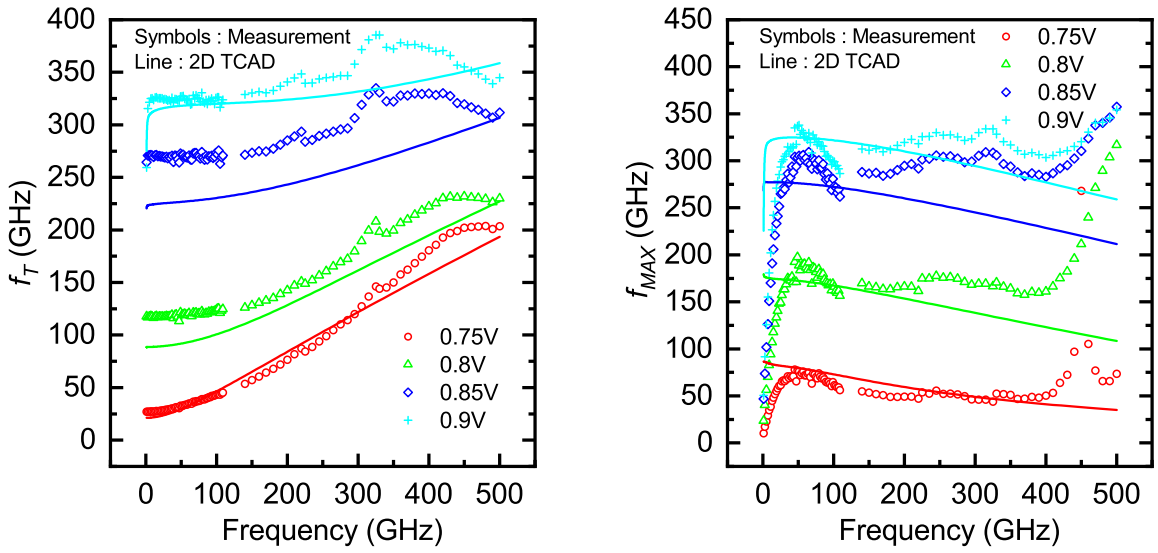


Figure 1.27: Frequency dependent magnitude and phase of S_{12} & S_{22} -parameters at V_{CB}=0 V and V_{BE}= 0.8 V and 0.9 V for the SiGe HBT up to metal-1: measurement and TCAD comparison.

- Up to 100 GHz, a good correlation can be observed for both V_{BE} bias points and for all four S-parameters (magnitude as well as phase).
- In the frequency range of 50-100 GHz, we can observe there is a sudden reduction of f_{MAX}; it can be observed from S_{12} parameter, which is strongly related to f_{MAX} and can be attributed to cross-talk. A similar dip is observed in f_T also in 50-100 GHz range can be due to minor difference in probe angle and position in both real and simulation environment. It can also be observed in the magnitude of S_{21}.
- When changing the frequency bands (especially at the transition into the 140 GHz to 220 GHz band), discontinuities are observed with the actual measurements while comparing with TCAD and virtual measurements. Therefore, it can be concluded that the TRL calibration along with short-Open de-embedding are not able to correct the coupling with HF probes and the wafer surface accurately for the given test-structure and the HF



(a) (b)

Figure 1.28: Frequency-dependent transit frequency (f_T) and maximum oscillation frequency (f_{MAX}) at V_{CB}=0 V and V_{BE}=0.8 V and 0.9 V for the SiGe HBT up to metal-1: measurement and TCAD comparison.

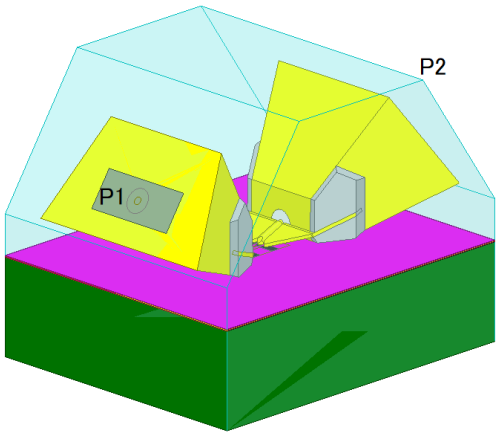
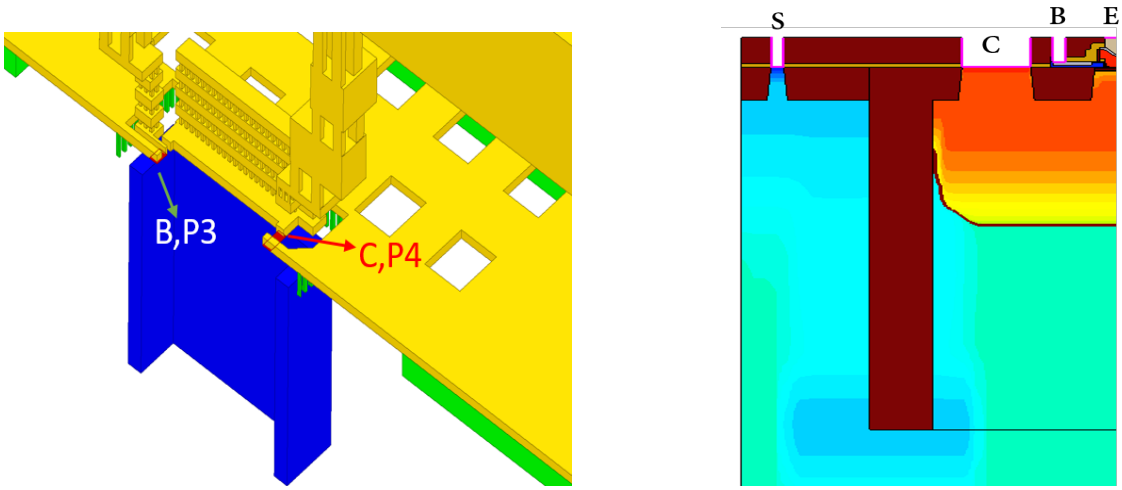


Figure 1.29: Side view of the EM simulation setup for probes enclosed within an air-box. P1 and P2 probes are marked which are a part of 4-port EM simulation.

probes used. Also, the measurements are noisy and can be attributed to some difficulties to have a reliable contact between the probes and pads, and in particular, considering the small-signal pad surface 40 $\times$25 μm^2 and the thickness of the passivation layer. In fact, the very short tips of the WR3.4 probes and their angle do not allow us to have a reliable contact on our BICMOS test structures (please note that this is not the case on calibration test structures).

- The discontinuous transition of actual measurements at 220 GHz is well predicted by the virtual measurement, especially for the magnitude of S_{12}-parameter. This discontinuity has already been studied in [64]. WR5.1 (140-220 GHz) uses a technology which is closed to the probe designed for 110 GHz band, while the WR3.4 (220-325 GHz) uses a very different technology which is much more scaled. We have demonstrated in [64] that the



(a) (b)

Figure 1.30: Half structure from EM simulation (left) depicting the other two ports P3 as Base and P4 as collector, below which the active device starts. S, C, B and E in the TCAD structure (right) corresponds to Substrate, Collector, Base and Emitter contacts respectively.

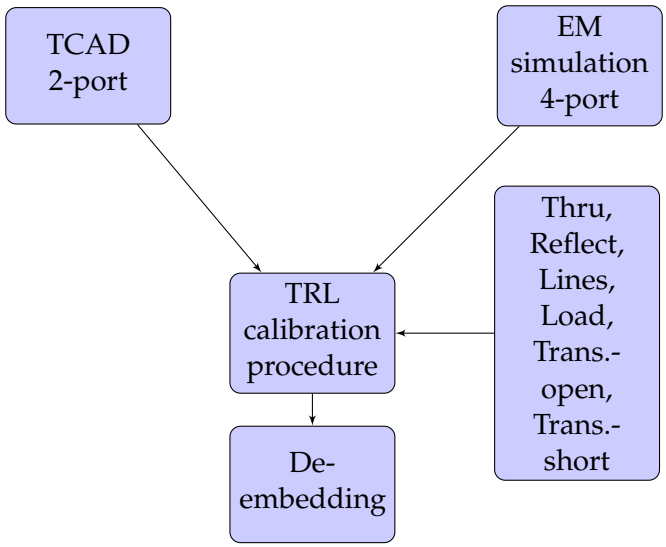


Figure 1.31: TCAD-EM co-simulation flow leading to a virtual measurement.

coupling with the environment is much stronger when the EM field is not confined by a highly scaled probe.

- Because of different resistance of bias tees from the probe to probe, there can be a change in bias current. The bump in the actual measurement data of f_T characteristics (Fig. 1.34) in the 320-420 GHz range can be a due of this resistance difference.
- While moving to higher frequencies beyond 400 GHz, a deviation of the expected behavior can be observed both in virtual as well as actual measurements compared to TCAD data. Even though no perfect agreement for the actual or virtual measurement data is achieved with the reference TCAD data, at least the trend is reasonably predicted. We suppose that the cross-talk is the source of these discrepancies since the magnitude of S_{12} is increasing more than what is predicted by the TCAD. Please note the TRL calibration

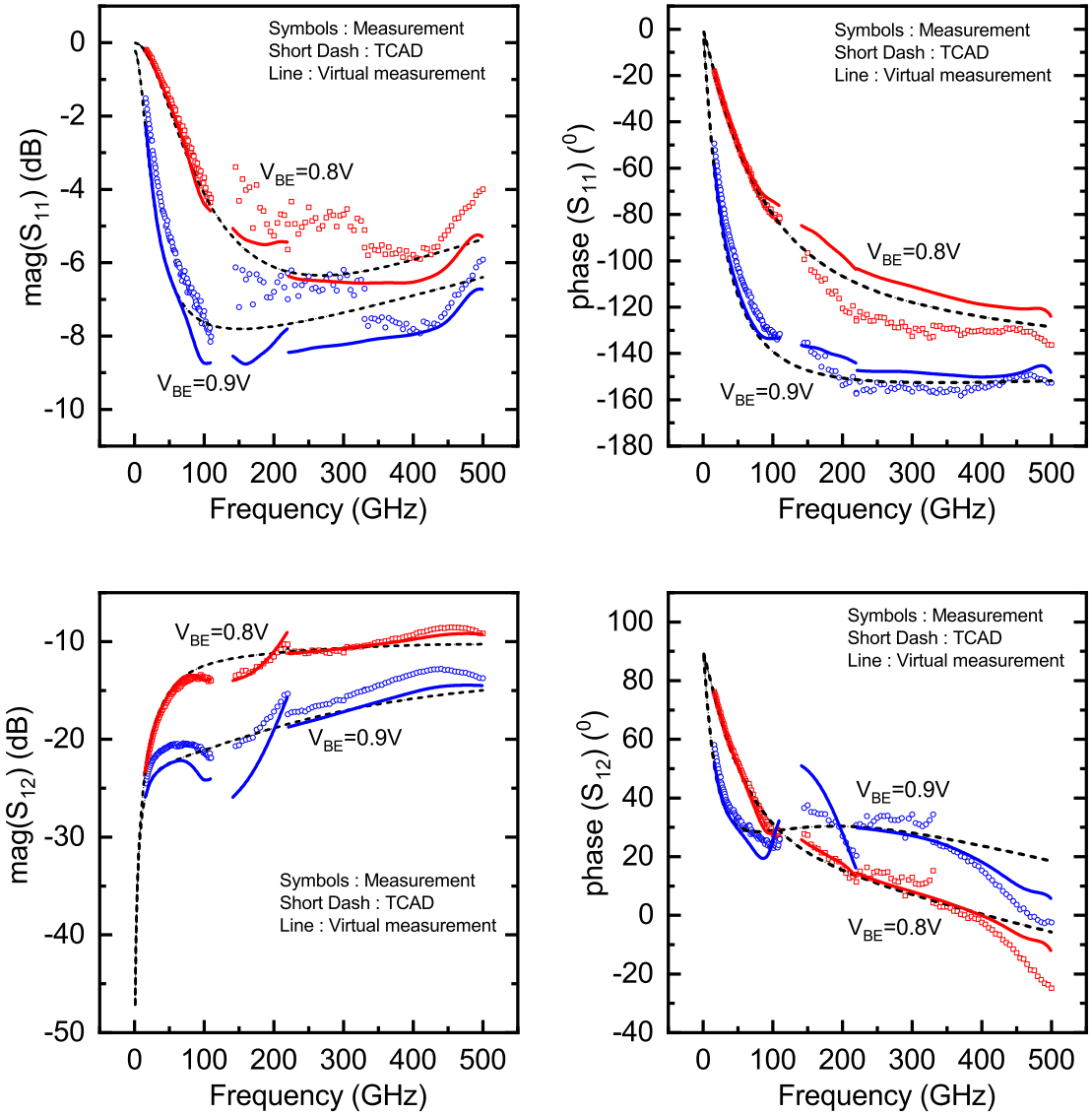


Figure 1.32: Frequency dependent magnitude and phase of S_{11} & S_{12} -parameters at V_{CB}=0 V and V_{BE}= 0.8 V and 0.9 V for the SiGe HBT up to metal-1: comparison of de-embedded actual and virtual measurements with TCAD simulation.

does not correct the cross-talk. The distance from pads to pads must be increased to reduce this effect. If with the picoprobe 110 GHz with 100 μm pitch, the cross-talk start to influence the results above 70 GHz, using a highly scaled probe such as the picoprobe WR2.2, the cross-talk start only above 400 GHz.

This brings us to the conclusion that the calibration and de-embedding procedures are still needed to be improved, especially at the higher frequency bands, and require further improvements in order to correct all the demonstrated errors introduced by the measurement environment. Two roads of improvement for measurement accuracy needs to be explored in parallel, one involving the design of the test structures the other one is about the design of the probes:

- On one side, the design of the test structures already uses a boundless ground plane, a large space between each structure (200 μm in X- ,and 133 μm in Y-direction) may still be

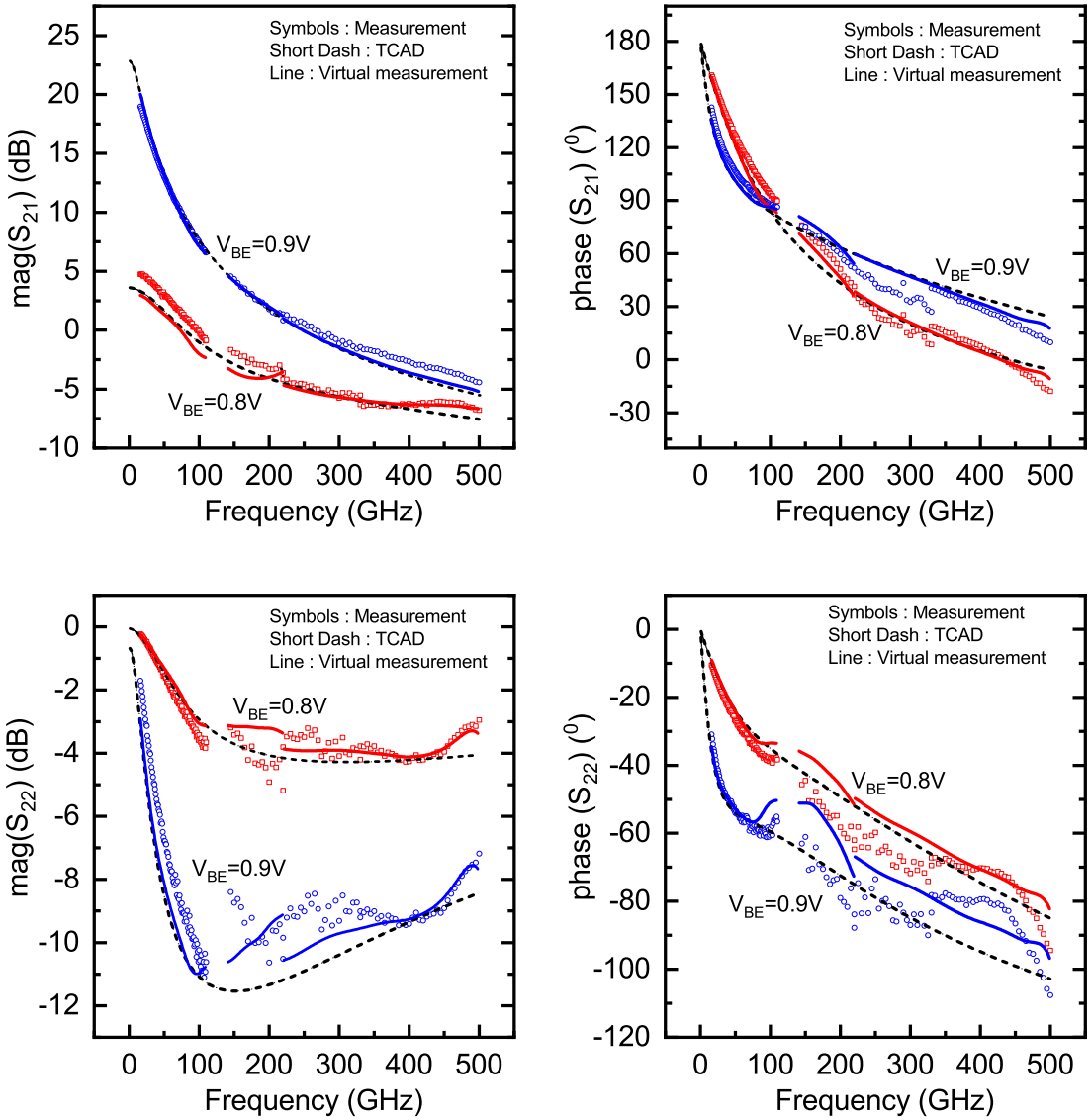
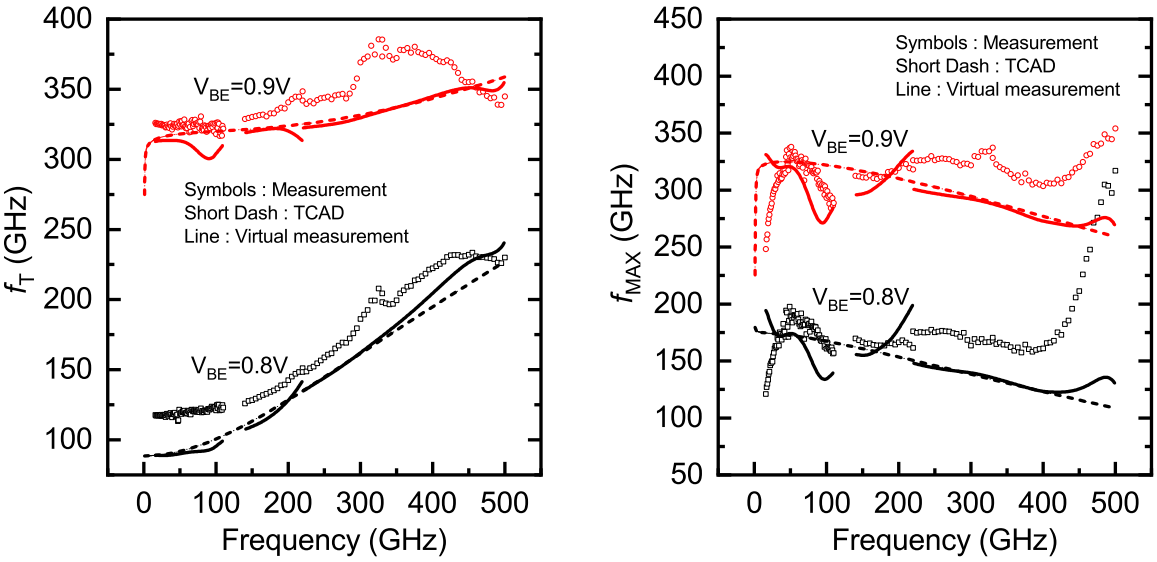


Figure 1.33: Frequency dependent magnitude and phase of S_{21} & S_{22} -parameters at V_{CB}=0 V and V_{BE}= 0.8 V and 0.9 V for the SiGe HBT up to metal-1: comparison of de-embedded actual and virtual measurements with TCAD simulation.

improved on the two following points: the distance from pad to pad to reduce the cross-talk and the design of the pad itself which is mainly empirical. Please note that most of the improvement leads to very costly test structures.

- On the other side, The WR3.4 and WR2.2 probes are highly scaled probes that give very reasonable results [64] but require much higher manufacturing cost. If one can develop the design of a similar probe keeping the same geometry scaling, but with the connector for the lower frequency band (especially 140-220 GHz), the measurement accuracy would be much improved.

This brings us to the conclusion that the calibration and de-embedding procedures are not absolutely flawless, especially at the higher frequency bands, and require further improvements in order to correct all the demonstrated errors introduced by the measurement environment.



(a) (b)

Figure 1.34: Frequency-dependent (a) transit frequency (f_T), and (b) maximum oscillation frequency (f_{MAX}) at V_{CB}=0 V and V_{BE}=0.8 V and 0.9 V for the SiGe HBT up to metal-1: comparison of de-embedded actual and virtual measurements with TCAD simulation.

The impact of these unavoidable errors on the circuit level will be explored in the next section.

1.6 G-band Amplifier Design

The preceding analysis makes it clear that there may be a mismatch between the characteristics of the intrinsic device (up to metal-1) and the actual measurement (from which the influences of the probes and pads in the calibration process are not perfectly eliminated). However, parameter extraction is performed on actual measurement data, and the process design kit contains an electrical compact model parameter set for the transistor based on these imperfectly de-embedded measurement data.

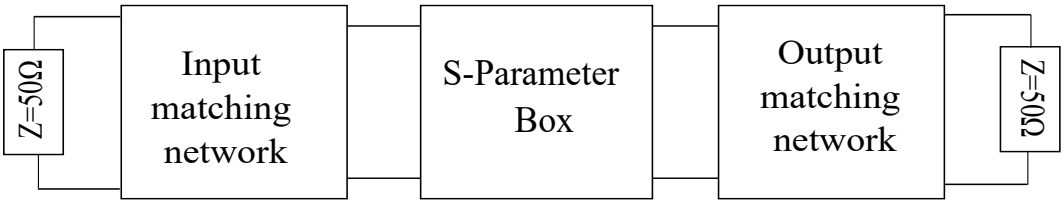
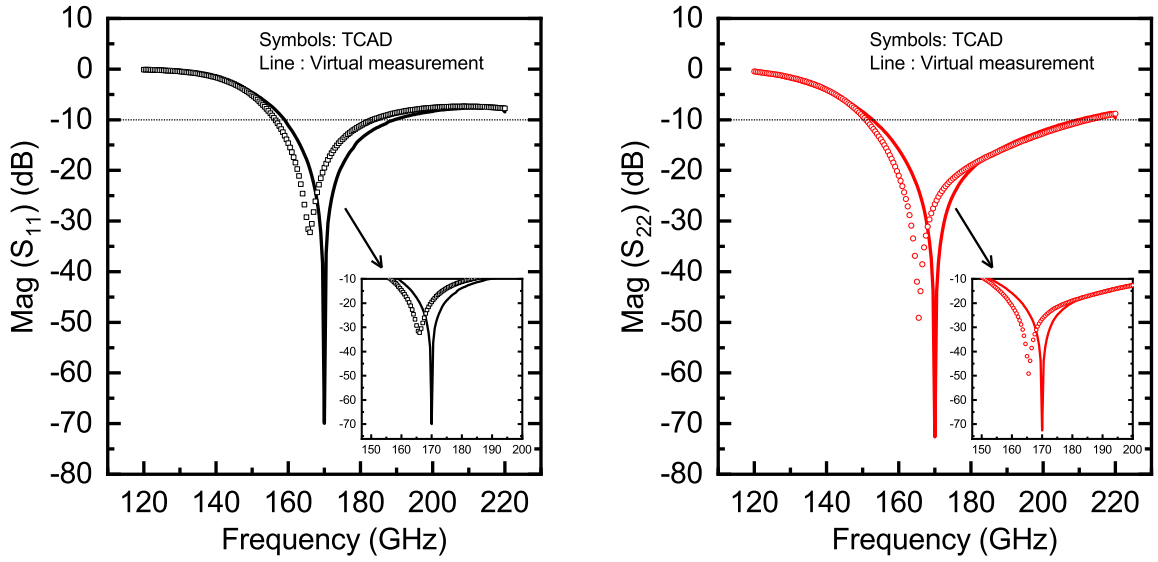


Figure 1.35: Block diagram representation of an single stage amplifier circuit with input and output matching networks.

In the following, we evaluate the effects of using a slightly erroneous parameter set for circuit design through a case study on a single-stage amplifier operating at 170 GHz. The same design has been re-simulated with a perfect parameter set representing only the actual device up-to metal-1 that is obtained from TCAD simulation.

The S-parameter box in Fig.1.35 contains the virtual measurement data simulated at V_B=0.85 V and V_{CB}=0.5 V, which stands for the optimum gain bias conditions, below the breakdown voltage. Before going ahead with the design of the matching network, K-$\triangle$ stability is checked for both virtual measurement and TCAD data. The required conditions of $K > 1$ and $\triangle < 1$ are met for both data sets at 170 GHz.



(a) (b)

Figure 1.36: Comparison of frequency-dependent (a) S_{11}, (b) S_{22} obtained from two amplifier circuits designed at V_{BE}=0.85 V and V_{CB}=0.5 V: one amplifier considers a transistor virtual measurement data while the other corresponding to the TCAD simulation.

Using the virtual measurement data for the transistor, input and output matching networks for the G-band single-stage amplifier are designed for the port matching with 50 Ω termination at 170 GHz (see Fig.1.35). Now the virtual measurement data is replaced with the TCAD data, keeping the matching networks and bias point fixed. The amplifier circuit is simulated using both TCAD and actual measurement data for the transistor, and the obtained results are evaluated subsequently. First, we check that the matching constraint S_{11}, S_{22} < -10dB is met for the amplifier circuit using both data sets for the transistor (as shown in Fig.1.36). Considering a -10 dB bandwidth, a 30 GHz bandwith is observed (160-190 GHz). In this 160-190 GHz range, the gain S_{21} in Fig.1.37) shows a difference of around 1 dB between TCAD simulation and virtual measurement, which is significant at such a high frequency. In addition, the maximum gain value is shifted by around 10 GHz towards a higher frequency for the amplifier's virtual measurement.

It is apparent from the results that if someone designs an amplifier based on model simulation in millimeter-wave range, the measured circuit performance may not meet the predicted performance. Particularly, in the case of a narrow band application, such a performance deviation of the measured circuit may potentially lead to unsatisfactory specification requirements.

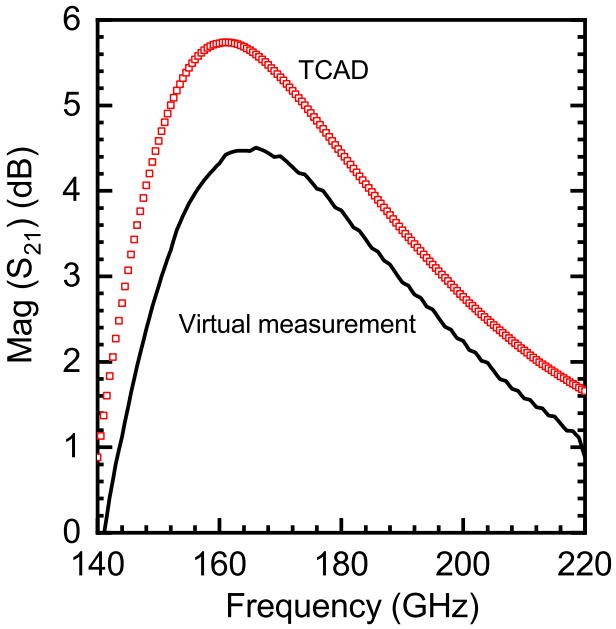


Figure 1.37: Comparison of frequency-dependent gains (S_{21}) obtained from TCAD simulation (red line) and virtual measurement (black line) of the G-band single-stage amplifier

1.7 Conclusion and Discussion

For the first time, we analyzed a very high frequency (>110 GHz) measurement data using two finite element tools, one solving the EM equations and the other one solving the semiconductor equations. The presented methodology provides one with sufficient confidence in the adopted characterization techniques and results. More precisely, it allows one to differentiate between accurate and erroneous characterizations. This work is a step forward (i) to adopt an improved characterization technique and (ii) to validate the TCAD result in high frequency; this will help the modeling of non-quasi static and other high frequency effects, which has been already demonstrated in [84] work. As a whole, the analysis outlines a prescription in order to make the researchers aware of certain precautions while designing the test structures, carrying out the process of characterization, and verifying the corresponding results. The detailed methodology, if adopted, will certainly help in the accurate extraction of high-frequency compact model parameters.

Chapter 2

Novel SiGe based Nanowire Bipolar HBT for THz Frequency Applications

2.1 Introduction

Since the evaluation of SiGe BiCMOS technology, it has been serving the continuous demand for higher functionality and front-end performance quite well at low cost and medium to low volumes. As the operating frequency of SiGe HBTs exceeds 300 GHz (e.g., [36],[85],[86]), it allows critical circuits to operate beyond 100 GHz, which is called the lower end of THz gap. The upper limit of the THz gap extends up to 30 THz. Within this THz gap range a lot of applications are envisioned like (i) *THz imaging and sensing* comprising of security, medical and biochemistry, material reliability and meteorology [87], [88] (ii) in *measurement equipment* like in ultra high bandwidth analog to digital converters [89]. With the emerging mm-wave and THz market in sight, there is no fundamental device physics-related reason why HBT development cannot be persuaded aggressively as done for CMOS. However, predicting the analog HF application is not as straightforward as digital CMOS applications. The following section discusses some of the literature on SiGe HBT performance prediction.

2.2 Literature Survey

Various works have already been published to evaluate the performance limit of a bipolar device. Research by Hoeneisen et al. [90] correctly predicated avalanche breakdown, power dissipation, doping density fluctuation, and EM effects that ultimately limit the performance of a device. However, it was computed using DD transport theory and limited to Si-based BJTs in digital circuit applications. Limitations of this approach are that; it is based on the DD model and not predicated on a single device. Another effort by Gaur et al. [91] using 2D TCAD based simulation with DD transport model predicted $f_T/\ f_{MAX}$ of 17/10 GHz, for a device with a base width of 30 nm and emitter window of 400 nm. This performance comes with some physics-based limitations, including the computational one. Today's device performance has exceptionally outperformed the above-discussed performance limitations. After the 1980s, the focus was shifted mostly to MOS-based devices until SiGe HBT became widespread with hybrid BiCMOS technology. Among all the most cited one is Johnson's limit, which has been estimated to be 200 GHzV for the silicon-based transistors [22]. Nevertheless, this prediction was too early and based on 1D drift-diffusion (DD) transport. This limit has already been exceeded long back by fabricated Si and SiGe-based bipolar devices.

In the work by Shi et al. [92], using 2D-simulation, an HD transport model predicted f_T/f_{MAX} of 760/1090 GHz, with the emitter opening of 60 nm. However, it has been estimated with

some limitations like, (i) the discussion about the percentage of the complete device simulated is unclear, (ii) single-valued energy relaxation time was used, (iii) f_{MAX} estimated from the simple standard equation, and (iv) self-heating was not included during simulations. Schroter et al. [93] using vertical and then lateral scaling predicted a f_T/f_{MAX} of 1/1.5 THz. The vertical profile of the device presented in [85] is aggressively scaled and uses the HD model until reaching an optimized response. Then maintaining the same optimized profile, aggressive scaling of the lateral dimension is done and overall predicted such a high f_T/f_{MAX} of 1/1.5 THz. To conclude, we need to scale or remove as many parasitics as possible to improve the RF performance towards a maximum limit. With vertical scaling, f_T improvement, and with lateral scaling f_{MAX} improvement can be achieved.

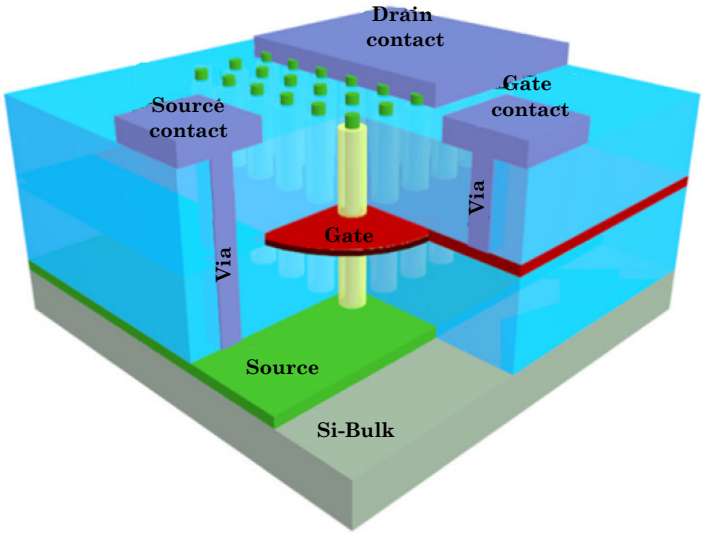


Figure 2.1: The nanowire device proposed by [94].

More realistic approaches for the performance prediction of STM's B55 device [33] are discussed in this chapter while preserving the same vertical profile and removing the probable lateral parasitics. Since this device is intended to be fabricated following the approach described in [94] (refer Fig.2.1), it will have a nanowire architecture. To the authors' knowledge, no research group has presented a SiGe HBT-based nanowire device to date. As a result, the literature on MOS devices based on nanowires for RF applications will closely match the performance analysis. As metal-oxide-semiconductor (MOS) devices' downscaling reached the International Technology Roadmap for Semiconductors (ITRS) limits, multi-gate device architectures emerged as an emerging technique due to improved gate electrostatic control over the channel. Amongst all advancements, gate-all-around architecture is one of the most pronounced technologies, and experiments on device fabrication and physics continue. So, in the MOS domain, the purpose of the gate-all-around gate nanowire (NW) is to offer improved gate electrostatic control over the channel [95]. Nanowires are classified to be solid rods with a diameter of less than 100-200 nm [95]. A semiconductor nanowire can compose multiple semiconducting materials. However, the problem is that a single nanowire device cannot carry enough current to be used as an efficient device due to its area constraints. So there are some proposals by different research groups to have an array of nanowires together contributing to high driving current and required performance [94].

As recently demonstrated by IBM, for the 2 nm technology node, the state-of-the-art MOS-based device fabrication is shifted towards the gate all around (GAA) nanosheet transistors

[96]. The stacking or array of more nanosheets gives the potential to derive more current than that possible from the FinFETs of the same size [97]. Interestingly, the nanowire structure provides a substantial surface-to-volume ratio that makes them unique for many applications discussed in [98].

The fabrication of nanowire devices can be categorized into two different methods : (i) the bottom-up approach and (ii) the top-down approach. In the case of the bottom-up approach, the nanowires are synthesized, mostly using the vapor-liquid-solid (VLS) chemistry [99], typically with the help of metal as a catalyst. Researchers have tried different approaches, including templated growth or using the etching method. In the etching method, the platinum masks were prepared with the help of the supper lattice nanowire pattern transfer process (SNAP) [100], [101]. The extensive discussion on bottom-up approach is discussed in [99], [102].

The nanowires are prepared using the lithography and etch process in the top-bottom approach, followed by trimming or stress-limited oxidation technique. Integration of top-down fabricated NWs in circuit functionality is straightforward [103], while the bottom-up approach is challenging while assembling the wires for the circuit approach.

In a loose tone, silicon-based nanowires are primarily used for logic applications, and III-V-based nanowires are intended to be used for high current driving applications, and RF applications for their respective advantage [104]. Some of the proposed NW devices for RF applications are discussed below. Egard et al. [105] proposed a vertical InAs device with gate length of 100 nm that shows a f_T>7 GHz and f_{MAX}> 20 GHz, where the frequency performance is limited by un-optimized geometry. It means that the gate capacitance is dominated by the parasitic capacitances between the gate and drain/source electrodes. Keeping the device technology unchanged Eric et.al [106] proposed another layout optimization technique and through simulation demonstrated an f_T, f_{MAX} > 500 GHz. Linjie et al. [107] proposed an SOI wafer-based silicon nanowire array with individual nanowire width of 25 nm, the gate of 130 nm, and TiN and Nickel Silicide for gate and drain/source contact, respectively. The key device RF performance of the MOS-based nanowire is f_T=14.8 GHz and f_{MAX}=1.75 GHz. Fredrik et al. [108] proposed a III-V device with Λ-ridge spacers demonstrated a f_T, f_{MAX} of 75 GHz and 100 GHz respectively for a device with 32 nm gate length. Kilpi et al. [109] proposed a III-V nanowire device on Silicon with a channel diameter of 30 nm and gate length of 120 nm. Extracted f_T, f_{MAX} of the device are 125 GHz and 130 GHz, respectively.

All these examples concern MOSFET transistors which use nanowires to improve the electrostatic effects in the transistor. Here, we propose to reuse these technological blocks to build a bipolar transistor. Indeed, the shape of the nanowire allows, in particular, in the case of the bipolar transistor, to strongly reduce the base resistance key point allowing to increase the f_{MAX}. Also, integrating III-V devices with silicon technology is an expensive and cumbersome task. So here we are proposing a SiGe-based bipolar nanowire device predicting f_{MAX} above 900 GHz. Silicon-Germanium (SiGe) devices are always preferred over III-V because they can be integrated smoothly with existing silicon technology, which is less expensive and gives the advantage of band engineering. So in this work, we proposed a SiGe-based nanowire HBT in the line of bipolar technology. As per the author's knowledge, such a device has not been proposed yet by any research group. In this proposed bipolar device, the purpose is different from the MOS. In MOS nanowire, where the target is to have improved electrostatic control over the channel, here the objective is to reduce the external parasitic by removing the extrinsic component of a conventional bipolar device to achieve improved RF performance.

As discussed in [93], [110], [111], the scaling of the lateral dimension will improve the maximum oscillation frequency, and the scaling of the vertical dimension will mainly improve the

transit frequency. In this proposed device, due to removal of the lateral extrinsic regions of the device in [49] improves the f_{MAX}. As it affects the extrinsic base region, the noise figure given by equations (2.1) can also be improved [110] which is beyond the scope of this work.

$$F_{opt} = 1 + \frac{1}{\sqrt{\beta_0}}\sqrt{1 + 2g_m r_b}, f_{MAX} = \sqrt{\frac{f_T}{8\pi r_b C_{BC}}}. \tag{2.1}$$

Here f_T is the transit frequency, C_{BC} is the base-collector capacitance, β_0 low-frequency ac current gain in common emitter configuration, g_m is the trans-conductance (I_C/V_T) and r_b is the series base resistance. The device is proposed to be fabricated by a top-down plasma etching approach presented in [94] with a doping profile already published in [33], [49]. But since it's a bipolar device, it needs a direct contact to be drawn from the base, unlike a gate oxide separating the channel from contact in the MOS device. Deposition of isotropic polysilicon layer to extend a base contact poses a severe fabrication limitation. Known deposition methods of polysilicon are only isotropic [discussion with [94]. It means that the whole emitter will be covered with the polysilicon which is undesired. Hence, another proposal is to use a metal deposition which is anisotropic and would allow to deposit metal only at the right place. Hence, a metal contact needs to be added directly to the base, which brings some limitation in terms of the base current degradation.

Here, the following section discusses device simulation setup and a solution to avoid such a problem.

2.3 Device Simulation Setup

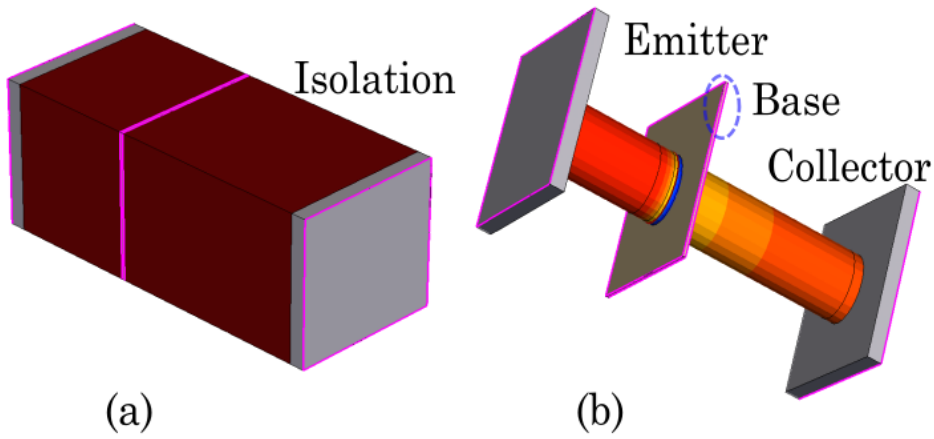


Figure 2.2: (a) 3D view (with Isolation) of SiGe nanowire device (the base contact is marked) (b) 3D view (without isolation) of the device depicting different regions and contact of the device. Note: Devices are vertical; they have been placed tilted for a better 3D view.

The SiGe nanowire HBT device is developed and analyzed using commercially available 3D Sentaurus TCAD. All simulations were conducted in a well-calibrated environment [76]. The results published in [49] substantiate our confidence on the choice of the adopted model parameters and overall calibration of the device simulation. The hydrodynamic model (HD) solves the current density equations. The Slootboom bandgap narrowing model accounts for the effect of carrier concentration at heavy doping caused by substantial doping in the emitter and buried collector layer. Along with the above-discussed models, Shockley-Read-Hall and Auger

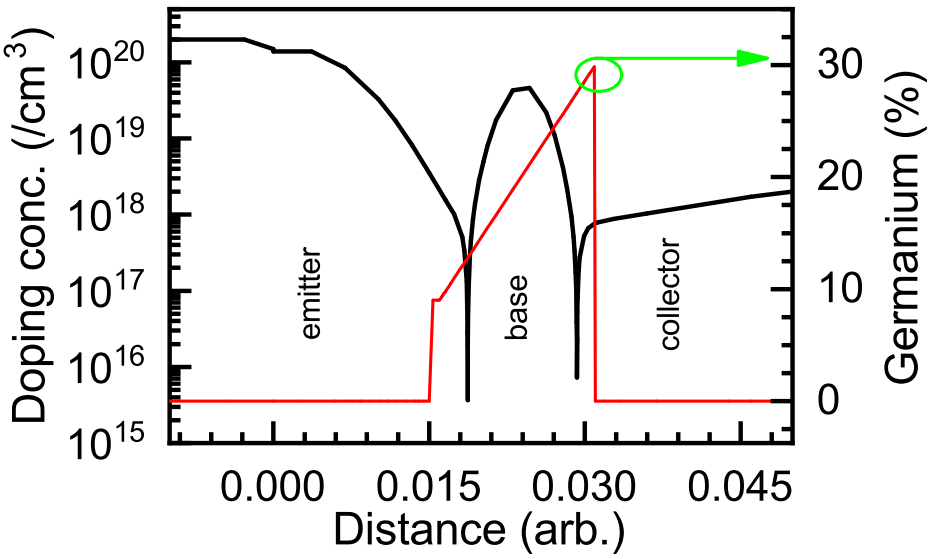


Figure 2.3: The doping profile of the device in cm^{-3}, germanium mole fraction (in %) shown on the right of y-axis.

recombination models are used to account for the recombination components; the Caughey-Thomas mobility model has been incorporated in the device physics deck as mobility plays a vital role in the bipolar device operation.

Fig.2.2 (a) and (b) illustrate the 3D schematic of the device with and without isolation oxide, respectively. It consists of an emitter, a base, and a collector region; additionally, the isolation oxide prevents contact from shorting. The following section discusses the precise size of the device's various areas. Arsenic is used for doping both the emitter and collector. For SiGe base realization, the base is boron-doped with germanium, and the contact extensions are made of nickel silicide. The device process flow will be identical to that of the intrinsic part of ST Microelectronics' B55 device [40], [33] in order to achieve the same vertical doping profile. The doping profile displayed in Fig.2.3 is identical to that of B55, which was previously published in our study [49], but with a reduced collector height to accommodate a feasible nanowire.

2.4 Approach, Results and Discussion

The Gummel characteristics of three different devices corresponding those shown in Fig.2.4 are presented in Fig.2.5. The devices are classified as follows: (i) device with a uniform radius of 40 nm from emitter to collector, with polysilicon as base contact (refer Fig.2.4(a)), (ii) same device as in Fig.2.4(a), with polysilicon contact replaced with metal as base contact (refer Fig.2.4(b)), and (iii) metal base contact same as in Fig.2.4(b) with both complete etched emitter and partially etched base (EPEB) (refer Fig.2.4(c)).

In case Fig.2.4(c), the collector radius is the same as 40 nm, but the EPEB radius is 25 nm. Now, suppose we observe the Gummel characteristics of the device corresponding to Fig.2.4(a) where the base contact is polysilicon. The base current is quite low, yielding a high DC current gain of approximately $\sim$5500. When the same device in Fig.2.4(a) is compared to the device of

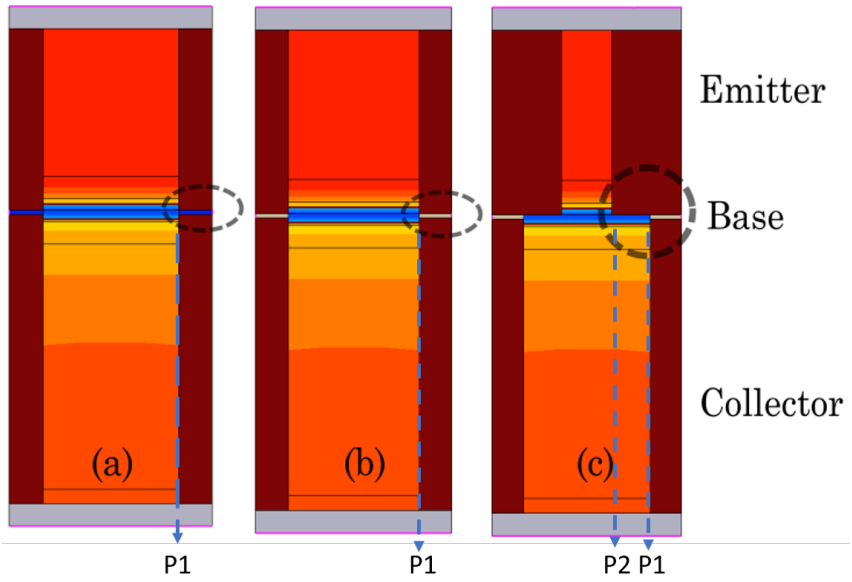
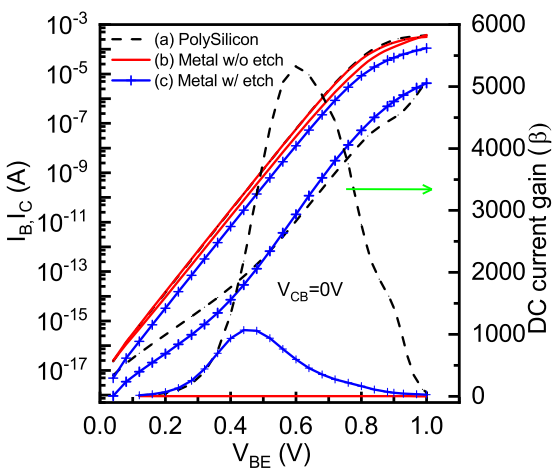


Figure 2.4: (a) Device with a polysilicon base contact (b) with metal base contact (c) with fully etched emitter and partially etched base (EPEB).

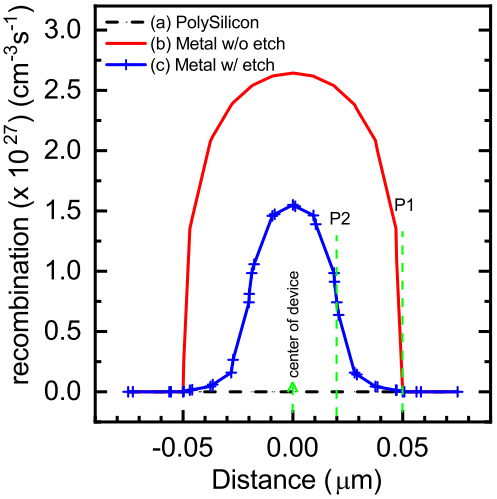
Fig.2.4(b) where the polysilicon base contact is replaced by a metal contact, although the collector current of both the devices remains approximately same, the base current increases significantly, leading to a significant increase in emitter current. The need for a device Fig.2.4(b) or a metal base contact instead of polysilicon contact arises because during fabrication, depositing an isotropic polysilicon layer is nearly impossible, and the device in Fig.2.4(a) cannot be realized with existing state-of-the-art technology. However, the problem with the device, as discussed in Fig.2.4(b), is that, while applying metal base contact, the base current increases and the DC current gain is reduced significantly. This is because electrons while getting transferred from the emitter to the collector, recombine at the base-metal peripheral interface. Indeed, the recombination at the metal-semiconductor interface is nearly infinite. To solve this issue, one needs to introduce a barrier in the carrier flow. The proposed technique shown in Fig.2.4(c) consists of an intentionally etched emitter and a partially etched base. This method partially suppresses the base recombination current, improving the DC current gain. Although due to this reduction in the emitter area (A_E), collector current reduces (not the collector current density) but most importantly, the recombination current in the base reduces, which is the primary requirement for a bipolar device to be useful for applications.

2.4.1 Development of the proposed device

Once the issue of base contact is resolved, one can compare the nanowire device's RF performance with the experimental B55 device from which the doping profile is adopted, as shown in Fig.2.3 and 2.6. We observe a significant improvement in the f_{MAX} while f_T is remaining nearly the same (see Fig.2.7). The purpose of comparing both the devices is that one may envisage using the B55 technology, realizing a large area transistor in its FEOL, and then applying the double etch process for fabricating the nanowires. Due to the reduction of lateral parasitics, i.e., reducing the distance from internal base to external base, the f_{MAX} is strongly improved according to the following equation (2.2):



(a)



(b)

Figure 2.5: Gummel characteristics (left) and recombination (right) comparison for the three devices, where (a), (b) and (c) corresponds to the same devices shown in Figure 3 respectively at V_{CB}=0 V.

$$f_{MAX} = \sqrt{\frac{f_T}{8\pi r_b C_{BC}}}. \tag{2.2}$$

In fact, due to the structural differences between the 5 μm width complex rectangular B55 and cylindrical NW device, normalizing the base resistance and base-collector (BC) capacitance is not as straightforward as it may appear. However, we can deduce from Fig.2.6 that there is a significant reduction in lateral parasitics, which contributes to the nanowire's f_{MAX} improvement.

As demonstrated in Fig.2.7 the change in current density (J_C=I_C/A_E) in NW compared to that of B55 is negligible. The collector current density of a bipolar device is given by the charge control relationship:

$$J_C = -\frac{q \exp\left(\frac{qV_{BE}}{kT}\right)}{\int_0^{W_B} \frac{p(x)}{n_i{}^2(x) D_{nB}(x)} dx}, \quad G_B = \int_0^{W_B} \frac{p(x)}{n_i{}^2(x) D_{nB}(x)} dx. \tag{2.3}$$

Here G_B known as base Gummel number and is calculated from position-dependent hole concentration $p(x)$, intrinsic carrier concentration $n_i(x)$ and minority carrier diffusivity $D_{nB}(x)$ together integrated over the base from 0 to W_B [112], [113]. As the doping and dimension are kept the same, hence base Gummel number G_B remains un-changed, which keeps the current density unchanged, although due to 3D effect there can be a little variation as already reflected however is negligible (valid when access resistance are negligible from low to medium current density).

2.4.2 Case I: Effect of collector area on DC and RF characteristics

As discussed previously, looking into the fabrication of a prospective EPEB device is a better choice and a good compromise for a practical solution; thus, the next step is to optimize the device's dimensions. To do so, the effect of varying the collector radius (c_r) (50 nm, 60 nm, and 70 nm) (refer Fig.2.8(a)) while maintaining a fixed emitter radius (e_r) of 25 nm on the device's various figure of merits is discussed.

The Gummel characteristics are plotted in Fig.2.8(b). From Fig.2.8(b), it is evident that when the emitter radius is fixed, there is no variation in the collector current (I_C); however, the base current (I_B) is reducing while changing from D1 to D3. As explained in the previous subsection, this is due to the greater distance between the emitter electron flow and the base metal contact, which reduces the recombination of these electrons. For the same case the transit frequency and maximum oscillation frequency are plotted in the Fig.2.9(a) at V_{CB}=0 V. It is observed that there is negligible variation in f_T (up to critical current density) with increase in collector radius from 50 nm to 70 nm. It is easily understood from (2.4), since emitter radius is fixed it is yielding negligible variation in I_C, g_m (=I_C/V_T), and C_{BE} as illustrated in Figs.2.8(b), 2.9(b) and 2.10(a).

$$f_T = \frac{1}{2\pi\tau_{ec}}, \tau_{ec} = \frac{C_{diff} + C_{BE} + C_{BC}}{g_m} + (r_e + r_c)C_{BC} \tag{2.4}$$

Here symbols carry their usual meaning.

While C_{BE} capacitance remained constant amongst the three devices, C_{BC} is simply scaling with the collector area as expected. Above peak f_T, we can observe that the device having the smaller collector area shows a more pronounced Kirk effect due to a reduced collector current spreading effect. Due to the expanded collector area, while changing from D1 to D3 the peak current density changes from 21 $mA/\mu m^2$ to 28 $mA/\mu m^2$. However, a variation of ∼180 GHz in f_{MAX} is observed.

Given that (refer to equation 2.2) the variation of f_T is negligible, two dominant parasitics of f_{MAX}, the base resistance (r_b) and collector-base capacitance (C_{BC}), have been analyzed to gain a better understanding of the trend of f_{MAX} variation with dimensions. The base resistance can be determined using different methods, as described in [114]. The circle impedance method (CIM) derives base resistance directly from the AC terminal characteristics. However, if we observe our device, we will notice that each device has an identical intrinsic base region opening, but the extrinsic base region dimension varies. Thus, if the base resistance is extracted from AC terminal characteristics, separation of the intrinsic and extrinsic components is nearly impossible [115]. Thus, the results extracted using the conventional method (= $\rho\frac{l}{A}$) described in [116] are presented here. The resistance is calculated from the physical parameters like carrier density and mobility using device simulation at each bias points in the main operating range. It's clear from Fig.2.11 that the r_{bi} is almost constant for the three devices; however, in the r_{bx}, the difference between D1 (c_r=50 nm) and D3 (c_r=70 nm) is significant, with a difference of ∼1.2 $k\Omega$. Due to D3's larger collector area, the base-collector capacitance in Fig.2.10(b) is the highest among D1, D2, and D3. Consequently, the f_{MAX} (refer Fig.2.9) for device D3 is the lowest of the three, owing to an increase in r_b and C_{BC}. The following subsection examines the effect of simultaneously varying the emitter and collector area on device performance.

2.4.3 Case II: Effect of both collector and emitter area on DC and RF characteristics

In this case, both the radii of the emitter and collector are varied, as shown in Fig.2.12(a), while maintaining the same collector to emitter area ratio. As shown in the Gummel plot in

Fig.2.12(b), the base and collector currents improve as the emitter and collector area increase. The increase in collector current can directly be attributed to the increase in the emitter area (A_E); in fact, the base current should also directly scale with A_E [21], however this is not true in this case. Indeed, the dominant part of the base current is the recombination component, hence a decrease in the base current is observed. Indeed, the increased gap between the metal contact and intrinsic region (P1-P2 as marked in Fig.2.4) helps in reducing the base current.

For the same case f_T and f_{MAX} are plotted in the Fig.2.13(a) at V_{CB}=0 V. It is observed that there is an increasing trend in f_T with increase in emitter radius from 25 nm to 35 nm, while the variation in f_{MAX} is less compared to the devices discussed in the last subsection. At low current densities, $\tau_F = C_{diff}/g_m$ in 2.4, is nearly constant for the devices. So the variation in f_T, can be understood given the variation in I_C, g_m (=I_C/V_T), and (C_{jEi}+C_{jCi})/g_m as illustrated in Figs.2.12(b), 2.13(b) and 2.14(a). For a better explanation of this variation in f_T, it will be better to apply geometry scaling based on the concept of the perimeter over area (PoA) separation. The general idea of PoA separation is based on the transistor partitioning as described in [6],[114]. For example, the total BE junction capacitance

$$C_{jE} = \bar{C}_{jEi} A_E + C'_{jEp} P_E, \tag{2.5}$$

is described by the internal (area normalized) and the perimeter normalized components $\bar{C}_{jEi}$ and C'_{jEp}, respectively. Where $C_{jEi} = \bar{C}_{jEi} A_E$, and $C_{jEp} = C'_{jEp} P_E$. $A_E = \pi r^2$, $P_E = 2\pi r$, and r denote the window area, perimeter, and radius, respectively. In eq (2.5), $\bar{C}_{jEi}$, and C'_{jEp} components are unknowns and needs to be determined from the known total capacitance C_{jE} by normalizing to area A_E,

$$\frac{C_{jE}}{A_E} = \bar{C}_{jEi} + C'_{jEp} \frac{P_E}{A_E} \tag{2.6}$$

is obtained, which can be used to perform the PoA separation. In fact, eq (2.6) is a linear equation with the axis intercept $\bar{C}_{jEi}$, and slope C'_{jEp}. Using different device sizes as in D1, D4 and D5 the above unknowns are calculated. Now plotting only the internal part (C_{jEi}+C_{jCi})/g_m as shown in Fig.2.14(a) we can observe in case of D5, this component is the lowest which leads to higher f_T.

To understand the f_{MAX} performance, we need to understand the base resistance variation clearly. For a cylindrical nanowire structure, the internal base resistance is geometry independent, and the observed difference should come from the external part only [117] which can be verified from Fig.2.11(a). Hence, we need to investigate the following case. let's consider r_1 inside radius, r_2 outside radius as shown in Fig.2.15 and σ is the resistivity, t thickness (the base length ∼12 nm from Fig.2.3). r_{bx} is the resistance between the inner and the outer ring (doughnut-shaped sample) with radial current flow: r_{bx} is given by: $dr_{bx} = \sigma dr\ / A(r)$, $A(r)$: surface perpendicular to current flow which reads $A(r) = 2\pi rt$. Integrating from r_1 to r_2 yields $r_{bx} = \sigma/2\pi t(ln(r_2/r_1))$, in this case, for all the investigated nanowire configurations $r_2/r_1 = 2$ i.e., constant, thus the external base resistance should remain same for D1, D4, D5. The numerical calculation results in a little variation, which is less than 4% of the total base resistance, which can be accounted to extraction error and can be ignored. That is the reason why in case I (refer Fig.2.11(b)), we can observe significant variation in r_{bx} compared to this case. While changing from D1 to D5, through D4, we can observe, on the one hand, an increase of f_T (which entails an increase in f_{MAX}), and on the other side, we observe an 45% increase in C_{BC} (from D1 to D5) due to an increase in the collector area, and nearly constant r_b, which entail a decrease of f_{MAX} from D1 to D5, through D4.

To observe the performance prediction, at higher reverse bias (V_{CB}) the f_T and f_{MAX} at V_{CB}=0.5 V for the devices in both case I & II are plotted in Fig.2.16. Due to increased reverse bias, base resistance (r_b) will not get affected; only the base-collector capacitance (C_{BC}) will vary, which is important for f_{MAX} extraction. This concludes that the f_T value remains nearly constant as device dimensions vary, and based on the f_{MAX} value, device D1 outperforms others; hence, D1 is chosen for the nanowire array performance study in the following section.

2.4.4 Effect of device to device separation on DC and RF performance

Due to the tiny dimensions of individual nanowires, they are unable to deliver the high current levels required for RF applications; however, an array of nanowires may deliver the requisite current level of I_T= number of nanowires (n) $\times$ individual nanowire current (I). The spacing between individual devices is critical for RF performance when arranging an array of devices. The devices in Fig.2.17 are separated by a distance w, which will be varied. Because the devices' contacts are biased at same voltage, the DC characteristics will remain constant regardless of the device-to-device spacing gap. It is worth noting that the individual device dimensions in the array are maintained at 25 nm for the emitter and 50 nm for the collector, which is the optimized choice from our previous studies. The simulation is computed with a high mesh count at the individual and overall array level for more accurate FoM extraction. If we plot the total capacitances including the isolation, $C_{BE,Total}$ and $C_{BC,Total}$ in Fig.2.18(a) for various device-to-device separation gaps, we can see that both junction capacitances increase with increasing w. This increase in capacitance decreases the f_T and f_{MAX} values. f_T decreases by $\sim$142 GHz ($\sim$50%) and f_{MAX} decreases by $\sim$215 GHz ($\sim$26%) when the device separation gap increases from 150 nm to 600 nm (refer to Fig.2.18(b)). As a result, a trade-off must be made between the minimum device-to-device distance that can be maintained from a manufacturing standpoint and the performance that the device engineer expects.

2.5 Proposed Fabrication Methodology

The nanowire array can be fabricated in both a bottom-up or top-down approach. The bottom-up approach is accomplished by the use of chemical vapour deposition (CVD) on a substrate. The top-down approach involves fabricating the device by selectively etching a patterned planner wafer. The advantage of bottom-up approach is we can have great freedom in choosing different materials, but in the case of top-down approach, it can be integrated quickly into an existing technology with more densely spaced nanowires. Because the device is intended to be manufactured using existing technology, a top-down approach is the best choice. The suggested technology combines e-beam lithography, plasma etching, and self-limited sacrificial oxidation and, as proved by the [94] procedure based on stress delayed oxidation, enables the manufacture of regulated nanowires with reduced dimensional variability.

It starts with a wafer having a doping profile, as shown in Fig.2.19, which is the same as Fig.2.3, and then it can be plasma etched up to the first layer, as shown in Fig.2.19(b). In the next step, isolation oxide can be deposited up to level 1 (refer Fig.2.19(c)). A nickel silicide layer of a thin metal layer is to be deposited on top of it throughout the surface only to bring out the contact for the base. This is a very challenging task considering the base thickness is less than 15 nm, which needs a quasi nanometer accuracy. The silicon layer is again etched to get the narrow emitter region (EEPB region as discussed in Fig.2.4). Then on top of it, a second isolation oxide layer is deposited throughout the surface. Finally, the top contact region of 15 nm to realize the final SiGe HBT nanowire array.

2.6 Conclusion

For the first time a SiGe HBT device based on a vertical nanowire architecture is proposed. The issue with fabricating the device is discussed, and solutions are proposed. The device dimensions are optimized varying the collector and emitter radius. The influence of parasitics on the RF performance of a nanowire array is analyzed using TCAD simulation. While devices placed with a smaller separation gap provide superior RF performance while fabricating, the proximity to which they can be placed must be carefully determined.

One major challenge with the device realization is that the base contact formed is relatively thin, which can project some challenges during fabrication.

Overall a f_{MAX} above 900 GHz is projected, which may further be increased by optimizing the device's dimension and buried layer. Also, in this case, the vertical profile is kept the same as B55. If a vertical profile with a thinner base is considered, it can improve the f_T and consequently the f_{MAX}. Suppose a method to deposit polysilicon anisotropically can be developed in the future; in that case, it will make the device realization easier, saving a lot of chip area and delivering improved RF performance.

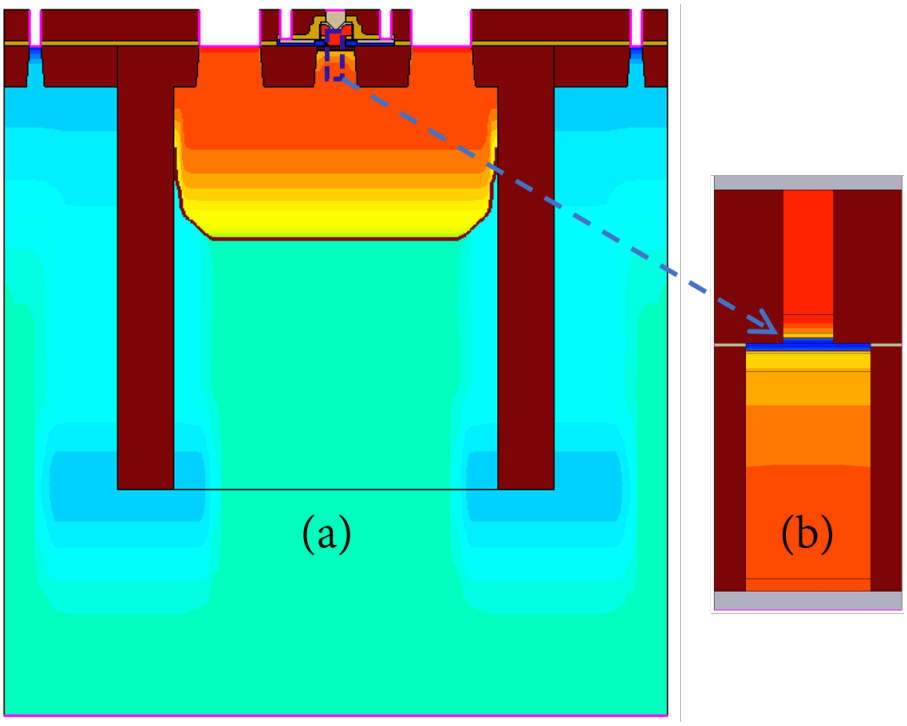
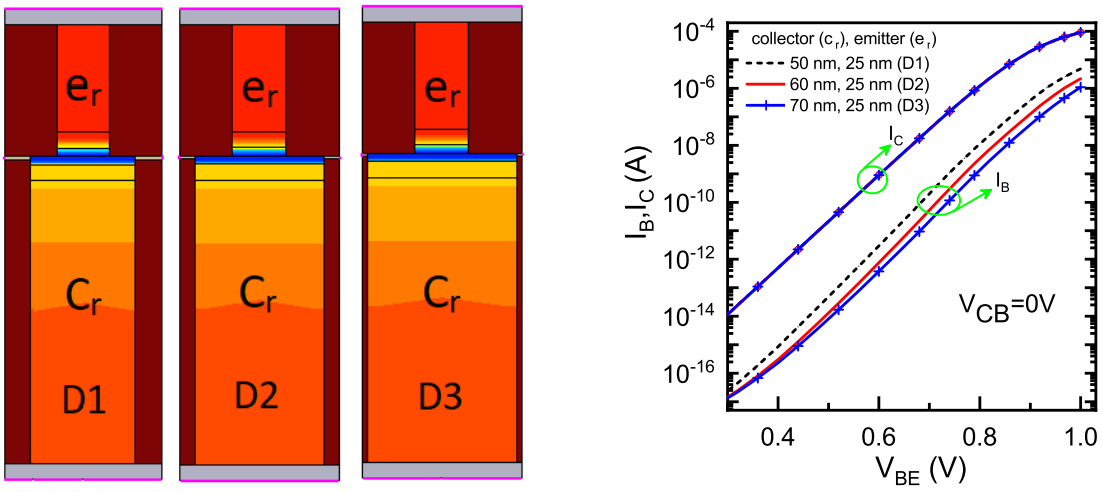


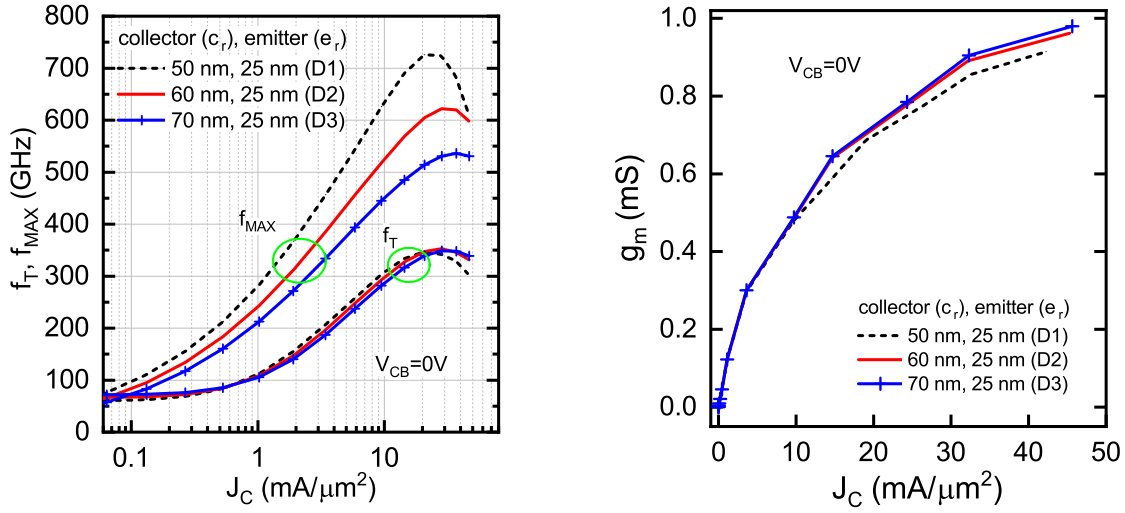
Figure 2.6: (a) STM's B55 experimental device [49], (b) Proposed nanowire device.

Figure 2.7: Comparison of f_T and f_{MAX} of B55 device (A_E=0.09 $\mu m \times 4.8 \mu m$) and proposed nanowire (A_E=$\pi(0.025)^2$, where the current density are normalized to corresponding areas.



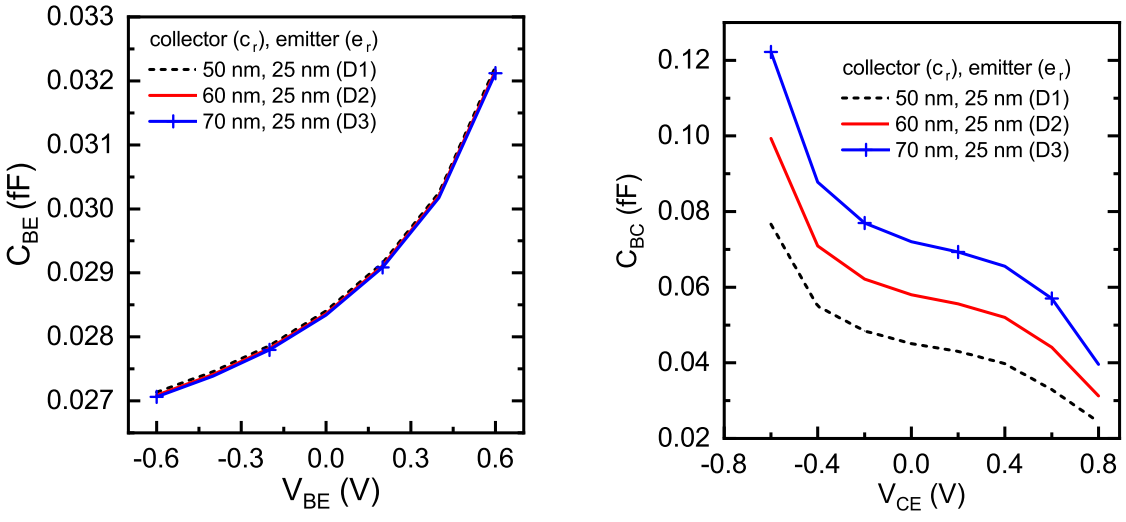
(a) (b)

Figure 2.8: (a) Nanowire device as finalized from last section with fixed emitter radius (e_r) of 25 nm, and variation in collector radius (c_r) as 50 nm (D1), 60 nm (D2), 70 nm (D3), (b) comparison of Gummel characteristics corresponding to D1, D2 and D3 as shown in (a).



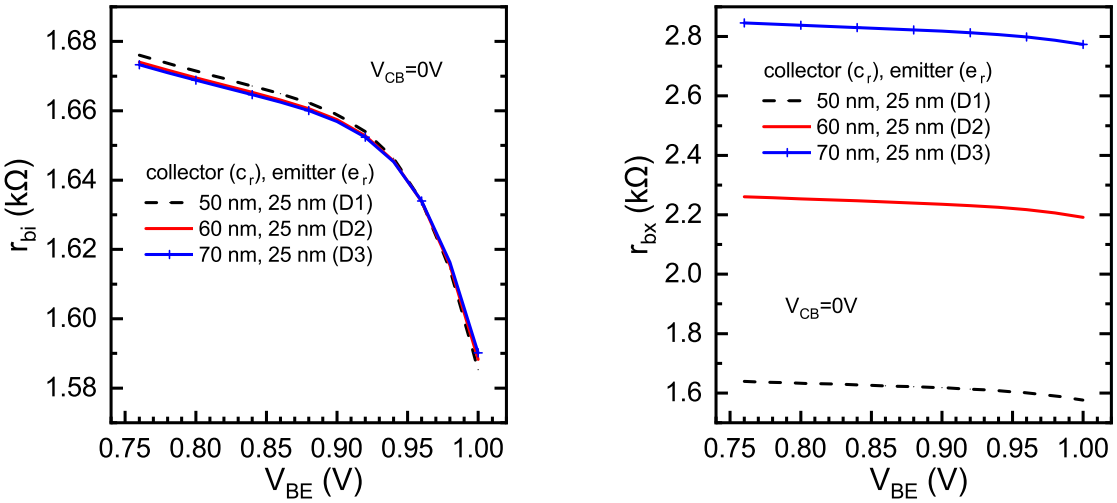
(a) (b)

Figure 2.9: (a) f_T and f_{MAX} (b) g_m vs. normalized J_C w.r.t device emitter area at (a) V_{CB}=0 V.



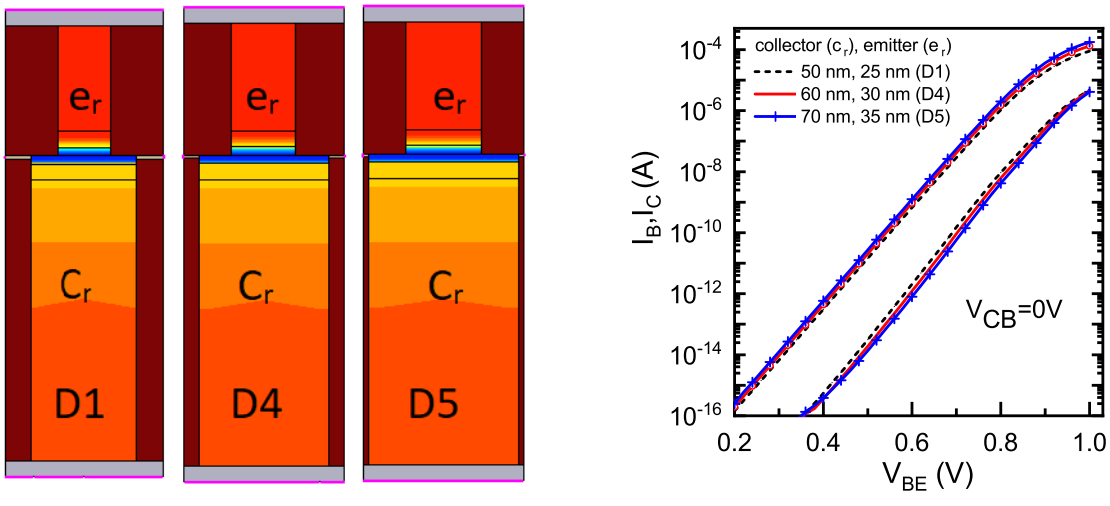
(a) (b)

Figure 2.10: (a) base-emitter, and (b) base-collector capacitance, corresponding to D1, D2 and D3.



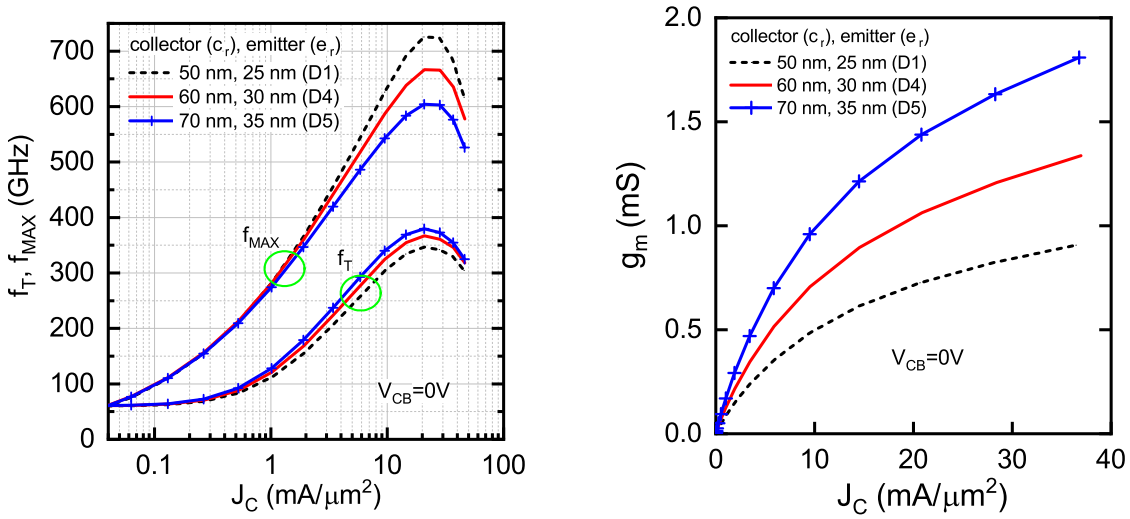
(a) (b)

Figure 2.11: (a) Variation in internal base resistance (r_{bi}) (b) only extrinsic base resistance (r_{bx}), for D1, D2 and D3.



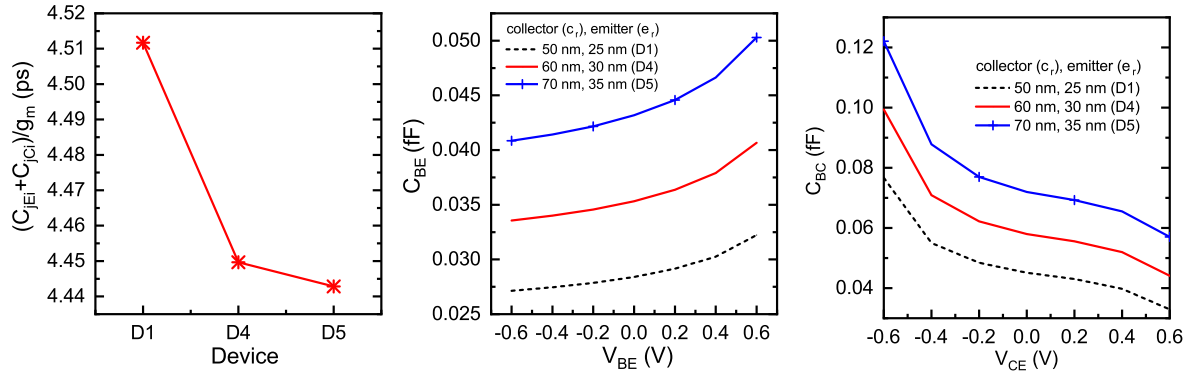
(a) (b)

Figure 2.12: (a) Devices with both varying EPEB and collector radius (e_r, c_r), c_r=50 nm, e_r=25 nm (D1), c_r=60 nm, e_r=30 nm (D4), c_r=70 nm, e_r=35 nm (D5) keeping same area ratio (b) comparison of Gummel characteristics corresponding to D1, D4 and D5 as shown in (a).



(a) (b)

Figure 2.13: (a) f_T and f_{MAX} (b) g_m vs. normalized J_C w.r.t device emitter area at (a) V_{CB}=0 V.



(a) (b) (c)

Figure 2.14: (a) Internal component $(C_{jEi}+C_{jCi})/g_m$ (b) base-emitter, and (c) base-collector capacitance, corresponding to D1, D4 and D5.

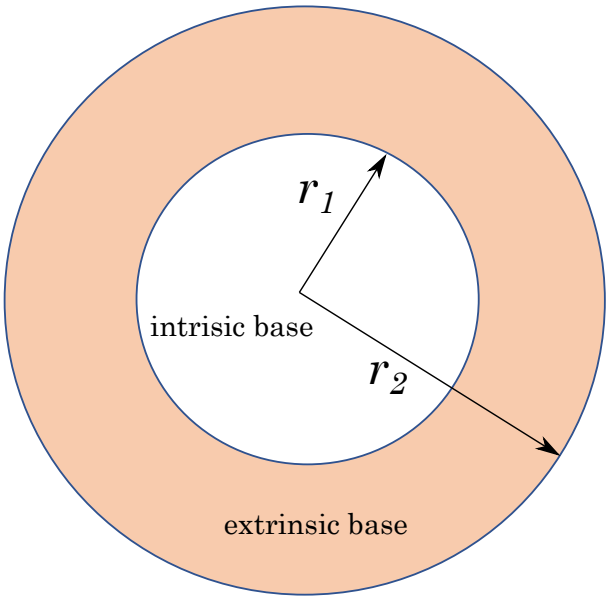
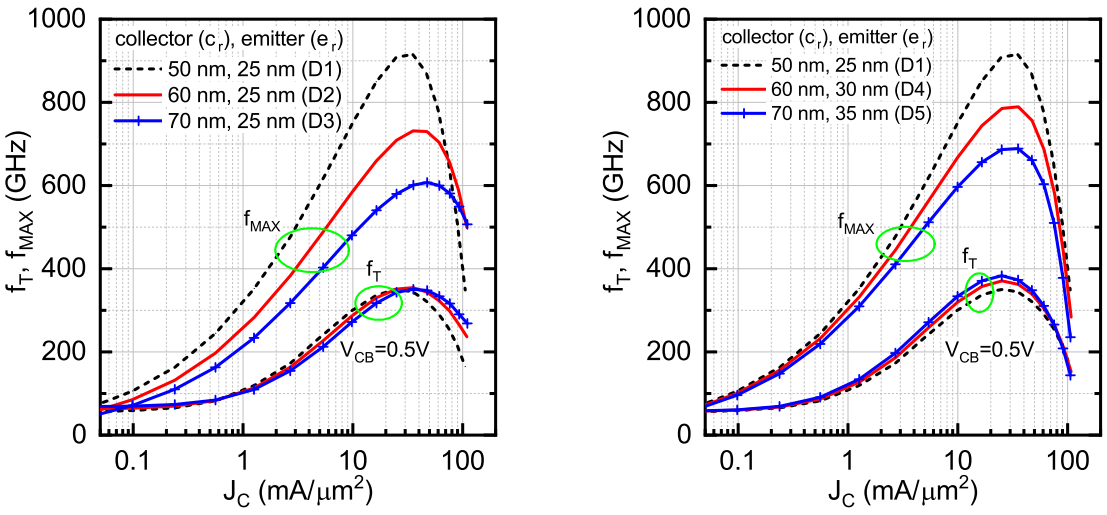


Figure 2.15: Depiction of different regions for analyzing the base resistance.



(a) (b)

Figure 2.16: f_T and f_{MAX} vs. normalized J_C w.r.t device emitter area at V_{CB}=0.5 V (a) for only c_r variation and (b) both e_r, c_r variation.

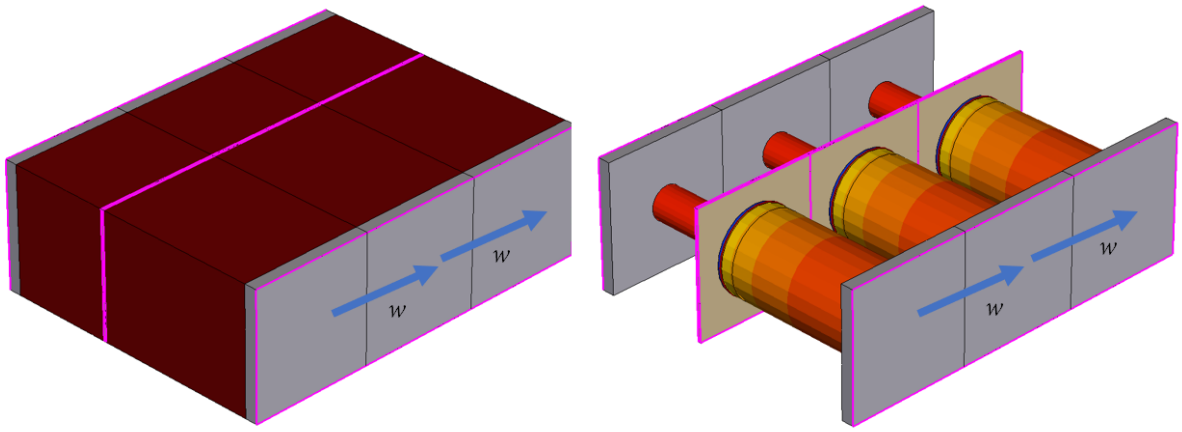


Figure 2.17: A set of 3 nanowires (a) with SiO_2 isolation (b) without isolation, the devices are separated by a distance (w) as marked. Note: The devices are vertical only for the 3D visualization effect placed tilted.

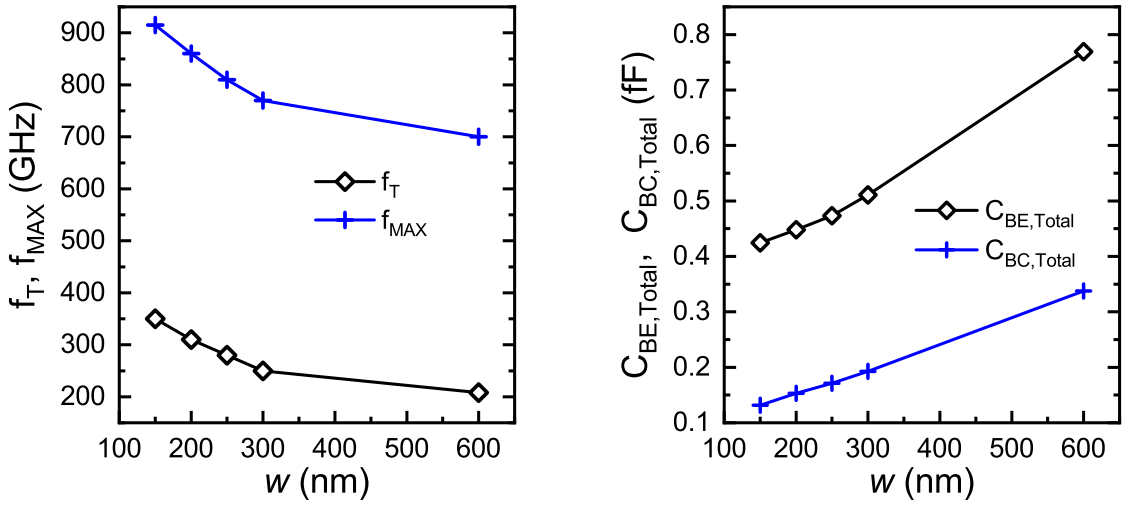
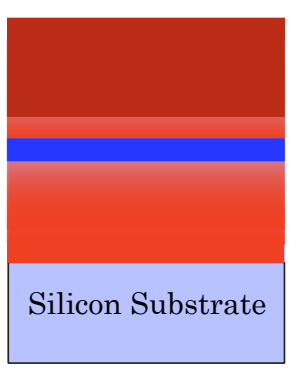
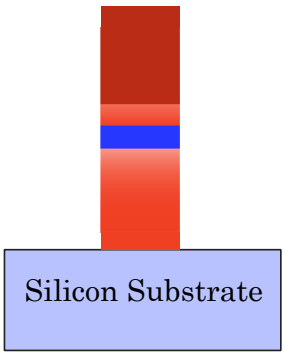


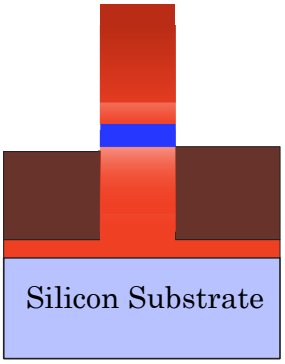
Figure 2.18: (a) f_T, f_{max} and (b) $C_{BE,Total}$, $C_{BC,Total}$ variation as a function of wire to wire gap (w) as shown in figure at V_{CB}=0.5V.



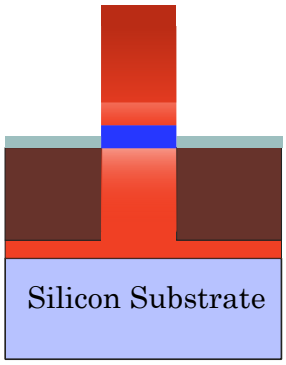
(a) Silicon substrate with doping and SiGe base



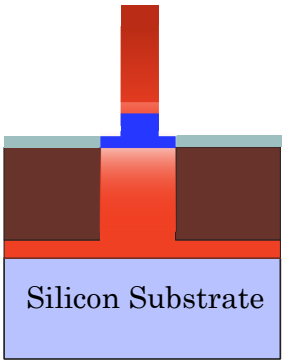
(b) Etching for the nanowire formation



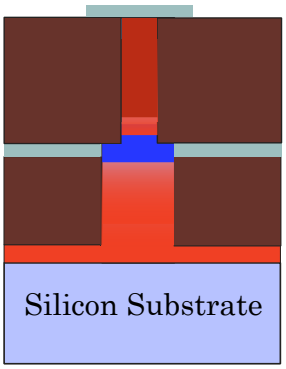
(c) Isolation oxide growth



(d) Base contact



(e) Etching for narrow emitter



(f) Oxide growth for isolation and emitter contact

Figure 2.19: The proposed fabrication steps.

Chapter 3

Substrate Effect on Asymmetric Lateral SiGe HBT for THz Applications

3.1 Introduction

The exponential growth in wireless communication (4G, 5G, and upcoming 6G) forced the need for smaller and faster transistors that can fuel the development of improved radio frequency integrated circuit (RFIC) technology. Now the need is to accommodate as many RFIC subsystems on a silicon substrate as possible. This will help the technology and economics to move in pace and fulfill the researchers' dream of truly system-on-chip (SoC) realization. BiCMOS process allows more flexibility in integrating complex circuits in a single chip. Bringing together structurally different CMOS and bipolar devices in a BiCMOS process demands an expensive high mask count process. This process is becoming increasingly expensive with shrinking technology nodes since the lithography resolution is getting scaled to achieve smaller and faster MOS devices. Further the mask count has increased due to additional passive components and other related issues. Therefore, each generation lithography advancement demands increased wafer cost that poses a significant obstacle in shifting towards RF SoC [118]. For SoC applications, silicon on insulator (SOI) substrate is best suited due to (i) superior quality of device isolation, (ii) reduction in leakage current, (iii) reduced cross-talk and noise due to substrate, and (iv) increased packing density.

Therefore, it is not surprising that most SOI-based electronics are CMOS-based VLSI and ULSI chips. CMOS circuits fabricated on SOI enjoy lower power dissipation in standby and operating modes, improved speed, and a wider temperature range. Some prime examples of complex circuits realized on SOI include the high-performance $POWER7^{TM}$ microprocessor [119], general-purpose UART/JTAG ($GPIO\ 3V3$), printer/LED ($SRC\ 3V3$) and many more can be found from ST Microelectronics's 28 nm FD-SOI technology application catalog [120]. The general consensus is that, compared to bulk silicon, SOI offers 20-30% performance improvement at the same operating voltage. If compared at similar low-power dissipation, the gain can easily be doubled. This is very significant since the performance improvement is similar to a generation of lithography migration in IC processes. SOI-based circuits realized by the n^{th} generation lithography process can outperform Si-based circuits with the $(n+1)^{th}$ generation process. The ultimate advantage of SOI is not limited to its superior performance compared to silicon. Due to the additional costs of SOI wafers, most ICs are still built on silicon wafers. However, SOI offers one most important advantage that is unmatched by bulk-Si substrate. The thin silicon film which hosts the active devices can be adjusted as a tunable parameter for device miniaturization. SOI-based CMOS devices can be shrunk in the planner direction by lithography, and a thinner SOI layer can shrink the device thickness.

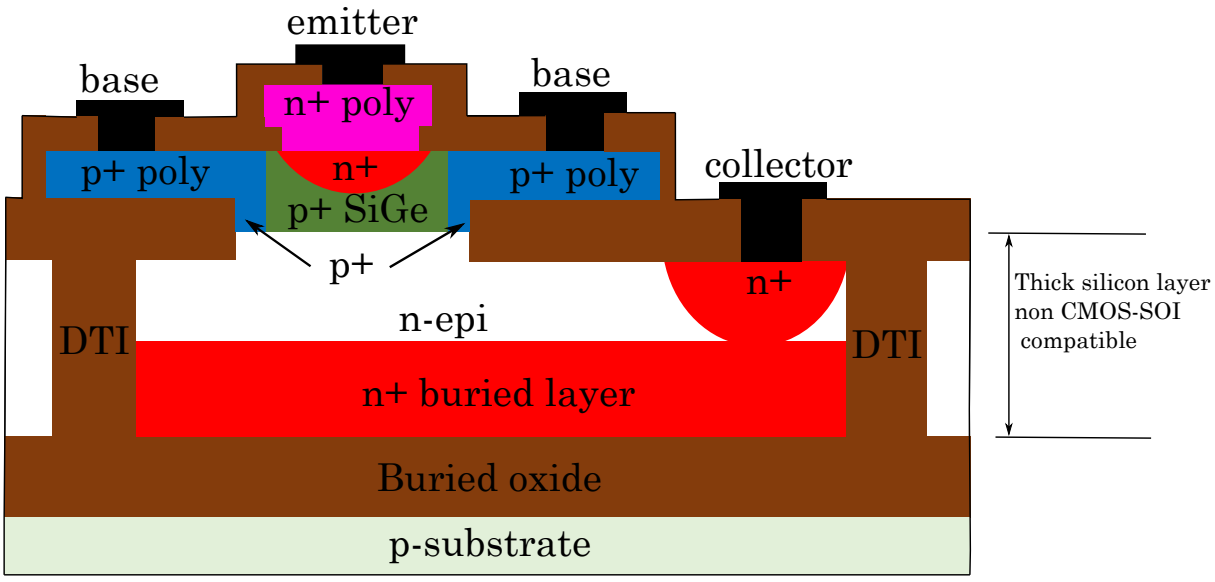


Figure 3.1: Schematic of SiGe HBT on thick SOI layer to accommodate collector buried layer of vertical device.

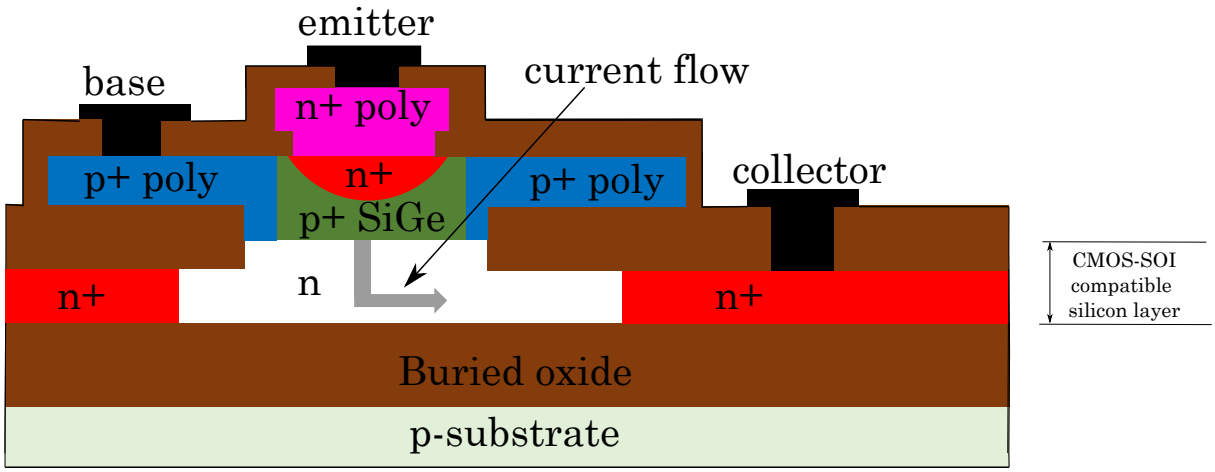


Figure 3.2: Schematic of SiGe HBT on thin SOI layer to make it CMOS compatible.

Although the CMOS transistors enjoy the above discussed added advantages of SOI, integrating SiGe HBT with CMOS on SOI is still in progress and has not yet matured. HBT on SOI are implemented with two approaches. One method is to maintain the original vertical HBT structure, which requires a deep n-well and buried layer, as shown in Fig.3.1. Comparatively, no significant advantage is reported in f_T and f_{MAX} when migrating from bulk-silicon to SOI [121]. The thickness of SOI is 1 μm, which is sufficiently thick in order to accommodate the vertical collector. However, with a very thick SOI layer, the benefits of short-channel effect suppression and parasitic capacitance reduction of CMOS disappear. This is undesirable in most cases since CMOS devices are the backbone of the digital block and occupy a major part of any mixed-signal design. The second method uses a thin-film SOI-CMOS compatible substrate to build the HBT. The collector region is built laterally along the SOI layer, as opposed to a vertical collector in the conventional design. During operation, the n-collector is fully depleted so that electrons travel laterally towards the n^+ collector contact, after been injected vertically from emitter to the base [121]. The cross-sectional schematic is shown in Fig.3.2, for direct comparison with thick SOI-based SiGe HBT in Fig.3.1. This method is ingenious in the

sense that the epitaxial structure of the HBT can remain intact to produce high performance base-emitter regions. Also, with a SOI-CMOS compatible SOI layer, no compromise is needed when implementing both CMOS and HBTs.

Though IBM and ST Microelectronics have already demonstrated vertical SiGe HBT on SOI layer, its architecture cannot be implemented on advanced SOI nodes such as the 28 nm technology having a thin Si film of less than 10 nm [122], [123], [42]. Therefore, in this situation, lateral SiGe heterojunction bipolar devices come out to be a more interesting and economical choice. Here the current flows horizontally, similar to the MOS transistors. Additionally, the key figures of merit (FoMs) of the high frequency vertical SiGe HBTs are limited by the RC time constants originating from the external parasitics [93]. Adapting a lateral bipolar structure, the parasitic junction capacitances and collector & base resistances can be reduced, allowing further improvements in maximum oscillation frequency (f_{MAX}) and noise figure [110]. Here, the collector and base resistances are lowered because of the close proximity of the top contacts to active regions.

The chapter flow is as follows; first, an extensive literature survey related to the LHBT has been done, then the device architecture and a simulation setup are discussed, and in section 4.4, the influence of substrate bias on DC and RF performance, i.e., the complete device operation is described. Substrate doping also impacts device operation, so a comparison study at different substrate doping is done in section 4.5. The process simulation importance and the lateral SiGe HBT device process development are discussed in 4.6. Finally, the chapter is concluded in section 4.7.

3.2 Literature Survey

The work on the SOI-lateral bipolar junction transistor (LBJT) stems from the fact that integrating bipolar laterally on SOI, together with CMOS, offers the best analog and digital devices on the same substrate. Both CMOS and LBJTs are built so that current flows horizontally, parallel to the wafer surface. Therefore, shrinking lithography, as well as thinner SOI films, benefits both devices. Strained SOI that offers increased mobility have shown to benefit MOSFETs, and it remains to be seen if such technology will also benefit LBJTs. One of the first works LBJT on

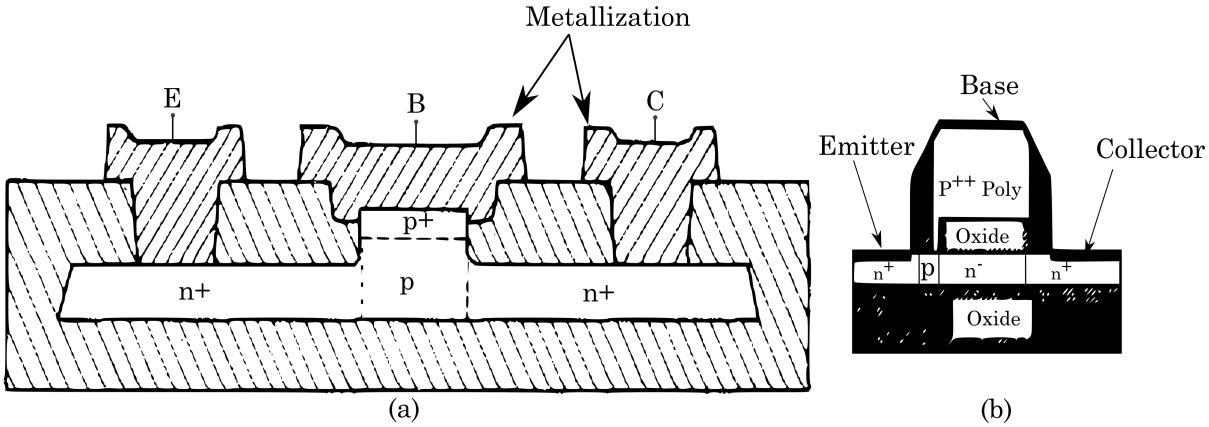


Figure 3.3: (a) Basic structure of a SOI LBJT proposed by [124], (b) SOI-LBJT with double-diffused base and a narrow emitter size proposed by [125].

SOI is realized by Rodder et al. [126]. This is one of the first works on silicon-on-insulator bipolar transistors. Sturm et al. [124] employed a structure that allows the self-aligned base contact

to be made with a minimum geometry in the base width. This elementary topology is as shown in Fig.3.3(a). It was widely adopted, and various modifications were made based on this work. Although a meager base resistance of 20 Ω is reported, the current gain was significantly low. Shahidi et al. [125] demonstrated an LBJT with double-diffused base and a narrow emitter size (refer Fig.3.3(b)) that was determined by the SOI thickness. The self-aligned base was realized using poly-side-wall-spacer (PSWS) for minimal parasitic junction capacitance. The achieved f_T reaches 20 GHz, but the BV_{CEO} breakdown voltage is a significant trade-off at only 2.8 V.

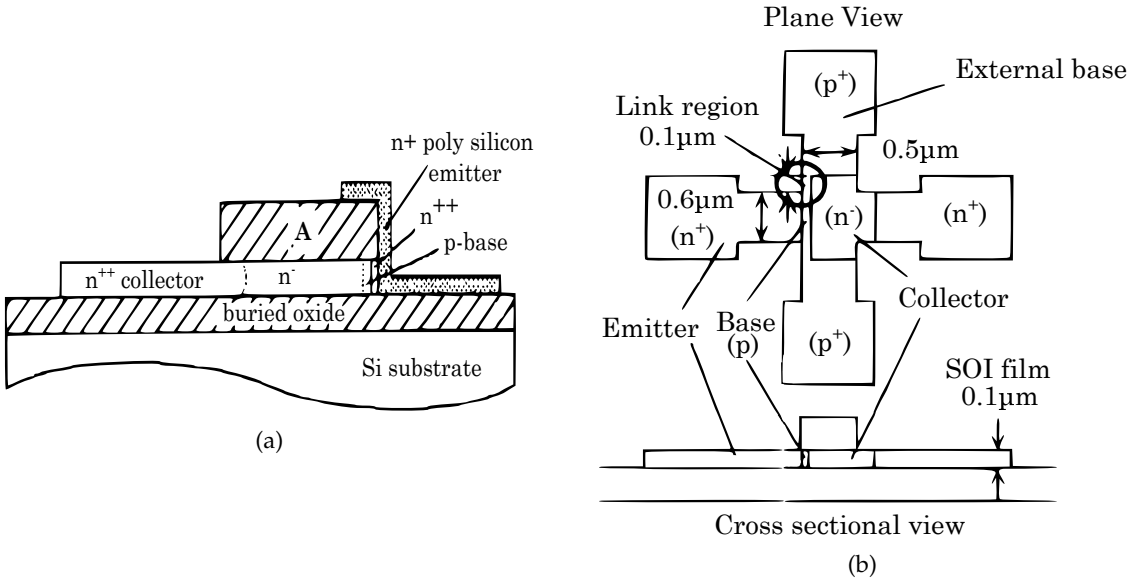


Figure 3.4: (a) A schematic after collector, base laterally diffused under the oxide and diffusion of the poly-silicon emitter [127], (b) proposed structure of lateral SOI BJT by [128].

[127] et al. realized an ultra-low-power LBJT as shown in Fig.3.4(a) employing a small active emitter area. The device with an emitter area of 0.15 μm^2 can operate at a peak f_T/f_{MAX} of 16/13 GHz, using only 15 μA of current. However, the problems are that the BV_{CEO} is only 2.5 V, and the ratio of emitter area to overall device area is significantly higher than other vertical or lateral bipolar junction transistors. Shino et al. [128] realized a low power SOI-LBJT by employing a self-aligned external base formation process (refer Fig.3.4(b)). This novel topology uses a tilted implant to form the shallow p-base region. Due to the reduced parasitics, a comparatively high f_{MAX} of 31 GHz is demonstrated. Nii et al. [129] further optimized this process with cobalt-silicide base electrode contacts (refer Fig.3.5(a)), comparatively high f_{MAX} of 67 GHz is achieved. The device also delivers sufficient f_T and BV_{CEO} of 12 GHz and 5.3 V, respectively. The most interesting LBJT structure is demonstrated by Suligoj et al. [130], which realized a high-performance LBJT on bulk-silicon substrate. This horizontal current bipolar transistor (HCBT) was realized using sophisticated processing sequence and chemical and mechanical processing (CMP) to provide silicon pillar surrounded by isolated oxide islands (see Fig.3.5(b)). The HCBT is built laterally on a silicon pillar, which has a very similar topology to an SOI-LBJT. The reported performance are f_T/f_{MAX} of 30/35 GHz, with BV_{CEO} of 4.2 V. Novel poly silicon side wall spacer architecture on SOI (see Fig.3.6), fabricated by [121] et al. demonstrated a f_T/f_{MAX} of 16/46 GHz, with a BV_{CEO} of 11.9 V.

The fact that LBJTs are implemented laterally on SOI has limited the migration from a homojunction bipolar into a heterojunction bipolar, significantly benefiting the vertical BJTs. In order to build a SiGe heterojunction for enhanced speed performance, the SOI film will either have to be doped with Ge, or the SiGe epitaxial film has to be selectively grown vertically adjacent

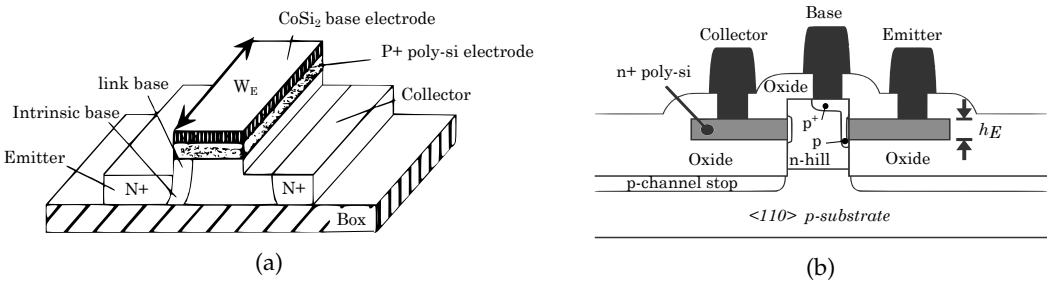


Figure 3.5: (a) Schematic of bipolar transistor with emitter size 0.12 $\times$ 3.0 μm^2 and $CoSi_2$ base electrode [129], (b) final HCBT structure after CVD oxide deposition, CMP and base implantation by [130].

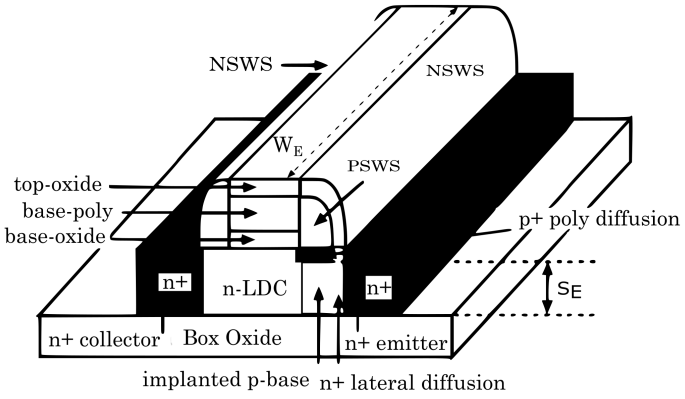


Figure 3.6: 3D cross sectional view of the poly silicon side wall spacer (PSWS) SOI LBJT [118].

to the base opening. Experimentally, this is a challenging task since selective epitaxial can only be realized on a planner surface, and the implementation of Ge has some potential limitations. In a SiGe-HBT, the Ge doping required is significantly high in order to make a suitable SiGe composition. This poses a problem for Ge ion-implantation, as ion-implantation equipment that can handle high dose Ge are not easily available. Also, high temperature annealing is required to re-crystallize the implanted SiGe layer. Kumar et al. [131] demonstrated a lateral SiGe PNP on SOI. This is the only experimental work by 2004, that has incorporated SiGe into the fabrication of LBJTs. It is to be noted here that no RF performance FoMs are reported in the publication.

With the sophistication of modern device engineering, most modern transistors are designed with CAD tools and device simulators to help the design process or predict the performance. Hamel et al. [110] proposed a lateral SiGe HBT technology that can be fabricated on SOI substrate. Although no experimental data is demonstrated, numerical analysis shows impressive results that have lateral SiGe-HBT out-performing vertical SiGe-HBT in speed as well as power dissipation. It also demonstrated that the lateral SiGe HBTs can significantly increase f_{MAX} while consuming lower current. Further analysis reveals that lateral SiGe HBTs with advanced scaling can operate at f_{MAX} in the neighborhood of 500 GHz. This theoretical study by [110] confirms two important assumptions of the advantage of LBJTs over VBJTs.

- The LBJTs with fewer parasitics components can operate at a higher frequency while consuming less current.

- The f_{MAX} of an LBJT can be a few times higher than VBJT due to the intrinsic connection to the base that results in low base resistance.

This type of transistor potentially offers a tremendous advantage for wireless applications. It can significantly impact the future of integrated RF/mixed-signal communication systems on a chip.

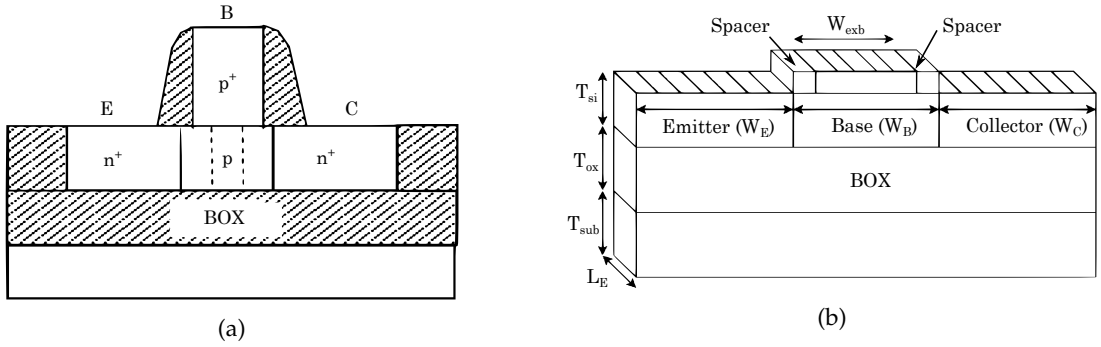


Figure 3.7: A symmetric SOI based LBJT proposed by [132], (b) 3D SOI-based lateral HBT proposed by [133].

Cai et al. [134], [132], [135] proposed a symmetric device (refer Fig.3.7(a)) that suppresses high current density degradation and base push out effects, and predicted a THz f_{MAX}. Raman et al. [77] and Derrickson et al. [133] with device topology as shown in Fig.3.7(b) predicted f_{MAX} up to 1.2 to 2 THz, respectively, with constant and graded germanium profiles. All performance factors published by different researchers as discussed in this section are summarized in Table 3.1 and 3.2.

Table 3.1: Summary of SOI Lateral BJTs published by different researchers/laboratories (1).

	[127]	[128]	[129]	[130]	[118]
Ext. base	Poly-Si	Poly-Si	Co-Silicide	Poly-Si	PSWS
Int. base	Silicon	Silicon	Silicon	Silicon	Silicon
$WB_{ext.}(\mu m)$	0.5	0.5	0.5	3	0.1
$f_T(GHz)$	15.5	9.8	12.6	4.4	12/16
$f_{MAX}(GHz)$	13	31	67	12	36/46
BV_{CEO}(V)	25.7/11.9	7.4	5.6	15.8	4.3/27.5
f_TxBV_{CEO}(GHz-V)	32.5	72.5	70.5	69.5	73.1/330

In [137], the proposed symmetric device (same topology as shown in Fig.3.7(a)) the collector and emitter are highly doped, making them insensitive to substrate back-bias, whereas the least-doped base region becomes highly sensitive to substrate bias. On the contrary, in this work, we propose a new asymmetric lateral SiGe HBT having a lightly doped collector that can be electro-statically adjusted by tuning substrate bias (V_{sub}). The light collector doping of 10^{17} cm^{-3} is very sensitive to the substrate back-bias and allows one to switch from a high-speed device to a high voltage device. This device achieves an f_{MAX} of 2.7 THz at V_{sub}=2 V with a BV_{CBO}=2.2 V and can switch to an f_{MAX} of 0.8 THz with a BV_{CBO}=3.6 V for a V_{sub}=-2 V. The high f_{MAX} observed is mainly due to the low internal base resistance. This is due to the

Table 3.2: Summary of SOI Lateral BJTs published by different researchers/laboratories (2)

	[136]	[134]	[132]	[77]	[133]
Ext. base	Poly-Si	Poly-Si	Poly-Si	Poly-Si	Poly-Si
Int. base	SiGe	SiGe	SiGe	SiGe	SiGe
$WB_{ext.}(\mu m)$	–	30	30	32	22
$f_T(GHz)$	47/47	270	350	812	1.2
$f_{MAX}(GHz)$	125/180	1000	700	1005	2000
BV_{CEO}(V)	3.9/2.5	–	–	–	1.9
f_TxBV_{CEO}(GHz-V)	183/190	–	–	–	3800

strong down-scaling of what is usually called the lateral dimension in conventional vertical HBT, equivalent to a technological node lower than 10 nm. Moreover, as explained above, the parasitics linked to the device's access are dramatically reduced. It is to be noted here in most of the cases f_{MAX} is higher than f_T and fortunately that's the demand by mm-wave and THz circuit researchers and circuit designers. People in the field are more interested in the power gain over current gain. In fact some researchers suggest that the circuit can operate at a frequency (f_{op}) above f_T as long as f_{MAX} is between 50 to 100% higher than f_{op} [54]. However, a technology is called balanced when f_{MAX}=1.5×f_T or f_{MAX}>f_T>f_{MAX}/2 [93], [138].

3.3 Simulation Setup and Device Architecture

In this work, the lateral SiGe HBT structure as shown in Fig.3.8 is investigated. Device simulations have been carried out using commercially available Sentaurus TCAD in a calibrated environment [72]. The results published in [49] substantiate our confidence on the choice of the adopted model parameters and overall calibration of the device simulation. The hydrodynamic model (HD) is used for solving the current density equations and the Slootboom bandgap narrowing model is used to account for the effect of carrier concentration at heavy doping. Caughey-Thomas mobility model is included; also the excess carrier recombination (Auger and Shockley-Read-Hall) models and breakdown effects at the high electric field (using Lackner's model) are included following the works of [77], [49], [40]. More details about the model and an analysis of the impact of different model parameters are elaborately discussed in chapter 2 of this thesis.

As shown in Fig.3.8, both the emitter and collector widths in the investigated SiGe HBT structure are fixed at 60 nm, while the physical base width and silicon layer thickness (T_{si}) are, respectively, kept at 28 nm and 7 nm in order to ensure compatibility with the STM's 28 nm FD-SOI technology. The collector is divided into two (a low doped (n) and a high doped (n^{++})) regions with appropriate doping levels given in Table 3.3. The BOX oxide thickness (T_{box}) is 25 nm and the device is extended across a width of 5 μm (in Z-direction).

In case of a lateral SiGe HBT, it is essential that the intrinsic and extrinsic base regions are fully self-aligned. Moreover, the extrinsic base must not be wider than the intrinsic base; else the unwanted widening of the base can cause a reduction in both transit frequency (f_T) and ac current gain at low frequency [110]. Therefore, in the simulation, the width of the extrinsic base is kept at 28 nm. Spacers are added on both sides of the base to avoid unwanted contact shorting. In the lateral npn SiGe HBT discussed here, the n-regions are doped with arsenic and

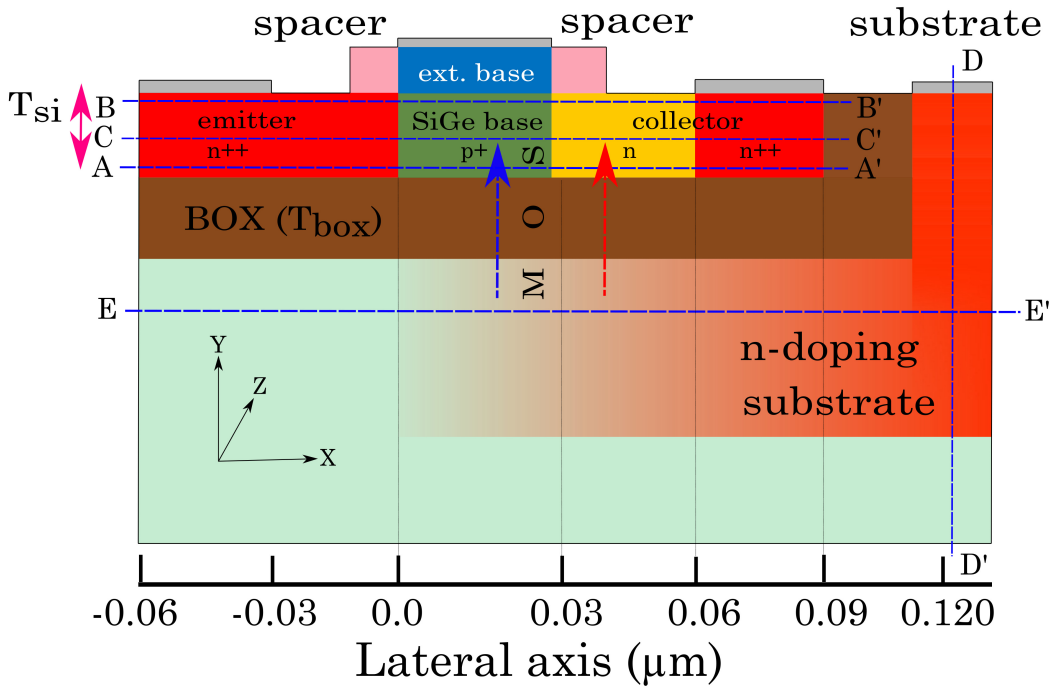


Figure 3.8: Cross sectional view of Lateral SiGe HBT device; the horizontal axis is marked in scale, T_{si} is 7 nm from the box/silicon interface. The blue colored dashed lines indicate different cut-lines used to plot spatial data in Figs.3.9, 3.10, 3.13(a), 3.14, and 3.24.

the p-region is doped with boron; and both have their respective advantages over the other possible dopants from a fabrication point of view. Additionally, the n-type substrate is doped with a Gaussian profile as shown in Fig.3.8, can be verified from Fig.3.24. The detailed dopant values and the regional dimensions are given in Table 3.3.

3.4 Impact of Substrate Bias on Device Performance

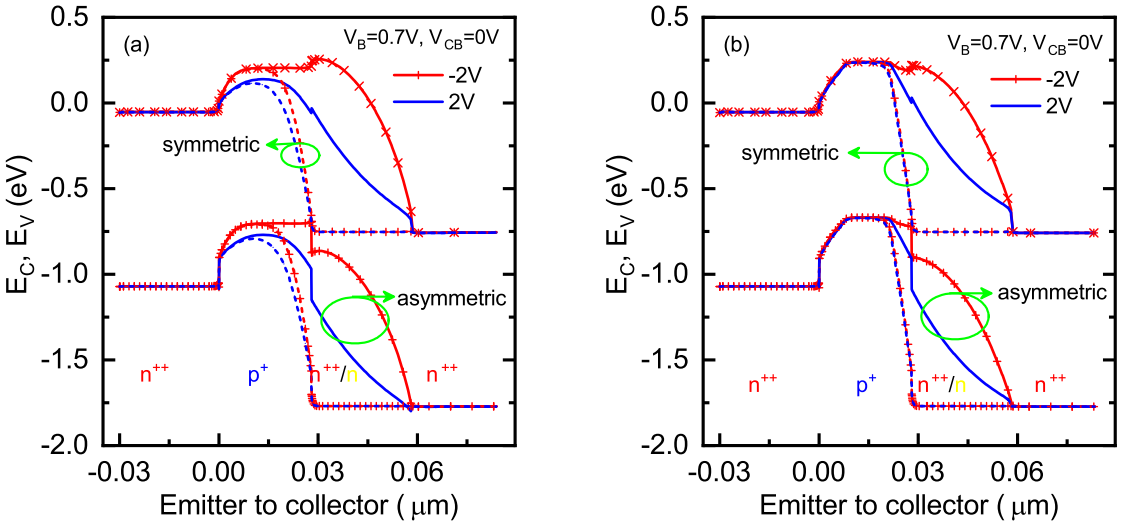
In this section first, we compare the physics of the symmetrical and asymmetrical HBT devices. It is necessary to mention here that the symmetrical HBT consists of identical highly doped emitter and collector regions, whereas in the asymmetrical device, the collector is subdivided into two sub-regions, (i) low doped (n) collector at a doping level as mentioned in Table 3.1, and (ii) heavily doped (n^{++}) collector region. It is to be noted here that, in later sections, unless mentioned specifically in the context of the asymmetric device, the collector will always mean to the low doped collector region. Afterwards, an in-depth analysis of the asymmetric device is given with an emphasis on achieving improved FoMs by tuning the substrate bias.

3.4.1 Comparison of Symmetric and Asymmetric Devices

In order to clarify the differences between the symmetric and asymmetric lateral HBTs, we need to understand the modulation of carrier density of different transistor regions under the varying substrate-bias. This is depicted in Fig.3.9 in terms of the conduction and valence bands plotted along the cut-lines A-A′ (Fig.3.9(a)) and B-B′ (Fig.3.9(b)) as mentioned in Fig.3.8.

Table 3.3: Dimensions and doping concentration of different regions of the device in Fig.3.8. In the symmetric device, the low doped collector (n) region is heavily doped and is identical to the emitter.

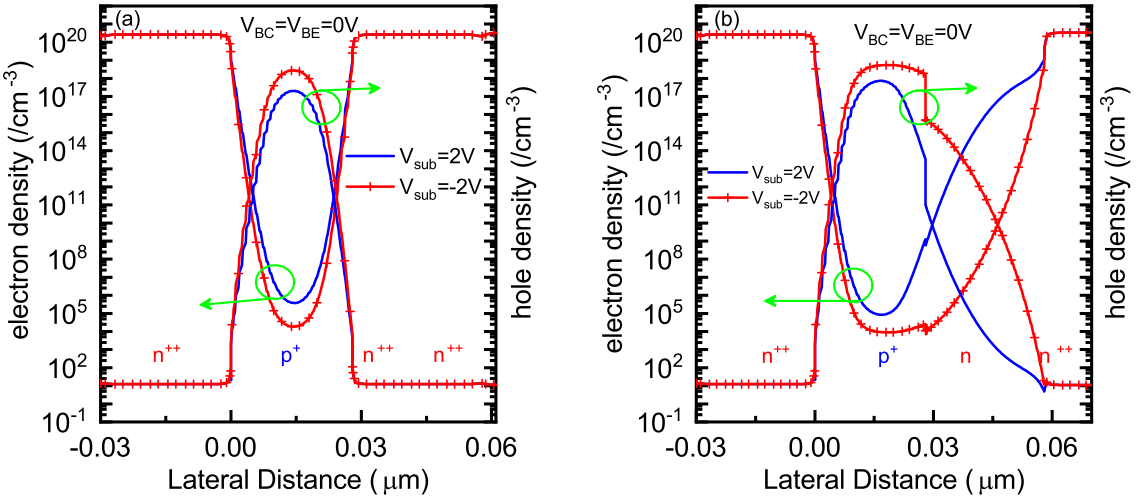
Region	Dimension (nm)	Doping (cm^{-3})
emitter (n^{++})	60	5×10^{20}
collector (n^{++})	30	5×10^{20}
collector (n)	30	1×10^{17}
base (p^{+})	28	3×10^{18}
extrinsic base	28	2×10^{20}
substrate	NA	1×10^{20} to 1×10^{15}
Ge	28	20%
T_{si}	7	–
T_{box}	25	–



(a) (b)

Figure 3.9: Conduction band and valence band (a) at the bottom (A-A′ in Fig.3.8) of the thin film SOI, (b) at the of the top (B-B′ in Fig.3.8) thin film SOI.

Note that the carrier density in the heavily doped emitter and n^{++} collector must remain insensitive to the substrate bias effects. We understand from the band diagram of Fig.3.9(a) that at close proximity of BOX/Si interface (cut-line A-A′ of Fig.3.8), carrier density only at the base region of the symmetric device is modulated whereas in case of the asymmetric device, the electron density at the relatively lower-doped collector region (having N_D=$1 \times 10^{17} cm^{-3}$) is impacted more compared with that of base (with $N_A = 3 \times 10^{18} cm^{-3}$). For the asymmetric device, it is observed that the substrate bias modulates the middle of the base region by about 67 meV whereas the middle of the lower-doped collector region by around 500 meV. On the other hand, Fig.3.9(b) shows that the band diagram at the top of the thin film (cut-line B-B′ in Fig.3.8), is hardly modulated within the base region for both the devices. Essentially the substrate bias-dependent field-effect is significantly reduced at a larger distance within the Si-film



(a) (b)

Figure 3.10: Spatial electron/hole density modulation in the base and collector regions with substrate bias variation (a) in symmetric (b) asymmetric devices at the middle of silicon (refer C-C′ -Fig.3.8).

away from the BOX. Comparing Fig.3.9(a) and (b) it appears that the low-doped collector of the asymmetric device is similarly modulated within all the thickness of the thin film due to its lighter doping. This modulation of the conduction band within the low-doped collector results into a modulation of the carrier density by more than three orders of magnitude as shown in Fig.3.10. The potential ($\phi_{B/C}$) within the Si film (base or collector) due to the applied substrate bias (V_{sub}) can be expressed as in (3.1)

$$\phi_{B/C} = V_{Sub} - \Delta Q_{B/C}/C_{BOX} \tag{3.1}$$

where $Q_{B/C}$ is the charge in the base or the collector region and C_{BOX} is the buried-oxide capacitance. Since the modulation of charge within the collector is much more than that in the base, the collector region potential is strongly affected. In the following part, we focus our attention to the asymmetric device to understand the impact of substrate bias onto its various performance FoMs.

3.4.2 DC Performance Analysis

Fig.3.11 presents the Gummel characteristics of asymmetric SiGe HBT under two different substrate biases. Due to the band lowering of the asymmetric device at the base-emitter junction (Fig.3.9(a)) at V_{sub}=2 V that allows more electrons and holes crossing from the emitter and base, respectively, it is observed in Fig.3.11 that both collector and base currents are increased. Additionally, the positive substrate bias results in a higher electron density in the low-doped collector region causing an increment in the space charge region inside the base. This leads to a shrinking in the quasi-neutral base (QNB) width subsequently increasing the collector current. Finally, the electron accumulation within the collector reduces the collector resistance (r_c) contributing to a higher collector current. Although the substrate field-effect tends to deplete the holes near the BOX/Si interface of the base region, the electron accumulation effect in the collector is significantly higher as the donor doping in the collector is more than one order less than the acceptor doping in the base region. From the output characteristics shown in Fig.3.12,

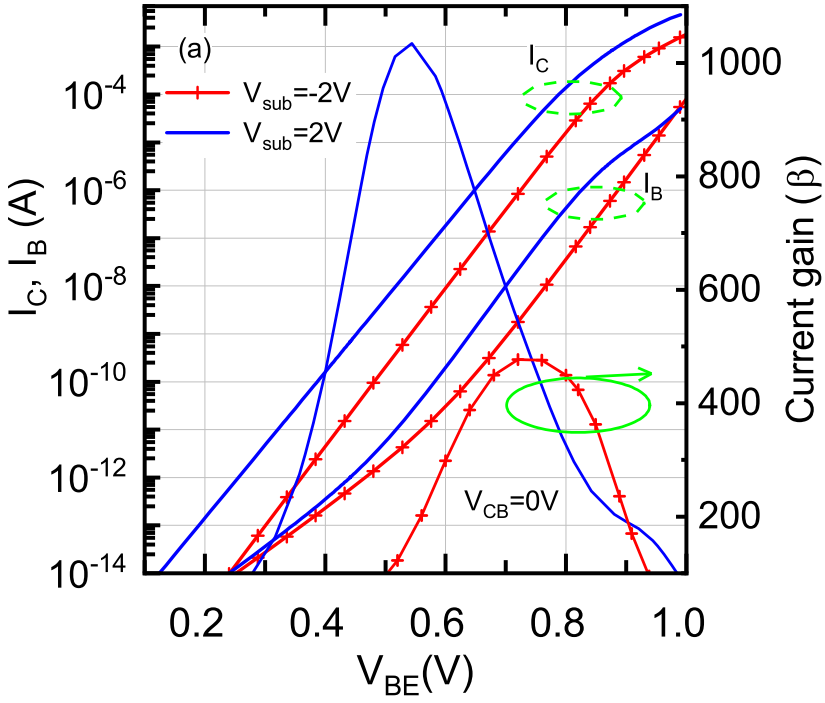


Figure 3.11: Gummel characteristics along with DC current gain (β) for asymmetric device at V_{sub}=-2 V and 2 V, $V_{CB} = 0\ V$.

it is observed that the collector current is increased by around 100% when the V_{sub} is varied from -2 V to 2 V; however, due to the reduction of the QNB width, Early voltage reduces and output conductance value increases.

The breakdown voltage determines the maximum operating voltage of a bipolar transistor in the particular transistor configuration, for different modes of operation. The current-voltage characteristics of the device measured under the condition of one floating electrode are used to determine the device's breakdown voltage. The BV_{CEO} is the collector-emitter breakdown voltage with the base open; the BV_{CBO} is the emitter-base breakdown voltage with collector open. Generally, the overall breakdown voltage of the transistor is dominated by BV_{CEO}.

Hence, we investigate the effect of V_{sub} onto the open base breakdown voltage (BV_{CEO}) that decides the maximum permissible collector-emitter voltage and subsequently the power rating at DC and RF operations. Electric field ($E(x)$) in the low-doped collector region as a function of V_{sub} can be expressed as

$$\frac{dE\,(x)}{dx} = \frac{q}{\varepsilon_{Si}}\left(N_C + \Delta n\,(V_{sub}) - \frac{j_n}{q v_n}\right) \tag{3.2}$$

where ε_{Si} is the silicon permittivity, N_C is the collector doping, v_n is the electron velocity, j_n is the electron current density and $\Delta n(Vsub)$ is the increased carrier density due to the positive substrate bias. In case of a negative V_{sub}, $\Delta n(Vsub)$ gets a negative value reducing the slope dE/dx in (3.2). This leads to a reduced electric field peak at the BC junction (see Fig.3.13(a)). Hence, the impact ionization effect becomes weaker resulting in a higher BV_{CE0} value. Fig.3.13(b) shows collector current at open base condition, for two different V_{sub}. It is observed that BV_{CE0} varies from 3.6 V to 2.2 V as V_{sub} changes from -2 V to 2 V. This way, one can achieve a relatively high

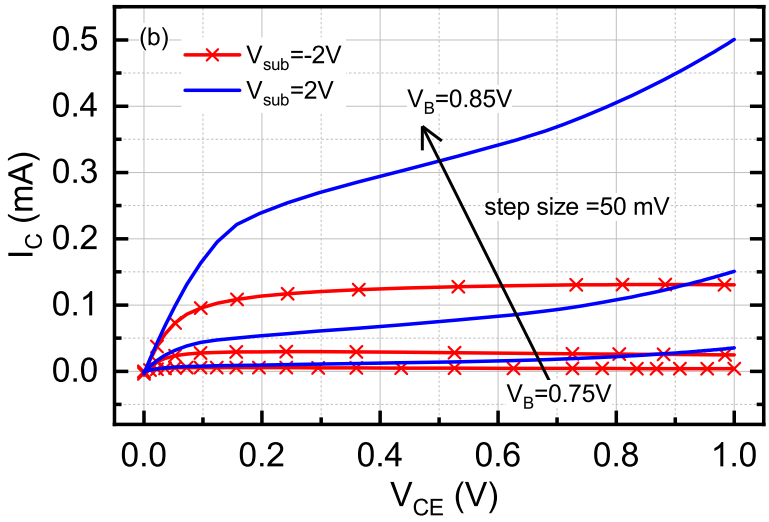


Figure 3.12: Forced-V_{BE} output characteristics at V_{sub}=-2 V and 2 V for asymmetric lateral SiGe HBT with W_{si}=5 μm.

power device just by tuning the substrate bias of a given lateral SiGe HBT.

3.4.3 Quasi Static Transit Time Analysis

The quasi-static analysis is one of the most helpful methods to investigate the RF performance of a high-frequency device. It is linked with the change in internal carrier densities due to applied terminal bias. To the applied ac signal, the device responds with modulation in internal stored charge. This change in stored charge causes a delay that limits the transistor gain at high frequencies. A relation between delay time (τ_{ec}) and stored charge in the transistor can be expressed as follows (3.3-3.4)

$$\tau_{ec} = \frac{\Delta Q_p}{\Delta I_C}. \tag{3.3}$$

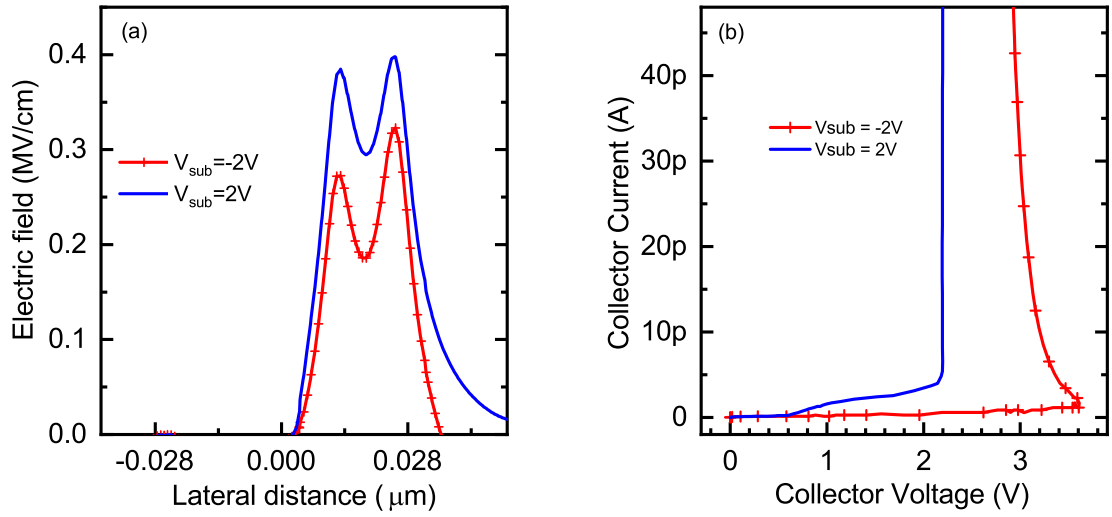
Due to charge neutrality ΔQ_p can also be written as ΔQ_n, so above equation becomes

$$\tau_{ec} = \frac{\Delta Q_n}{\Delta I_C}. \tag{3.4}$$

And this delay can be expressed in terms of transit frequency as (3.5)

$$\tau_{ec} = \frac{1}{2\pi f_T}. \tag{3.5}$$

Within the quasi-static approximation, ΔQ_n and ΔI_C can be considered as independent of frequency. This allows to calculate the device transit time from steady-state device simulation by a small change in DC signal. In the next section, a comparison is made between small-signal AC and quasi-static DC approximation. Since quasi-static approximation does not consider 2D effects, a deviation from small-signal AC extraction is obvious, but the trend is well explained.



(a) (b)

Figure 3.13: (a) Electric field distribution through C-C′ cut-line (see Fig.3.8) and (b) open base output characteristics showing BV_{CEO} values at V_{sub}=-2 V and 2 V.

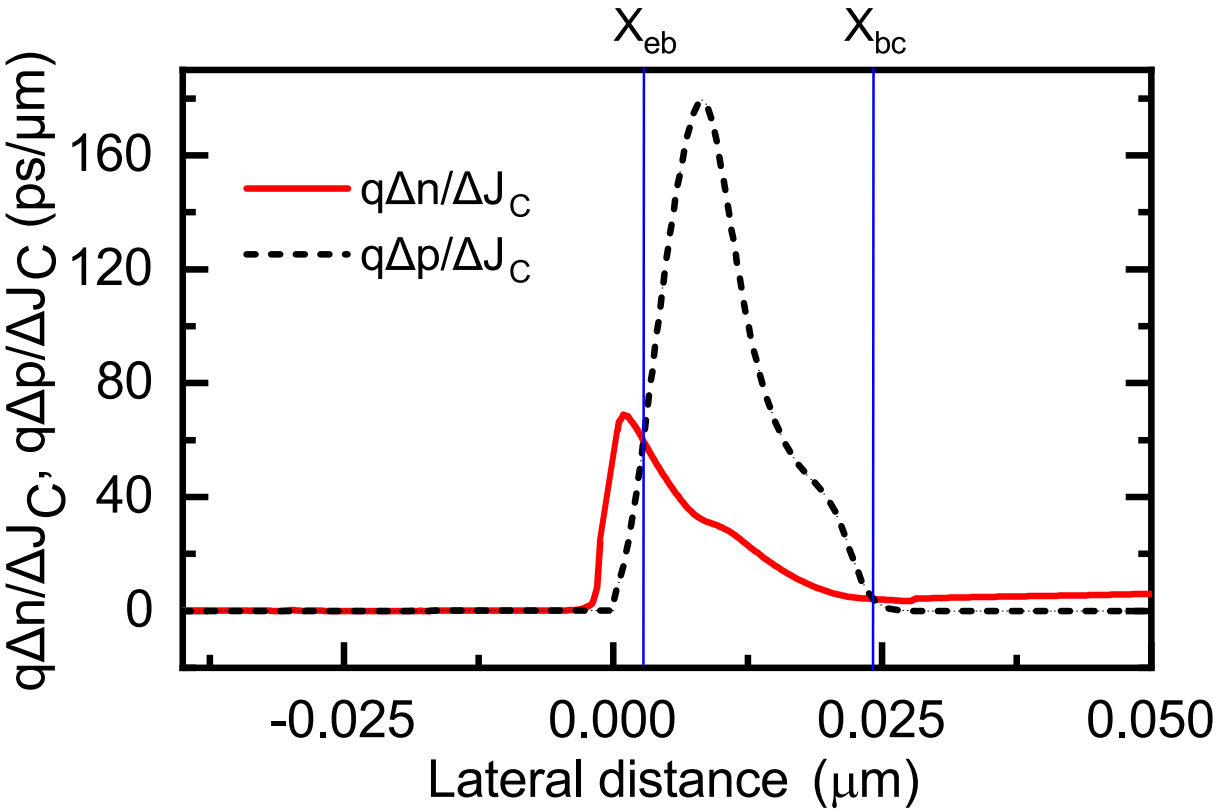


Figure 3.14: Spatial quasi-static electron and hole current density divided by total collector current density variation in the lateral direction, where X_{eb} and X_{bc} are indicating the boundary values for individual integration terms. Note, the extraction corresponding to this plot are done at peak f_T, similarly at each bias it can be extracted with updated X_{eb} and X_{bc}.

Regional Approach of Extracting Transit Time in SiGe Bipolar Device

Transit frequency (f_T) is the prominent figure of merit to characterize a high-frequency device. As already discussed, physically f_T is the reciprocal of the transit time it takes for the carries to

move from the emitter contact to the collector contact (τ_{ec}) under transistor action.

Due to application of a small perturbation of input bias variation, the variation in electron density (n) or hole density (p) yields τ_{ec}. In TCAD, quasi-static 1D simulation can help in extracting above parameters. As suggested by [139]

$$\tau_{ec} = q\int_{0}^{L}\left|\frac{dn}{dJ_c}\right|dx \tag{3.6}$$

or

$$\tau_{ec} = q\int_{0}^{L}\left|\frac{dp}{dJ_c}\right|dx \tag{3.7}$$

where, dn and dp is the change in electron and hole density, respectively, due to the applied perturbation, dJ_c is the change in total current density due to the same. And the condition is the small signal value of instantaneous collector-emitter voltage V_{ce}, i.e., v_{ce} must be zero. Integration limit $x = 0$ is the point at the emitter contact and $x = L$ at the collector contact.

The accumulated electron or hole transit time at any point "K" between $x = 0$ to $x = L$ can be expressed as [140]

$$\tau_{n,acc.}(K) = q\int_{0}^{K}\left|\frac{dn}{dJ_c}\right|dx \tag{3.8}$$

or

$$\tau_{p,acc.}(K) = q\int_{0}^{K}\left|\frac{dp}{dJ_c}\right|dx. \tag{3.9}$$

In this work, using the formulation described in (3.8-3.9) the transit time partitioning scheme demonstrated by [81] is used to investigate the individual delay times in different regions of the lateral device. In this approach the total transit time (τ_{ec}) is divided into five individual terms ($= \tau_e + \tau_{be} + \tau_b + \tau_{bc} + \tau_c$) consisting of terms, both related to the space charge regions (SCR) and neutral regions. The advantage of this partitioning scheme compared to other available schemes is that it just needs two boundary points for calculating all regional transit times [141]. The boundary points x_{eb} and x_{bc} are found from the crossing points of $q\frac{dn}{dj}$ and $q\frac{dp}{dj}$ as shown in Fig.3.11. Using this boundary points, all individual transit times can be calculated as

$$\tau_e = q\int_{0}^{x_{eb}}\left|\frac{dp}{dJ_c}\right|dx \tag{3.10}$$

$$\tau_{eb} = q\int_{0}^{x_{eb}}\left|\frac{dn}{dJ_c} - \frac{dp}{dJ_c}\right|dx \tag{3.11}$$

$$\tau_b = q \int_{x_{eb}}^{x_{bc}} \left| \frac{dp}{dJ_c} \right| dx \tag{3.12}$$

$$\tau_{bc} = q \int_{x_{bc}}^{L} \left| \frac{dn}{dJ_c} - \frac{dp}{dJ_c} \right| dx \tag{3.13}$$

$$\tau_c = q \int_{x_{bc}}^{L} \left| \frac{dp}{dJ_c} \right| dx. \tag{3.14}$$

The regional transit times like emitter (τ_e), base (τ_b), and collector (τ_c) components are related to the locally compensated minority charges in their respective neutral regions and are referred as diffusion charging times. The junction transit times at the base-emitter (τ_{eb}) and base-collector (τ_{bc}) regions are due to local uncompensated charge and are called depletion charging times.

There are other methods also suggested by [142], [143] for calculating the regional transit times separating both neutral and space charge regions (SCR). However, while applying these methods, finding the exact boundary points while calculating the SCR transit times during high injection operation regime is a cumbersome task as discussed in [144], [141].

Small Signal Analysis

To relate the transit time derived from quasi-static approach, an expression for τ_{ec} can be derived from the small-signal equivalent circuit model as shown in Fig.3.15. The high frequency current amplification is typically derived from the small-signal current gain (h_{21}), when the output terminal is shorted (see Fig.3.16).

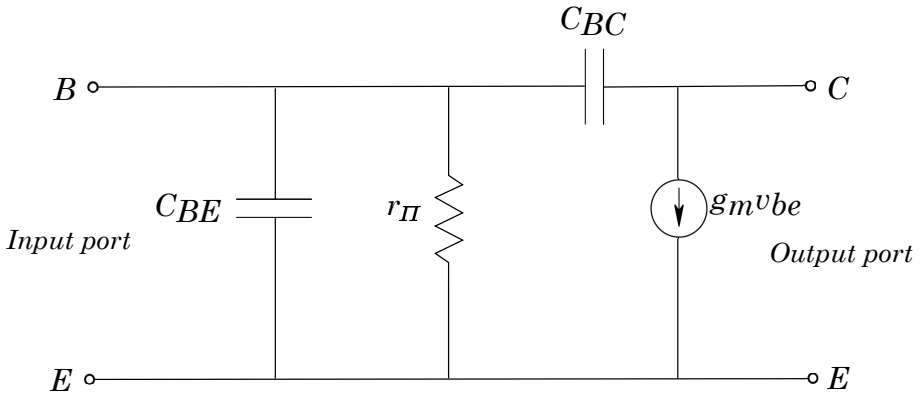


Figure 3.15: Equivalent circuit for small-signal analysis.

Imagine that the base/input terminal of the small-signal circuit is forced with a small-signal input current of $i_b = i_0 e^{j\omega t}$ while the output/collector terminal is shorted as shown in Fig.3.16. Using basic circuit theory the node voltage at base v_b can be derived as

$$v_b = \frac{1}{\frac{1}{r_\pi} + j\omega(C_{EB} + C_{BC})} i_b. \tag{3.15}$$

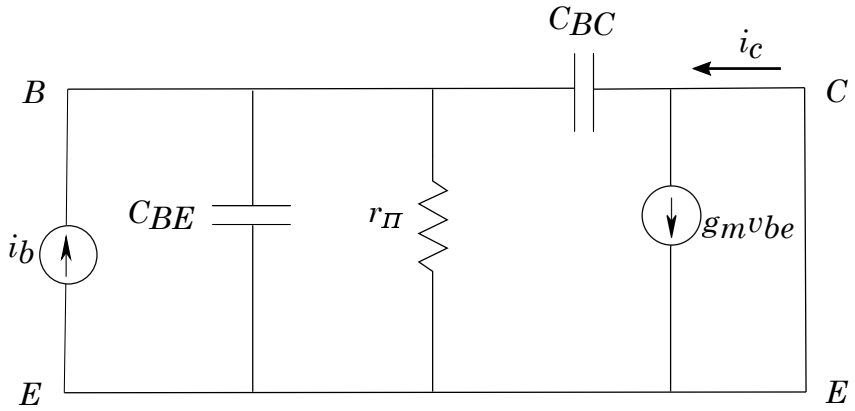


Figure 3.16: Equivalent circuit for small-signal analysis with output terminal shorted.

Similarly, the output current at collector i_c derived as

$$i_c \approx g_m v_b = \frac{g_m}{\frac{1}{r_\pi} + j\omega(C_{EB} + C_{BC})} i_b. \tag{3.16}$$

Therefore, the small-signal current gain (h_{21}) from eq(3.15-3.16) is given by

$$h_{21} = \left.\frac{i_c}{i_b}\right|_{v_{ce}=0} = \frac{g_m - j\omega C_{BC}}{\frac{1}{r_\pi} + j\omega(C_{EB} + C_{BC})}. \tag{3.17}$$

It is assumed that $g_m >> \omega C_{BC}$ at frequencies of interest, then the current gain equation is simplified to

$$h_{21} = \frac{g_m}{\frac{1}{r_\pi} + j\omega(C_{EB} + C_{BC})}. \tag{3.18}$$

Further, one assumes, $\frac{1}{r_\pi} << j\omega(C_{EB} + C_{BC})$ yielding

$$h_{21} = \frac{g_m}{j\omega(C_{EB} + C_{BC})}. \tag{3.19}$$

The transit frequency is the frequency at which small-signal current gain $|h_{21}|$ becomes unity, which can be expressed as

$$\frac{1}{2\pi f_T} = \frac{C_{EB} + C_{BC}}{g_m}. \tag{3.20}$$

The capacitance C_{EB} and C_{BC} can be split into both depletion (C_j) and diffusion capacitance (C_d) components. The sum of the diffusion capacitance is expressed in terms of the forward transit time (τ_f)

$$C_{dEB} + C_{dBC} = \tau_f g_m. \tag{3.21}$$

With this, the equation 3.20 can be re arranged as

$$\tau_{ec} = \frac{1}{2\pi f_T} = \tau_f + \frac{C_{jEB} + C_{jBC}}{g_m} \tag{3.22}$$

where C_{jEB} and C_{jBC} are the corresponding depletion capacitances. Now the forward transit time can be related to the regional transit times as discussed in the last section, with the following equation

$$\frac{1}{2\pi f_T} = \tau_{ec} = \tau_e + \tau_b + \tau_c + \tau_{bc,0} + \frac{C_{jEB} + C_{jBC}}{g_m} \tag{3.23}$$

where $\tau_{bc,0} \approx x_d/2v_{sat}$ is the current independent part of collector-base junction transit time, and x_d is the base-collection depletion width and v_{sat} is the saturation electron velocity.

The depletion time constants τ_{be} and τ_{bc} are due to the uncompensated charges at corresponding depletion regions; therefore, they can be expressed with their corresponding depletion capacitance as [145]

$$\tau_{eb} = \frac{C_{jEB}}{g_m}, \tag{3.24}$$

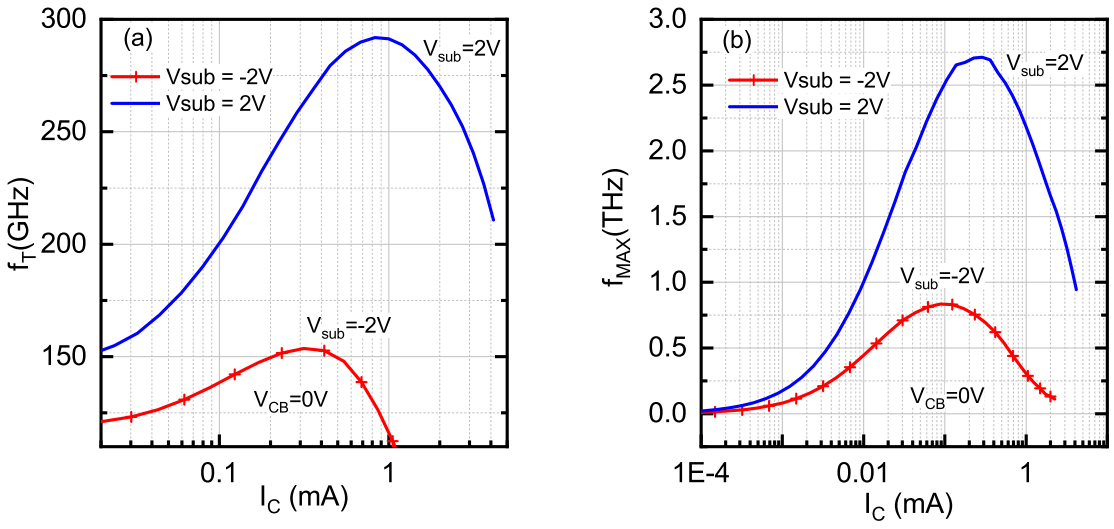
$$\tau_{bc} = \frac{C_{jBC}}{g_m} + \frac{x_d}{2v_{sat}}. \tag{3.25}$$

Now arranging all the terms, the expression derived from small signal equivalent circuit can be expressed as

$$\frac{1}{2\pi f_T} = \tau_{ec} = \tau_e + \frac{kT}{qI_C}C_{jBE} + \tau_b + \left(\frac{kT}{qI_C}C_{jBC} + \frac{x_d}{2v_{sat}}\right) + \tau_c. \tag{3.26}$$

Now for further analysis, the relation between regional transit time and the one derived from small-signal equivalent circuit can be compared term by term from this simplified expression.

3.4.4 RF Performance Analysis



(a) (b)

Figure 3.17: Collector current dependent (a) transit frequency and (b) maximum oscillation frequency at $V_{CB} = 0\ V$ and different V_{sub}.

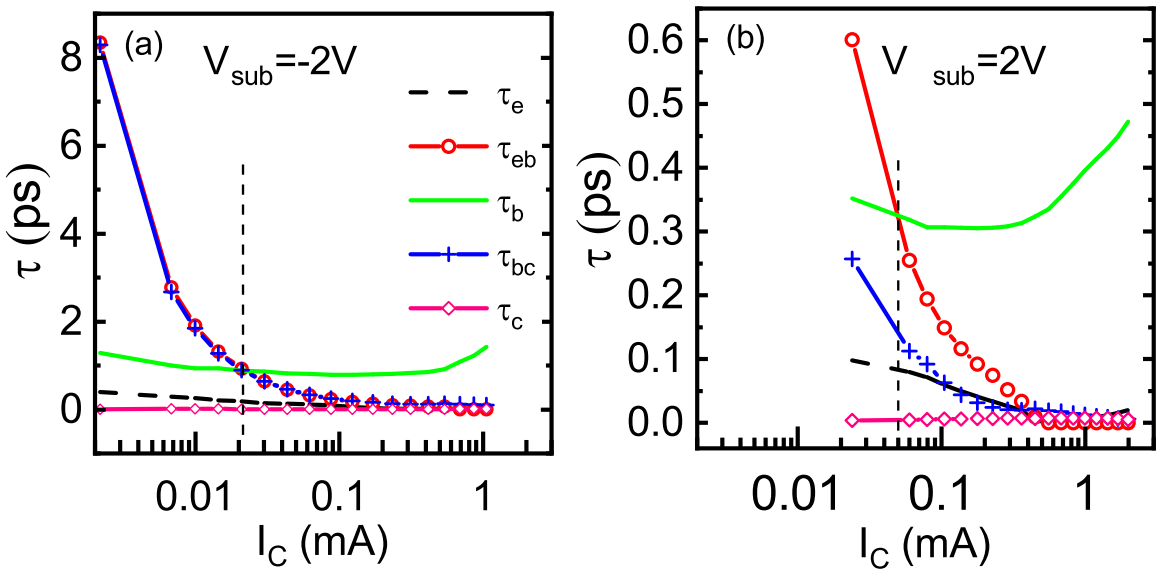


Figure 3.18: Regional transit time variation w.r.t substrate bias the transit time axis is in linear scale and the collector current axis is in log scale. The marked line indicates the change from low injection to intermediate operation. The analysis has carried out at $V_{CB} = 0\ V$.

The transit frequency (f_T) and the maximum oscillation frequency (f_{MAX}) are the main FoMs of a transistor from the perspective of its high-speed performances. Fig.3.17 shows the impact of V_{sub} onto the collector current (I_C) dependent f_T and f_{MAX} curves. It is observed that both f_T and f_{MAX} are significantly increased when V_{sub} changes from -2 V to 2 V. Also the peaks of f_T and f_{MAX} occur at higher I_C levels showing a change in the collector current corresponding to the onset of the Kirk effect that is directly related to the dE/dx in (3.2). Essentially, as $\Delta n(Vsub)$ increases with positive V_{sub} values, obtaining the null point of dE/dx requires higher electron current density j_n; hence the Kirk effect is pushed to higher I_C leading to the peaks of f_T and f_{MAX} occurring at higher I_C.

As already discussed, the relation between the f_T and emitter-to-collector transit time (τ_{ec}), can be expressed as

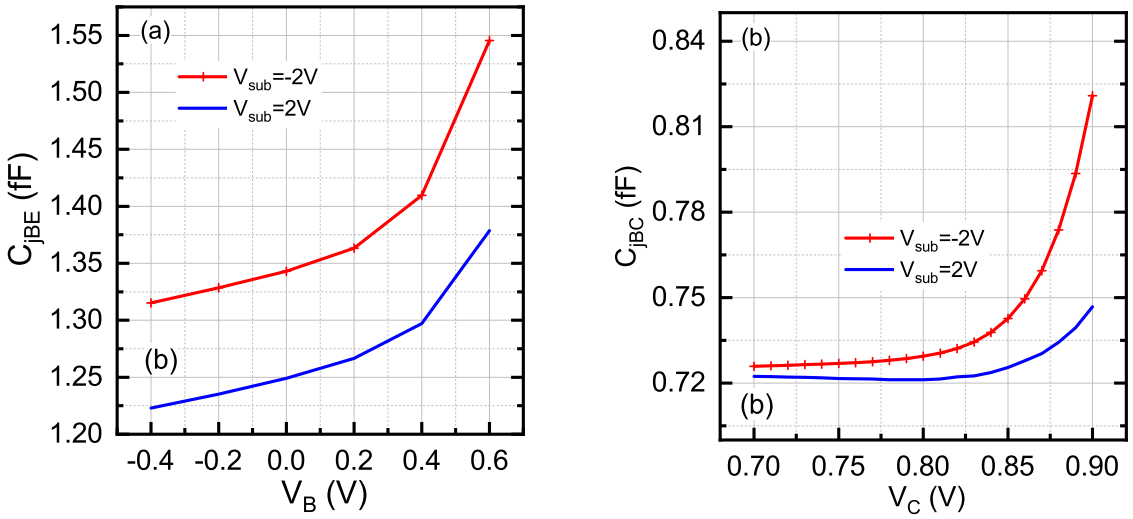
$$\frac{1}{2\pi f_T} = \tau_{ec} = \tau_e + \tau_{eb} + \tau_b + \tau_{bc} + \tau_c \tag{3.27}$$

where, the symbols hold the same meaning as discussed in the last sub-section. In order to investigate the impact of V_{sub} on various transit time components, we carried out a regional analysis at V_{CB}=0 V and different V_{sub}. We extracted the BE and BC metallurgical junctions (X_{eb} and X_{bc}) from Fig.3.14, as the intersection points of spatially $q\frac{dp}{dj_c}, q\frac{dn}{dj_c}$ curves (at the C-C′ cut-lines as mentioned in Fig.3.8) at each bias in the main operating region (0.7 V to 1 V). where p and n denote the hole and electron density, respectively, while j_c corresponds to the current density. Note that, the differential data are extracted following an infinitesimal change in the base-emitter voltage as reported in [81].

Fig.3.18 shows that in the low injection regime, the total transit time is dominated by the junction related transit time components (τ_{be} and τ_{cb}). Both τ_{eb} and τ_{bc} are found to decrease rapidly with increasing collector current which can be easily comprehended from (3.28) which was derived in the last section

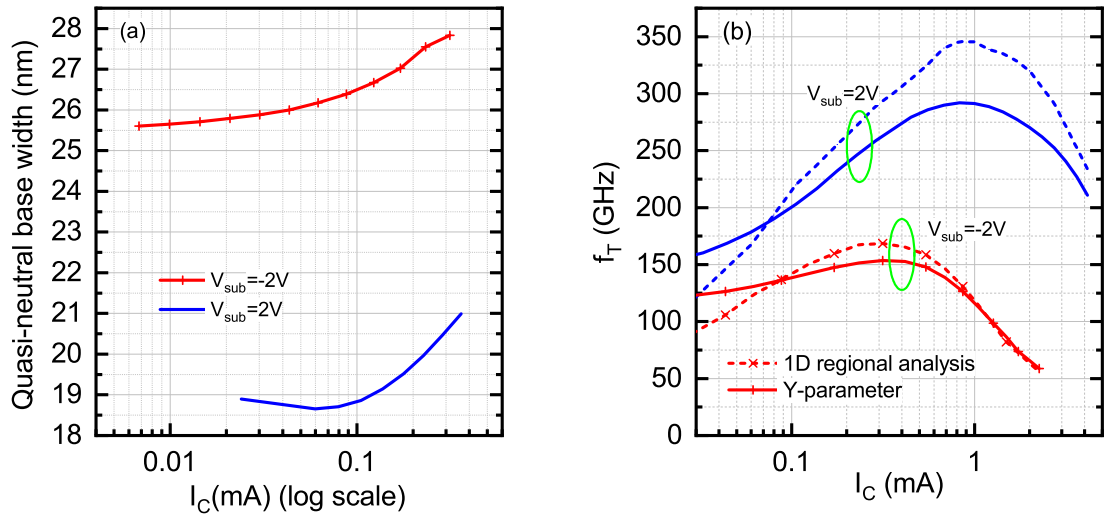
$$\frac{1}{2\pi f_T} = \tau_{ec} = \tau_e + \frac{kT}{qI_C}C_{jBE} + \tau_b + \left(\frac{kT}{qI_C}C_{jBC} + \frac{x_d}{2v_{sat}}\right) + \tau_c \tag{3.28}$$

where x_d, v_{sat} indicate the base-collector depletion width and the electron saturation velocity respectively, while C_{jBE} and C_{jBC} are junction capacitances. Note that the overall transit time values are significantly reduced as V_{sub} is changed from -2 V to 2 V.



(a) (b)

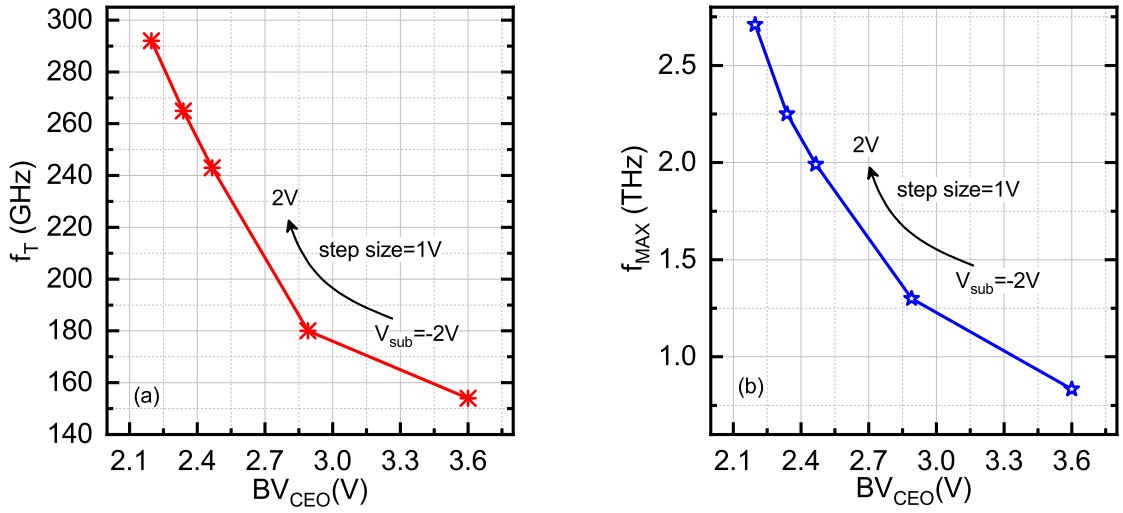
Figure 3.19: TCAD simulated (a) base-emitter and (b) base-collector capacitances at V_{sub}=-2 V and 2 V.



(a) (b)

Figure 3.20: TCAD simulated (a) effective quasi-neutral base width at different V_{sub}; (b) comparison of transit frequency extracted using 1D regional approach (dashed line) and that obtained from small-signal Y-parameters (as done in Fig.3.17).

Application of a positive V_{sub} reduces hole charge in the base (near the BOX/Si interface) while the carrier concentration within the highly doped emitter remains unchanged leading to an



(a) (b)

Figure 3.21: BV_{CEO} dependent f_T and f_{MAX} at different substrate bias; the extraction of f_T and f_{MAX} are carried out at V_{CB}=0 V.

enlargement of the BE-SCR within the base. This causes a reduction in C_{jBE} in (3.28) (see Fig.3.19(a)) consequently reducing the τ_{eb} in (3.27) (refer Fig.3.18(b)). On the BC side, positive V_{sub} causes relatively higher accumulation of electron density in the low-doped collector that results into a wider SCR within the base. Consequently, C_{jBC} and τ_{bc} reduce (see Figs.3.19(b) and 3.18(b)).

As shown in Fig.3.18, τ_b and τ_{bc} dominate at the medium injection regime. Finally, at the high injection level τ_b dominates and it counts almost 80% of the total transit time. At positive V_{sub}, the average neutral base width is reduced due to the encroachment of BE and BC SCRs within the base. Additionally, information of X_{jE} and X_{jC} obtained from the regional approach yields the exact width of QNB. Our estimate shows that the QNB is reduced from ∼26 nm to ∼19 nm as V_{sub} changes from -2 V to 2 V (Figs.3.20(a)). This leads to a significant increase in f_T as shown in Fig.3.17(a). Fig.3.20(b) compares the transit frequency extracted from 1D regional approach and that from the Y-parameters of the device. We observe a certain mismatch in the approximated 1D f_T values from the exact ones; however, the bias-dependent variation trends are in agreement. This confirms that the modulation of f_T is dominated by the intrinsic 1D device and that the contribution of parasitic capacitances is minimal. The substrate-related capacitances greatly depends on the thickness of the BOX oxide (T_{BOX}). For typical thin-film SOI devices, the BOX thickness is generally in order of nanometers. Therefore, the overall substrate-capacitances attributed will be much lower than any junction capacitances in the device [121].

Fig.3.17(b) demonstrates that an increase in V_{sub} from -2 V to 2 V causes an almost 225% (∼1.9 THz) increase in f_{MAX}. It is well known that f_{MAX} depends on f_T, base resistance (r_b) and collector-base junction capacitance (C_{jBC}) through the relation (3.29)

$$f_{MAX} = \sqrt{\frac{f_T}{8\pi r_b C_{jBC}}}. \tag{3.29}$$

Fig.3.19(b) shows that there is around 6% reduction in C_{jBC} as V_{sub} changes from -2 V to 2 V at

peak f_T bias point. Regarding the base resistance, we understand that it has a bias-independent extrinsic component (r_{bx}) that is estimated to be around 0.33 Ω (due to heavy doping) and a bias-dependent intrinsic component which requires special attention for the device under investigation. For the conventional device, typical emitter width and base thickness are about ∼90 nm and ∼20 nm, respectively. In case of our device, these dimensions are 7 nm and 28 nm, respectively. Hence, the dominant hole flow direction is horizontal (and not vertical in this lateral device) and accordingly, the r_{bi} formulation is modified as

$$r_{bi} = \frac{\rho_B L_{Beff}}{3 W_{si} T_{si}} \tag{3.30}$$

where ρ_B is the base resistivity [146], L_{Beff} is the QNB length in the lateral direction that varies from ∼19 nm to ∼26 nm as V_{sub} changes from 2 V to -2 V (Fig.3.20(a)), W_{Si} is width of the device in the Z-direction, and T_{si} is the silicon thickness of the base region. The intrinsic base resistances calculated using 3.30 are 4.16 Ω and 2.8 Ω for V_{sub}=-2 V and 2 V, respectively. This estimates a reduction of 40%. The 1/3 term in base resistance calculation is included to account the 3D distribution effect. It is clear that due to a change in V_{sub} from -2 V to 2 V, f_T increases whereas C_{jBC} and r_b reduce effectively improving the f_{MAX} following (3.29). For certain applications that demand transistors with balanced f_T and f_{MAX}, the silicon thickness (T_{si}) needs to be increased as reported in [110], [132].

Table 3.4: Summary of variations in figures-of-merit due to substrate bias variation at $V_{CB} = 0\ V$.

V_{sub}	I_C in μA (at peak f_T)	β	f_T (GHz)	f_{MAX} (THz)	BV_{CEO} (V)
-2V	87.81	440	154	0.834	3.6
2V	361.4	1035	292	2.71	2.2

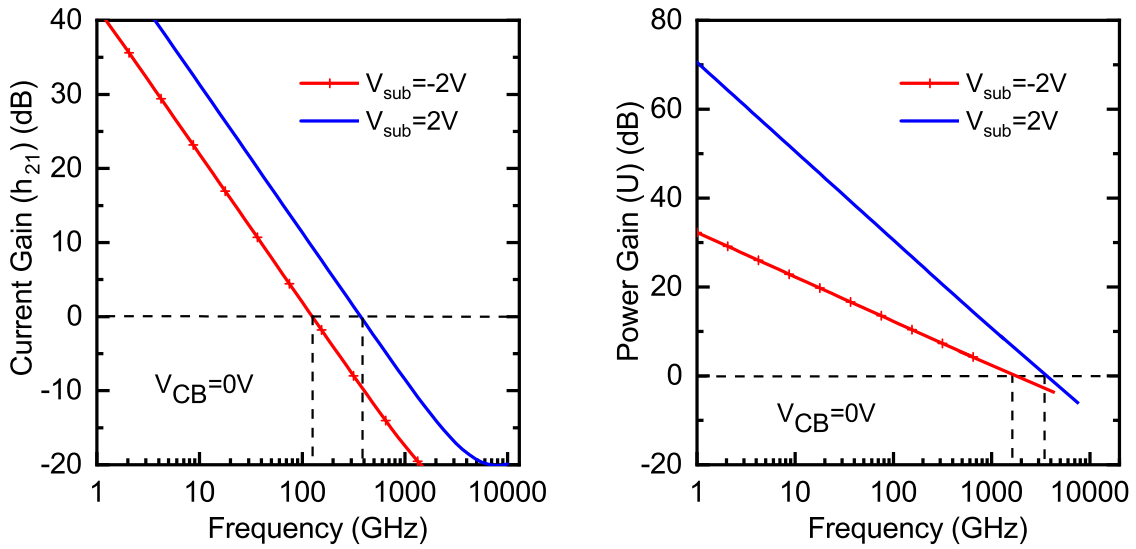


Figure 3.22: Extrapolation of (a) f_T from h_{21}, and (b) f_{MAX} from U at V_{sub}=-2 V, 2 V and $V_{CB} = 0\ V$. Here h_{21} and U are calculated from small-signal equivalent circuit corresponding to the LHBT.

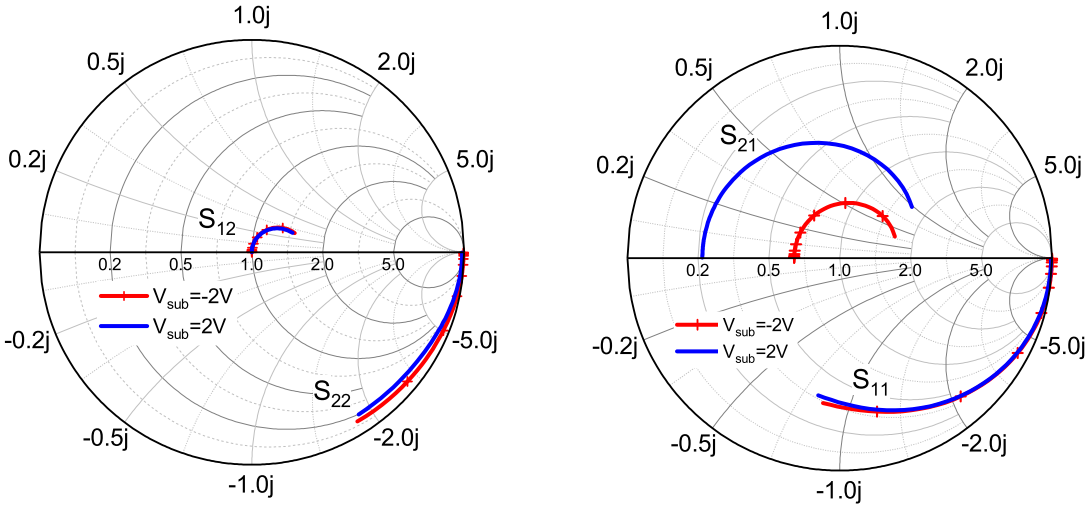


Figure 3.23: Effect of substrate bias on S_{12}, S_{22}, S_{11} and S_{21} parameter; the results are plotted at V_C=V_B=0.85V and from 1 GHz to 1 THz frequency range.

Table 3.5: Performance comparison of reported devices with this work. Note: lat. for lateral, vert. for vertical, asym. for asymmetric and sym. for symmetric.

Parameter	This work	[132]	[36]	[33]	[147]
Device	asym. lat.	sym. lat.	asym. vert.	asym. vert.	sym. lat.
Dimension	T_{si}=7nm	T_{si}=20nm	0.105×1 μm^2	55nm BiCMOS	28nm FDSOI
Bias	V_{CB}=0V, V_{sub}=2V	V_{CB}=1V	V_{CB}=0.25V	V_{CB}=0.5V	V_{DS}=1V
f_T (GHz)	292	200	505	325	246
f_{MAX} (THz)	2.7	1	0.72	0.375	0.359
B_{VCE0} (V)	2.2	–	1.6	1.5	–

Calculating the equivalent circuit parameters [148] of the device from the TCAD simulation, at both V_{sub}=-2 V and 2 V, the simplified equivalent small-signal model (same as in Fig.3.15) of the device is simulated in ICCAP. The resulting small signal current gain (h_{21}) and the power gain (U) w.r.t frequency are plotted in Fig.3.22. The f_T extracted by projecting the h_{21} intercept with the frequency axis at 0 dB gives a close match (120 GHz for -2 V, and 380 GHz for 2 V) with the results presented in Fig.3.17(a). Similarly, the f_{MAX} extracted from U gives the same trend for both the substrate biases projecting THz values. Although, the trend is same, there is a difference in exact values if compared with those of Fig.3.17. The reason can be cited to the use of approximate model, while ignoring the exact model that includes all parasitics.

Figs.3.21(a) and (b) present, respectively, the breakdown voltage-dependent variation of f_T and f_{MAX} obtainable by tuning the V_{sub}. Table 3.4 shows that with an optimal choice of device dimension and bias, the proposed device with substrate modulation technique can deliver the desired figures of merit. Following Fig.3.21, and Table 3.4, a device engineer can suitably

decide on the intended combination of breakdown voltage, f_T and f_{MAX} targeting a suitable application. The reduction of breakdown voltage with positive substrate bias is related to the equation described in (3.2) and well known Johnson's limit on bipolar transistor performance, where the product of breakdown voltage and transit frequency is roughly constant (approximately 340 GHZ.V in the case of silicon devices and higher than 500 GHz.V for SiGe-HBT [23]). This limit is material-related, and consequently is not subject to improvements in device design and process technology. The calculated Johnson's product values at both the substrate biases are 551 GHz.V and 642 GHz.V for -2 V and 2 V, respectively.

Scattering parameters on Smith chart is considered to be the most useful method to test the utility of a device for circuit design [149]. It is observed from Fig.3.23 that input and output reflection coefficients (S_{11} and S_{22}) are almost insensitive to the substrate bias. Therefore, it is easier to obtain a single matching network at the output for any substrate bias. However, the intended variation in gain is clearly visible from the forward gain (S_{21}) parameter.

In Table 3.5, we compare some of the already fabricated devices with the device under investigation. It shows that the proposed device has the potential to deliver higher RF performance and breakdown voltage with appropriate device engineering to suit intended applications.

3.5 Effect of Substrate Doping

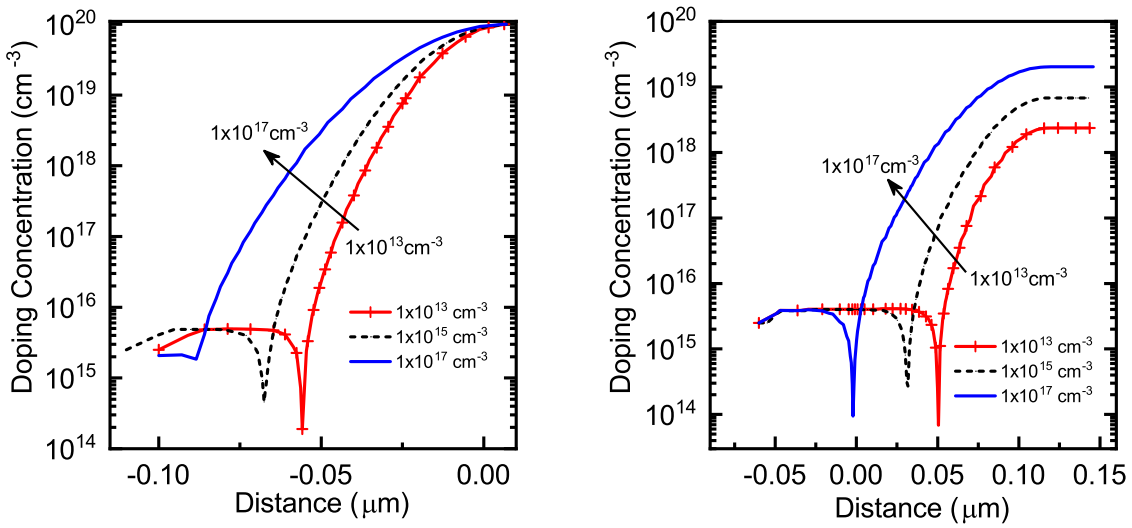


Figure 3.24: Doping profile at (a) vertical cut-section (DD') (b) horizontal cut-section (EE') as drawn on the schematic in Fig.3.8.

In the previous section, we observed that the substrate bias plays a vital role in terms of modulating the collector and base carrier densities which, in turn, changes the device performance. Similarly, this section investigates the impact of substrate doping variation. The variation in substrate doping can modulate the potential at the lower end of the BOX oxide (Fig.3.8) that effectively modulates the carrier densities in the base and collector regions leading to an impact on the device performance.

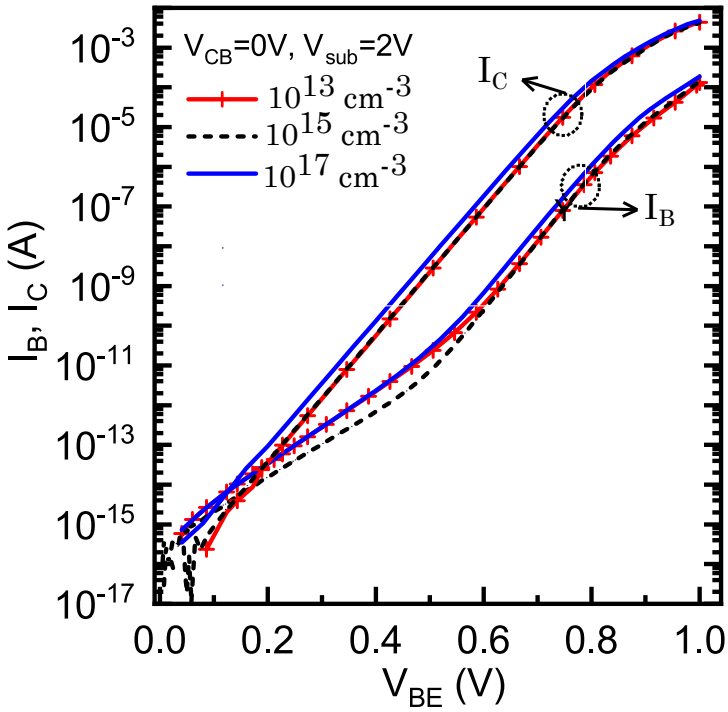
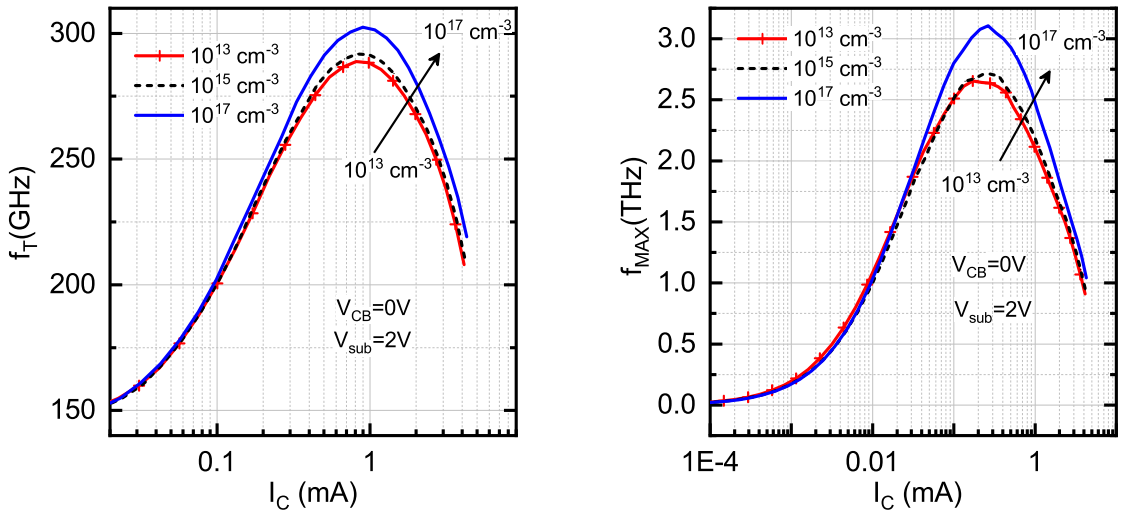


Figure 3.25: Effect of substrate doping on Gummel at V_{CB}=0 V, V_{sub}=2 V.



(a) (b)

Figure 3.26: f_T and f_{MAX} variation with substrate dopant at V_{CB}= 0V, V_{sub}=2 V.

The same device structure with a vertically and laterally varying substrate doping profile is presented for the analysis. In order to control the Gaussian profile, the doping value at a particular depth under the substrate contact is tuned, and resulting different doping profiles are generated. Figs.3.24(a) and (b) show, respectively, the vertical and horizontal doping profiles taken at the vertical and horizontal cut-lines on the device structure (as shown in Fig.3.8) after the above said doping level variation.

The doping value at the depth has been varied from 10^{13} to 10^{17} cm^{-3} with a step size of

10^2 cm^{-3}. It is observed from Figs.3.24(a) and (b) that with increased doping, the spreading of n-doping in substrate increases both in horizontal (i.e, towards the base and emitter w.r.t substrate contact) and vertical (i.e, towards the bottom of the device from the top substrate contact) directions accordingly causing a shift in the pn junctions.

The Gummel plot for the different substrate doping is shown in Fig.3.25 at V_{CB}=0 V and V_{sub}=2 V. As the doping varies from 10^{13} to 10^{17} cm^{-3}, the collector current increases. The f_T and f_{MAX} (see Fig.3.26) of the device also increase with higher substrate doping. Note that in section 4.5, we presented the results corresponding to a substrate doping of 10^{15} cm^{-3}. The resulting improvements in f_T and f_{MAX} indicates that substrate doping can be appropriately tuned for performance improvement.

3.6 Introduction to Process Simulation and its Importance in Device Development

The process development to realize the proposed SOI-LBJT is described in this section. The main tools of TCAD are process simulators ($S-process$ in Sentaurus TCAD) and device simulators ($S-device$ in Sentaurus TCAD). Process simulations predict the structure that results from applying a sequence of processing steps [150], and the device simulator predicts the electrical behaviors (such as f_T, f_{MAX}, and breakdown voltages) of the transistor. The design flowchart is illustrated in Fig.3.27.

The device development starts with a novel idea and some performance targets. A set of processing steps are used to build the initial structure in the process simulator, and the completed structure is passed over to the device simulator for performance verification. The device simulator analyzes and predicts the electrical property and behavior of the device, which can be fed back to the process simulator for device re-design or optimization. This routine is repeated until design specifications are achieved. Then, the process simulator is used to verify the device's sensitivity to process variations. This is to ensure that a working prototype device can still be made when the manufacturing process drifts from standard conditions. Once the design is finalized, the test chip layout is designed w.r.t the required geometry. The necessary recipe for constructing the device is developed. Short loop experiments are used to test and tune the process recipe to the desirable conditions before the fabrication starts.

Once the test chip is completed, the fabricated transistors and test structures are analyzed, and the electrical behaviors are measured. If the transistor characteristics do not satisfy the target specifications, then the re-design and optimization cycles repeat. First, the physical information of the device (i.e., doping and dimensions) are used to calibrate the settings of the process simulator such that the simulated device matches the experimental device. Then, the device simulation settings are calibrated to match the measured electrical behavior. As the calibrated process/device simulator can emulate the fabricated device, any process modification can be accurately predicted. This facilitates the re-design process in a very short time, without any experimental run. However, as there are limitations to any TCAD tools, several iterations of such simulation and fabrication cycle might be necessary to realize a working device that meets the target specification. Certainly, the development does not need to stop at this point. If better performance is desired, the cycle continues until the targets are achieved.

3.6.1 SOI Lateral Bipolar Junction Transistor Process Flow

This SOI lateral bipolar junction transistor is developed keeping an eye on standard STM's 28 nm FDSOI CMOS process. The process flow is arranged by using as much CMOS compatible

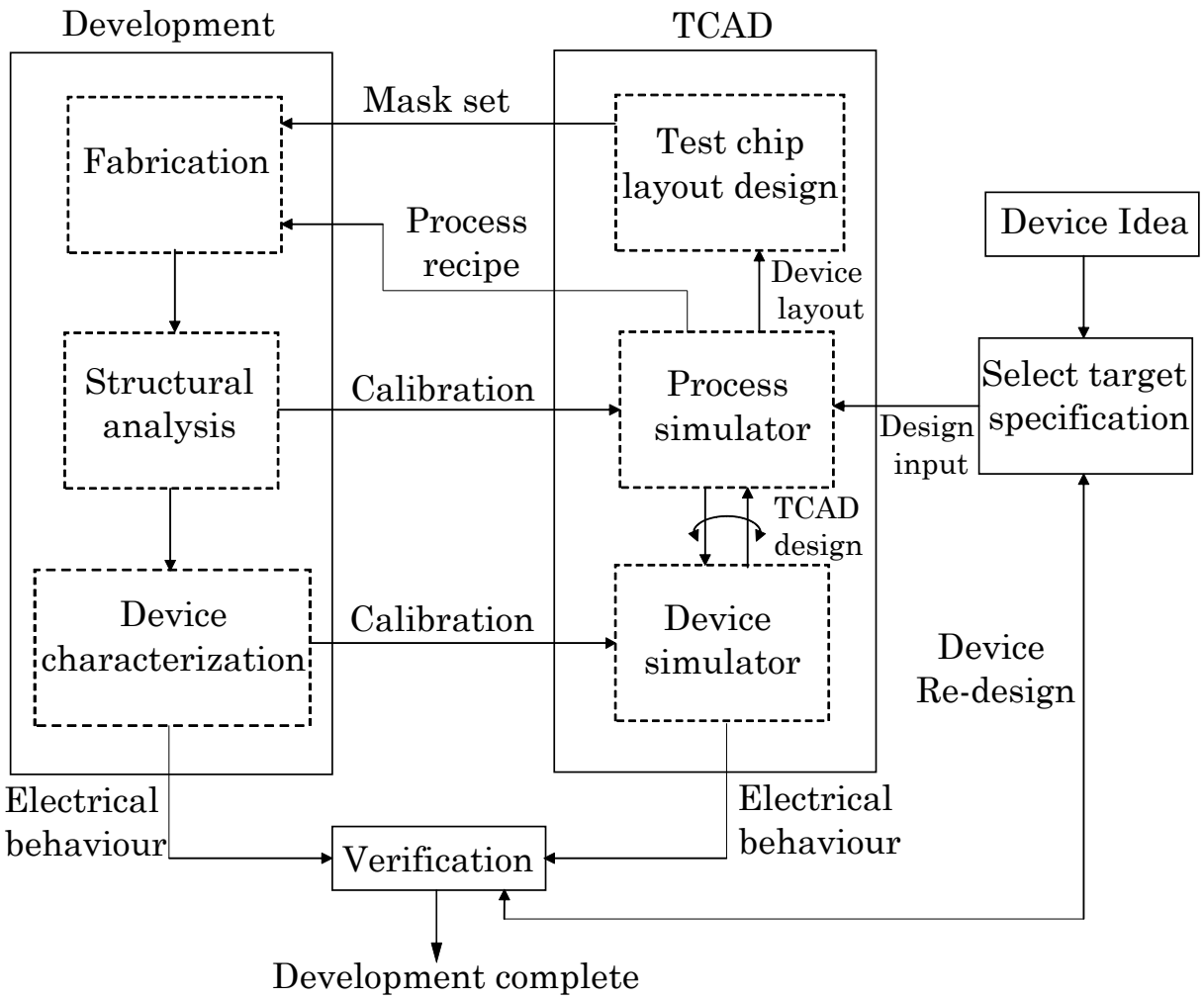


Figure 3.27: A high level flow chart for modern device development, with core deign done with TCAD.

procedure as possible in order to minimize the manufacturing complexity. The process flow is described in detail below:

Starting Substrate

The process starts with an SOI substrate, with SOI/BOX thickness of 20/25 nm, respectively (as shown in Fig3.28). The SOI film is p-type and is lightly doped to 10^{15} cm^{-3} and $< 100 >$ orientation. These particular specifications are chosen due to FDSOI 28 nm process [151], [152], [134]. This will allow a more compatible future integration with SOI-CMOS, and will provide a direct comparison to existing LBJT structures. In order to achieve the best possible performance, these dimensions can be further optimized.

Super Shallow Trench Isolation Formation

"Super-STI" , which is nothing but a very shallow STI. A mask (M1) is used to cover the wafer up to 180 nm (refer Fig3.28) and then both silicon and BOX oxide of the open region is etched down to remove the BOX oxide region mainly. Again silicon layer is deposited, and a chemical mechanical processing (CMP) step is done to create the space for super shallow trench isolation (SSTI) formation. The steps for SSTI are maintained the same as STI because of not having solid

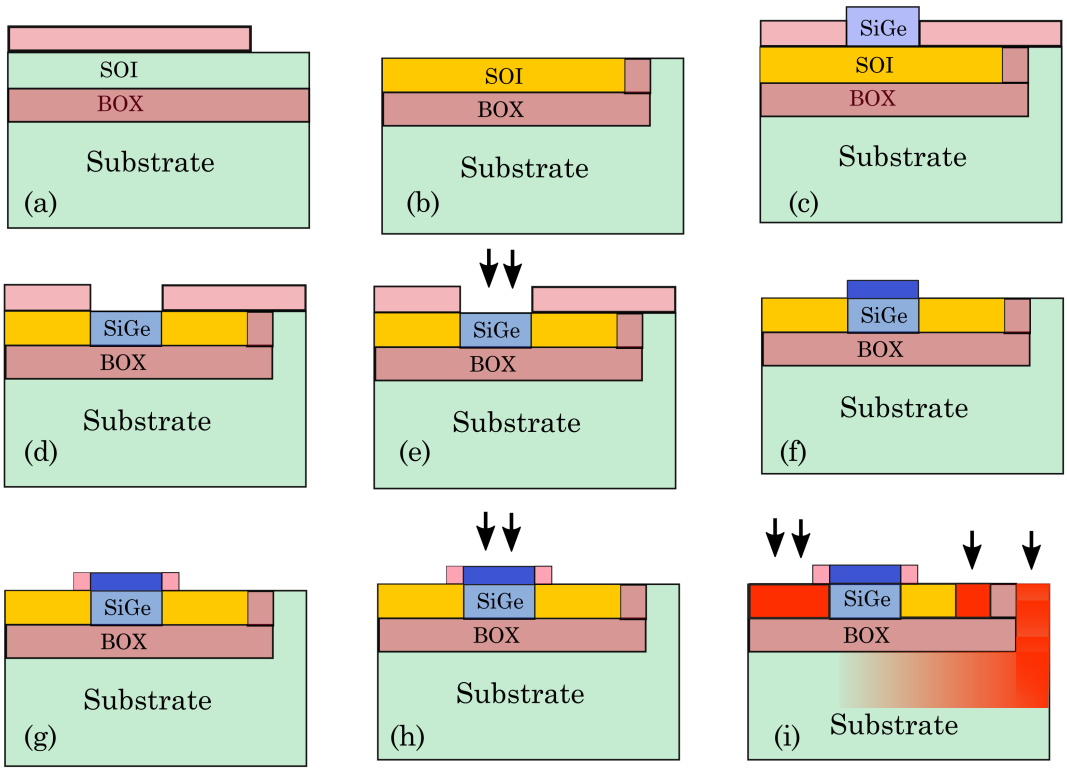


Figure 3.28: Complete process flow is as follows (a) SSTI formation (b) entire SOI doping (c) base pattern and SiGe deposition after oxidation (d) germanium condensation (e) base implant and rapid thermal annealing (RTA) (f) poly-base growth (g) spacer formation (h) base-poly doping (i) emitter, collector and substrate doping, and final device. To be noted: germanium mole fraction kept to be 20% which falls within the well suited range of high quality SiGe base formation.

information from any fab lab due to confidential reason. Initially, silicon is etched, and then a linear oxide is formed using thermal oxidation. The rest part of the trench is filled using oxide deposition. The purpose of choosing SSTI because of super thin SOI thickness of 7 nm.

After forming the SSTI trench, the entire wafer is implanted with arsenic to form the n-type lightly doped collector (n-LDC). The energy and angle should be such that arsenic can fully penetrate the SOI films and no p-type region remains at the SOI/BOX interface. Dose levels in the order of $10^{11}\ cm^{-2}$-$10^{12}\ cm^{-2}$ are used to produce a LDC concentration of roughly $10^{17}\ cm^{-3}$, in order to ensure sufficient trade-off of both breakdown and collector resistance. Also, such levels can push away the onset of the Kirk effect.

Base Region

As suggested in [77] while fabricating a lateral device within a BiCMOS framework, the main deviation from the CMOS process line occurs in the realization of the SiGe base. Fortunately for the lateral SiGe HBT, it does not require any additional mask saving the overall process cost. No epitaxial silicon is needed as it does not need a buried layer like the conventional vertical HBT.

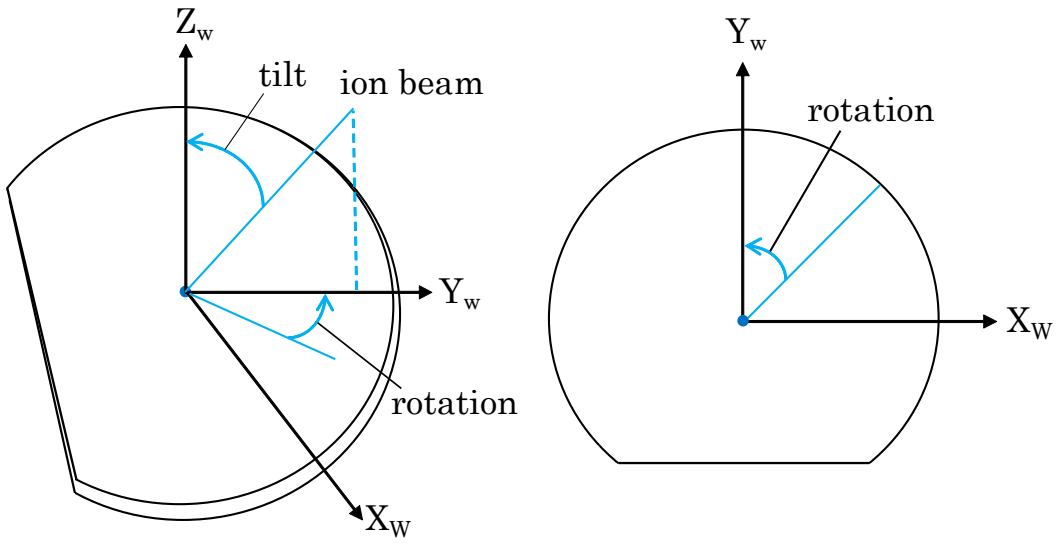
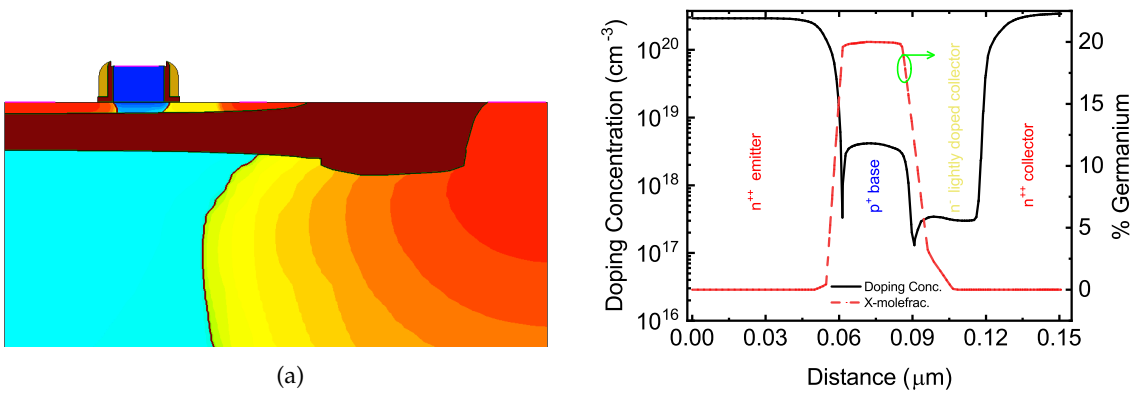


Figure 3.29: Difference between rotation and tilt are demonstrated. In this particular figure angles for implantation; beam angle shown corresponds to tilt = 20^0 and rotation = 45^0 [76].



(a)

(b)

Figure 3.30: (a) Final LBJT device with contacts (b) corresponding doping profile taken at the middle of silicon on insulator (T_{si}) of the device.

A mature material growth system to realize a practical lateral SiGe device has not yet been popularized. As discussed with engineers from ST Microelectronics it is possible to realize the SiGe base region, in fact [134] has demonstrated condensation to achieve high quality SiGe with a germanium mole fraction up to 30%. The base realization can start with pattering the base region and depositing with a low quality SiGe on top of the base by a chemical vapor deposition (CVD) method. This SiGe layer, when oxidized, will give an excellent crystalline SiGe and the residual SiO_2 can be removed in later steps. We kept this step simple during our process simulation as we have less information related to the exact procedure used at ST Microelectronics. SiGe base is formed through germanium condensation in the first step with a mask (M2), which will be reused during intrinsic and extrinsic base implantation.

Double implant of BF_2 and boron are used to produce a uniform doping vertical profile across the entire depth of SOI, with both dopant species covers different penetration depth. This is to ensure that the base region has a relatively constant lateral doping profile, and the device do not experience premature breakdown, or unnecessary wide base width. Unlike the default 7^0

Table 3.6: The complete process recipe for lateral SiGe device.

Region	Dopant	Dose and Tilt	Energy	Rotation
Substrate Implant	Phosphorus (31)	Total Dose=4×10^{15} cm^{-2}, Tilt=0^0	Energy=5 KeV Energy=10 KeV Energy=20 KeV	Rotation=0^0, 90^0, 180^0, 270^0
n-well	Arsenic (79)	Total Dose=4×10^{10} cm^{-2}, Tilt=7^0	Energy=0.1 KeV	Rotation=0^0, 90^0, 180^0, 270^0
		Total Dose=4×10^{10} cm^{-2}, Tilt=30^0	Energy=1 KeV	
		Total Dose=4×10^{10} cm^{-2}, Tilt=0^0	Energy=1.5 KeV	
		Total Dose=1.2×10^{11} cm^{-2}, Tilt=0^0	Energy=2 KeV	
Base	BF2 (49)	Total Dose=3.2×10^{12} cm^{-2}, Tilt=0^0	Energy=5 KeV	Rotation=0^0, 90^0, 180^0, 270^0
n-LDC	Arsenic (79)	Total Dose=4×10^{10} cm^{-2}, Tilt=7^0	Energy=0.1 KeV	Rotation=0^0, 90^0, 180^0, 270^0
		Total Dose=4×10^{10} cm^{-2}, Tilt=30^0	Energy=1 KeV	
		Total Dose=4×10^{10} cm^{-2}, Tilt=0^0	Energy=1.5 KeV	
		Total Dose=1.2×10^{11} cm^{-2}, Tilt=0^0	Energy=2 KeV	
n-emitter/collector	Arsenic (79)	Total Dose=4×10^{14} cm^{-2}, Tilt=0^0	Energy=0.1 KeV	Rotation=0^0, 90^0, 180^0, 270^0
		Total Dose=3.6×10^{14} cm^{-2}, Tilt=7^0	Energy=1 KeV	
Poly base	BF2 (49)	Total Dose=1.2×10^{16} cm^{-2}, Tilt=0^0	Energy=1 KeV	Rotation=0^0, 90^0, 180^0, 270^0

tilt for conventional implantation, this implant uses 0^0 tilt to minimize any shadowing effect [153]. This is because 7^0 tilt implant will offset the equal amount of dose intended for the base region, and produce inconsistent base widths. However, in the actual fabrication it is difficult to achieve perfect 0^0 implantation (refer Fig.3.29 for tilt and rotation difference). Small non-zero angle implantations can be used to fine-tune the process for better process yield [118]. Implantation has done at an energy of 5 KeV and total dose of $3.2\times10^{12}cm^{-2}$ for a target doping of $3 \times 10^{18}cm^{-3}$, and it is quick annealed for 1 sec at 1089^0 C.

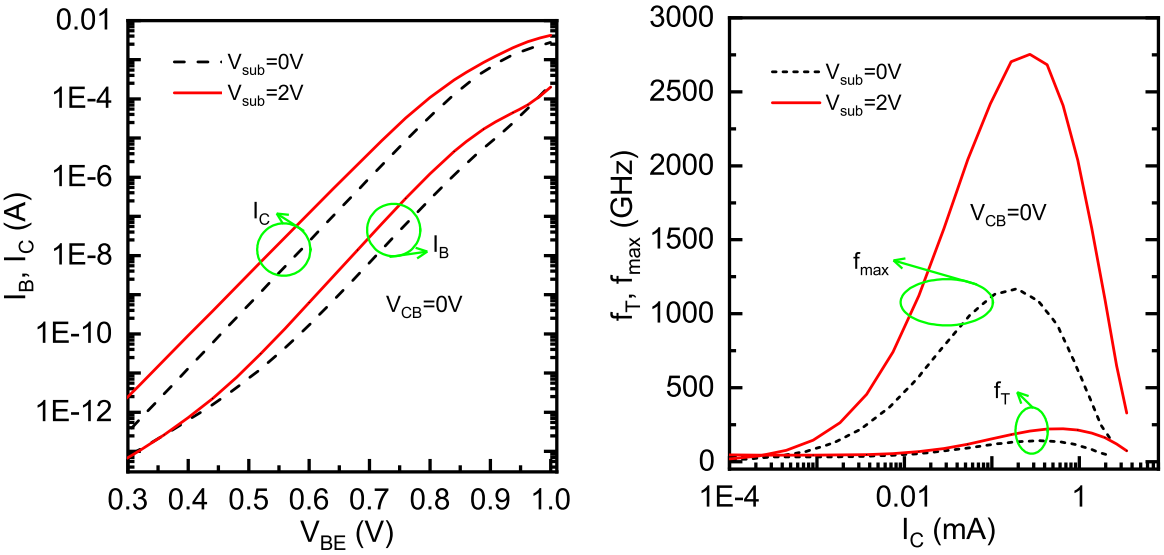
Poly-Silicon Base and Collector Emitter Contacts

Then poly-silicon deposition has been done with the same base mask (M2) to have a 20 nm extrinsic base region. Spacer has been formed around the anisotropically grown poly-silicon with 5 nm nitride and 3 nm oxide. Then a mask used to cover the base and n-LDC region and heavy does implantation of arsenic is done to achieve the emitter and collector regions (refer Table 3.6). Base poly-silicon is separately doped reusing mask (M1) to extract the base contact. This step is different from the CMOS development. In Fig.3.30 final device after all discussed steps along with doping profile is shown.

n-Substrate Doping

The substrate doping can be done in two ways, one while implanting for the heavily doped emitter and collector regions for contact, which will save one mask, but the depth of the substrate region doping will not be much because the dose and energy will be chosen most importantly based on the requirement for emitter and collector. If a specific substrate doping is planned, it can be done using a separate mask extending up to the right edge of SSTI. A mask up-to 0.218 nm (i.e right edge of SSTI) is used as shown in Fig3.28 to cover all other regions for a multi-step ion implantation of *phosphorous* to achieve desired n-substrate doping as presented in Table 3.6. Subsequent annealing is carried out for 10 seconds at 1022^0 C to activate the dopants.

3.6.2 Results



(a) (b)

Figure 3.31: (a) Gummel characteristics and (b) f_T, f_{MAX} vs. collector current I_C at substrate bias 0 V and 2 V, V_{CB}=0 V.

Figure Fig.3.31 (a) shows the Gummel characteristics at V_{CB}=0 V under varying substrate bias (V_{sub}). For V_{sub}=2 V the collector current (I_C) appears to be the highest with 80% improvement from I_C at V_{sub}=0 V. The transit and oscillation frequency improvement due to substrate bias is also presented in the same figure. We can observe a little variation in results presented in the earlier sections of this chapter where device simulation is considered. This is because the process development is not exactly optimized and calibrated to the same level. Following the process development steps presented in Fig.3.27 an optimized device can be developed, delivering results as required by the application.

3.7 Discussion and Conclusion

The proposed lateral SiGe HBT architecture suitably designed in advanced FDSOI nodes presents a number of advantages compared to the conventional vertical SiGe HBTs. The first advantage concerns the trade-off between C_{jBE} and r_{bi}. In a conventional vertical HBT, the base is highly doped to obtain a low r_{bi} and subsequently a high f_{MAX}. However, an increased base doping reduces the BE SCR leading to a higher C_{jBE}. Our new lateral HBT architecture is free from this trade-off. Since the proposed device has dramatically reduced dimensions in the direction perpendicular to the electron flow, r_{bi} component automatically becomes negligible allowing one to work with a relatively lower base doping that effectively reduces C_{jBE}. This results in a significantly high f_{MAX} reaching to THz level. The second advantage involves the nearly absent parasitic elements due to the significant reduction in the access of internal emitter, base, and collector regions. The peripheral junction capacitances and the buried layer for the collector access are suppressed. The third advantage is the simplified process for realizing an SiGe HBT since the proposed device can be suitably implemented in an advanced FDSOI technology.

Finally, the major advantage lies in the configurability of the device, thanks to the asymmetrical structure and a provision of back-bias from the substrate. Indeed, the results reported

in this work demonstrate that the substrate terminal can be judiciously used to adjust the device performance to the requirement of the circuit designer. For example, a positive voltage on the substrate terminal turns the device into a high-speed mode while a negative substrate bias modifies it into a high voltage transistor.

The major drawback of the device concerns its lower current level compared to a conventional HBT; however, a multi-finger architecture can partially address this problem. Also this drawback turns into an advantage in case of applications involving low power consumption and high-frequency performances. Therefore, this new architecture can generate plenty of opportunities for innovative circuit design on a SOI technology.

Conclusions and Scope for Future Research

Summary

In chapter 1, an extensive review of both architecture and performance of SiGe HBTs developed by various research labs and industries are discussed. In chapter 2, for the first time, a very high frequency (>110 GHz) measurement data is analyzed using two finite element tools, one solving the EM equations and the other one solving the semiconductor equations. The presented methodology provides one with sufficient confidence in the adopted characterization techniques and results. More precisely, it allows one to differentiate between accurate and erroneous characterizations. This work is a step forward (i) in adopting an improved characterization technique and (ii) in validating the TCAD result in high frequency. This will help the modeling of non-quasi static, and other high-frequency effects, which has been already demonstrated in [84] work. As a whole, the analysis outlines a prescription in order to make the researchers aware of certain precautions while designing the test structures, carrying out the process of characterization, and verifying the corresponding results. The detailed methodology, if adopted, will certainly help in the accurate extraction of high-frequency compact model parameters. Further, the need for proper calibration and de-embedding in high-frequency characterization is emphasized by investigating the S-parameters corresponding to a narrow-band amplifier at 170 GHz suitable for G-band radar applications.

While in chapter 3, using the similar vertical doping profile as STMicroelectronics' B55 technology ($f_{MAX} \sim$ 320 GHz), a nanowire SiGe HBT device is developed, and the issue of realizing the device and solution to overcome the problem is discussed. With the help of TCAD simulation, the parasitic effect on the nanowire array's RF performance is discussed. Devices placed with a small separation gap gives better RF performance, but it needs to be chosen carefully, looking into fabrication limitation. An f_{MAX} above 900 GHz (improvement of a factor 2.5) is projected, which may further be improved by optimizing the device dimension and buried layer of the device.

In chapter 4, we propose a new SiGe HBT based on FD SOI architecture. The device physics is compared with symmetric lateral SOI, and its advantage and usability are discussed. This asymmetric structure allows one to modulate the carrier densities in the collector region by applying substrate bias that causes significant improvements in the device performance. The transit frequency (f_T) is improved up to ∼90% by varying the substrate bias. The novelty of this device is that it achieves an f_{MAX} of 2.7 THz at V_{sub}=2V with a BV_{CEO}=2.2V and can be switched to an f_{MAX} of 0.8 THz with a BV_{CEO}=3.6 V at a V_{sub}=-2V. The bridge between the quasi-static analysis (to compare the small-signal performance) with the intrinsic (internal emitter to collector) device properties is discussed. Using regional analysis, the bias-dependent variation in-transit time is elaborately discussed. A formulation for base resistance calculation on thin SOI devices is proposed. Finally, an intended combination of breakdown voltage, f_T

and f_{MAX} is given, which will help a circuit designer in choosing proper substrate bias targeting a suitable application.

Scope for the Future Research

The well-calibrated B55 SiGe HBT device discussed in chapter 2 can be used to understand various physical effects (e.g., non-quasi static (NQS) effect) and analyze various breakdown mechanisms. The AC and Harmonic balance simulation can analyze noise analysis and large-signal device performance, respectively. Also, the RF performance can be predicted under cryogenic conditions and verified by measuring the actual device. Secondly, the cylindrical nanowire structure can help to develop an analytic model for SiGe HBT and base resistance calculation. Finally, a compact model for substrate modulated SOI-based asymmetric SiGe can be developed, and further, using the compact model, circuit performance can be evaluated. Once fully implemented as per STM's 28 nm FDSOI process, this new device will open the door for new THz applications, including the advantages of BiCMOS integration and actual THz operating frequencies.

Author Publication List

International journals

1. **Soumya Ranjan Panda**, Sebastien Fregonese, Marina Deng, Anjan Chakravorty, Thomas Zimmer, "TCAD and EM co-simulation method to verify SiGe HBT measurements up to 500 GHz" *Solid-State Electronics*, 174, 107915, (2020) **(Status: Published)**.
2. **Soumya Ranjan Panda**, Sebastien Fregonese, Anjan Chakravorty, Thomas Zimmer, "Novel SiGe based Nanowire Bipolar HBT for THz Frequency Applications" *xxxx*, xxx, xx, (2022) **(Status: To be submitted)**.
3. **Soumya Ranjan Panda**, Sebastien Fregonese, Pascal Chevalier, Anjan Chakravorty, Thomas Zimmer, "Substrate Bias Eeffect on Asymmetric Lateral SiGe HBT for THz Applications" *xxxx*, xxx, xx, (2022) **(Status: Submitted to TED)**.
4. Sebastien Fregonese, Marco Cabbia, Chandan Yadav, Marina Deng, **Soumya Ranjan Panda**, Magali De Matos, Anjan Chakravorty and Thomas Zimmer, "Analysis of high frequency measurement of transistors along with electromagnetic and SPICE co-simulation" *IEEE Transactions on Electron Devices*, 67, 4770-4776, (2020).
5. Bishwadeep Saha, Sebastien Fregonese, Anjan Chakravorty, **Soumya Ranjan Panda**, Thomas Zimmer, "Sub-THz and THz SiGe HBT Electrical Compact Modeling" *Electronics*, 10, 1397, (2021).

International conferences

1. **Soumya Ranjan Panda**, Sebastien Fregonese, Anjan Chakravorty, Thomas Zimmer, "TCAD simulation and assessment of anomalous deflection in measured S-parameters of SiGe HBTs in THz range", *2019 IEEE BiCMOS and Compound semiconductor Integrated Circuits and Technology Symposium (BCICTS)*, 1-4, (2019). **(Role-Speaker)**
2. **Soumya Ranjan Panda**, Marco Cabbia, Marina Deng, Sebastien Fregonese, Chandan Yadav, Anjan Chakravortyy, Thomas Zimmer, "SiGe HBT Device Characterization up-to 500 GHz: Procedure and Layout Improvement of Calibration Standards", *ICEE-IEEE*, xx-xx, (2020). **(Role-Speaker)**
3. Biswadeep Saha, Sebastien Fregonese, **Soumya Ranjan Panda**, Anjan Chakravorty, Didier Celi, Thomas Zimmer, "Collector-substrate modelling of SiGe HBTs upto THZ range", *019 IEEE BiCMOS and Compound semiconductor Integrated Circuits and Technology Symposium (BCICTS)*, 1-4, (2019). **(Role-Speaker)**

Workshops

1. **Soumya Ranjan Panda**, "TCAD and EM Co-Simulation Method to Verify B55 SiGe HBT Measurements up to 500 GHz", *STM-IMS workshop - 2020*. **(Role-Speaker)**
2. **Soumya Ranjan Panda**, Sebastien Fregonese, Chakravorty Anjan, Thomas Zimmer, "TCAD versus High Frequency Mesurements of SiGe HBTs", *AKB: BIPOLAR Working Group, 2019*. **(Role-Contributor)**
3. **Soumya Ranjan Panda**, Sébastien Fregonese, Marina Deng, Anjan Chakravorty, Thomas Zimmer, "Assessment of device RF performance and behavior using TCAD simulation", *IEEE BEE BRANCH*. **(Poster presentation)**
4. Sebastien Fregonese, Marina Deng, Marco Cabbia, Chandan Yadav, **Soumya Ranjan Panda**, Thomas Zimmer, "On wafer small signal characterization beyond 100 GHz for compact model assessment", *European Microwave Week Workshop Recent advances in SiGe BiCMOS: technologies, modelling & circuits for 5G, radar & imaging*. **(Role-Contributor)**

Bibliography

[1] N. Kukutsu and Y. Kado, "Overview of millimeter and terahertz wave application research," *NTT Technical Review*, vol. 7, no. 3, p. 6, 2009.

[2] M. Božanić and S. Sinha, "Millimeter-wave integrated circuits," 2020.

[3] DOTFIVE, *Eu project targets 0.5-thz sige bipolar transistor ee times europe print edition covering march 17 – april 6, 2008*, 2008-2012. [Online]. Available: https://www.eetimes.com/eu-project-targets-0-5-thz-sige-bipolar-transistor/.

[4] U. R. Pfeiffer and E. Öjefors, "Terahertz imaging with cmos/bicmos process technologies," in *2010 Proceedings of ESSCIRC*, IEEE, 2010, pp. 52–60.

[5] J. D. Cressler and G. Niu, *Silicon-germanium heterojunction bipolar transistors*. Artech house, 2003.

[6] T. Rosenbaum, *Performance prediction of a future SiGe HBT technology using a heterogeneous set of simulation tools and approaches*. BoD–Books on Demand, 2017.

[7] V. Trivedi, J. John, J Young, *et al.*, "A 90nm bicmos technology featuring 400ghz f max sige: C hbt," in *2016 IEEE Bipolar/BiCMOS Circuits and Technology Meeting (BCTM)*, IEEE, 2016, pp. 60–63.

[8] T. Zimmer, J. Böck, F. Buchali, *et al.*, "Sige hbts and bicmos technology for present and future millimeter-wave systems," *IEEE Journal of Microwaves*, vol. 1, no. 1, pp. 288–298, 2021.

[9] A. M. Arabhavi, F. Ciabattini, S. Hamzeloui, *et al.*, "Inp/gaassb double heterojunction bipolar transistor emitter-fin technology with fmax= 1.2 thz," *IEEE Transactions on Electron Devices*, 2022.

[10] J. Singh, J Ciavatti, K Sundaram, *et al.*, "14-nm finfet technology for analog and rf applications," *IEEE Transactions on electron devices*, vol. 65, no. 1, pp. 31–37, 2017.

[11] W Chakraborty, K. Aabrar, J Gomez, *et al.*, "Cryogenic rf cmos on 22nm fdsoi platform with record f t= 495ghz and f max= 497ghz," in *2021 Symposium on VLSI Technology*, IEEE, 2021, pp. 1–2.

[12] H. Kroemer, "Theory of a wide-gap emitter for transistors," *Proceedings of the IRE*, vol. 45, no. 11, pp. 1535–1537, 1957.

[13] B. G. Streetman, S. Banerjee, *et al.*, *Solid state electronic devices*. Pearson/Prentice Hall Upper Saddle River, 2006, vol. 10.

[14] J. C. Bardin, *Silicon-germanium heterojunction bipolar transistors for extremely low-noise applications*. California Institute of Technology, 2009.

[15] B. G. Malm, *High frequency characterization and modeling of sige heterojunction bipolar transistors*, 2002. [Online]. Available: http://www.diva-portal.org.

[16] W. Webster, "On the variation of junction-transistor current-amplification factor with emitter current," *Proceedings of the IRE*, vol. 42, no. 6, pp. 914–920, 1954.

[17] C. Kirk, "A theory of transistor cutoff frequency (f t) falloff at high current densities," *IRE Transactions on Electron Devices*, vol. 9, no. 2, pp. 164–174, 1962.

[18] Alam, *Bjt design*, 2018. [Online]. Available: https://nanohub.org/resources/5834/watch?resid=17046.

[19] R. F. Pierret, *Semiconductor device fundamentals*. Pearson Education India, 1996.

[20] D. J. Roulston, *Bipolar semiconductor devices*. McGraw-Hill College, 1990.
[21] P. Ashburn, *SiGe heterojunction bipolar transistors*. John Wiley & Sons, 2004.
[22] E. O. Johnson, "Physical limitations on frequency and power parameters of transistors," in *Semiconductor Devices: Pioneering Papers*, World Scientific, 1991, pp. 295–302.
[23] K. K. Ng, M. R. Frei, and C. A. King, "Reevaluation of the ftbv/sub ceo/limit on si bipolar transistors," *IEEE Transactions on Electron Devices*, vol. 45, no. 8, pp. 1854–1855, 1998.
[24] S. S. Iyer, G. Patton, S. Delage, S Tiwari, and J. Stork, "Silicon-germanium base heterojunction bipolar transistors by molecular beam epitaxy," in *1987 International Electron Devices Meeting*, IEEE, 1987, pp. 874–876.
[25] G. Patton, J. Comfort, B. Meyerson, *et al.*, *63-75 ghz ft*, 1990.
[26] D. Harame, K Schonenberg, M Gilbert, *et al.*, "A 200 mm sige-hbt technology for wireless and mixed-signal applications," in *Proceedings of 1994 IEEE International Electron Devices Meeting*, IEEE, 1994, pp. 437–440.
[27] B. Omer, Q. Liu, B Rainey, *et al.*, "A 0.13 pm bicmos technology featuring a 200/280 ghz (ft/fmax) sige hbt," 2003.
[28] M Khater, J. Rieh, T Adam, *et al.*, "Sige hbt technology with ft/fmax= 300/350 ghz and gate delay below 3.3 ps," IEDM, 2004.
[29] J. J. Pekarik, J Adkisson, P Gray, *et al.*, "A 90nm sige bicmos technology for mm-wave and high-performance analog applications," in *2014 IEEE Bipolar/BiCMOS Circuits and Technology Meeting (BCTM)*, IEEE, 2014, pp. 92–95.
[30] J.-S. Rieh, B Jagannathan, H Chen, *et al.*, "Sige hbts with cut-off frequency of 350 ghz," in *Digest. International Electron Devices Meeting,*, IEEE, 2002, pp. 771–774.
[31] B Heinemann, R Barth, D Bolze, *et al.*, "Sige hbt technology with ft/fmax of 300ghz/500ghz and 2.0 ps cml gate delay," in *2010 International Electron Devices Meeting*, IEEE, 2010, pp. 30–5.
[32] J Böck, K Aufinger, S Boguth, *et al.*, "Sige hbt and bicmos process integration optimization within the dotseven project," in *2015 IEEE Bipolar/BiCMOS Circuits and Technology Meeting-BCTM*, IEEE, 2015, pp. 121–124.
[33] P Chevalier, G Avenier, G Ribes, *et al.*, "A 55 nm triple gate oxide 9 metal layers sige bicmos technology featuring 320 ghz f t/370 ghz f max hbt and high-q millimeter-wave passives," in *2014 IEEE international electron devices meeting*, IEEE, 2014, pp. 3–9.
[34] A Fox, B Heinemann, R Barth, *et al.*, "Sige hbt module with 2.5 ps gate delay," in *2008 IEEE International Electron Devices Meeting*, IEEE, 2008, pp. 1–4.
[35] A. Fox, B. Heinemann, H. Rücker, *et al.*, "Advanced heterojunction bipolar transistor for half-thz sige bicmos technology," *IEEE electron device letters*, vol. 36, no. 7, pp. 642–644, 2015.
[36] B Heinemann, H Rücker, R Barth, *et al.*, "Sige hbt with fx/fmax of 505 ghz/720 ghz," in *2016 IEEE International Electron Devices Meeting (IEDM)*, IEEE, 2016, pp. 3–1.
[37] J. Korn, H. Rücker, and B. Heinemann, "Experimental verification of tcad simulation for high-performance sige hbts," in *2017 IEEE 17th Topical Meeting on Silicon Monolithic Integrated Circuits in RF Systems (SiRF)*, IEEE, 2017, pp. 94–96.
[38] P Chevalier, T. Meister, B Heinemann, *et al.*, "Towards thz sige hbts," in *2011 IEEE Bipolar/BiCMOS Circuits and Technology Meeting*, IEEE, 2011, pp. 57–65.
[39] A Fox, B Heinemann, and H Rücker, "Double-polysilicon sige hbt architecture with lateral base link," *Solid-state electronics*, vol. 60, no. 1, pp. 93–99, 2011.
[40] V. Vu, "Exploration and evaluation of a novel si/sige heterojunction bipolar transistor architecture for next bicmos generation," Ph.D. dissertation, Ph. D. thesis, Ph. D. dissertation, School Phys. Sci. Eng., Univ. Bordeaux . . ., 2016.

[41] B Geynet, P Chevalier, B Vandelle, *et al.*, "Sige hbts featuring ft > 400 ghz at room temperature," in *2008 IEEE Bipolar/BiCMOS Circuits and Technology Meeting*, IEEE, 2008, pp. 121–124.

[42] G. Avenier, T. Schwartzmann, P. Chevalier, *et al.*, "A self-aligned vertical hbt for thin soi sigec bicmos," in *Proceedings of the Bipolar/BiCMOS Circuits and Technology Meeting, 2005.*, IEEE, 2005, pp. 128–131.

[43] J Cai, A Ajmera, C Ouyang, *et al.*, "Fully-depleted-collector polysilicon-emitter sige-base vertical bipolar transistor on soi," in *2002 Symposium on VLSI Technology. Digest of Technical Papers (Cat. No. 01CH37303)*, IEEE, 2002, pp. 172–173.

[44] I. Mitrovic, O Buiu, S Hall, D. Bagnall, and P Ashburn, "Review of sige hbts on soi," *Solid-state electronics*, vol. 49, no. 9, pp. 1556–1567, 2005.

[45] T. Thibeault, E. Preisler, J. Zheng, *et al.*, "A study of ultra-high performance sige hbt devices on soi," in *2013 IEEE Bipolar/BiCMOS Circuits and Technology Meeting (BCTM)*, IEEE, 2013, pp. 235–238.

[46] H Rucker, B Heinemann, R Barth, *et al.*, "Integration of high-performance sige: C hbts with thin-film soi cmos," in *IEDM Technical Digest. IEEE International Electron Devices Meeting, 2004.*, IEEE, 2004, pp. 239–242.

[47] P. Chevalier, W. Liebl, H. Rücker, *et al.*, "Sige bicmos current status and future trends in europe," in *2018 IEEE BiCMOS and Compound Semiconductor Integrated Circuits and Technology Symposium (BCICTS)*, IEEE, 2018, pp. 64–71.

[48] IRDS, *International roadmap for device and systems: Irds roadmap*, 2020. [Online]. Available: https://irds.ieee.org/editions.

[49] S. R. Panda, S. Fregonese, M. Deng, A. Chakravorty, and T. Zimmer, "Tcad and em co-simulation method to verify sige hbt measurements up to 500 ghz," *Solid-State Electronics*, vol. 174, p. 107 915, 2020.

[50] S. Fregonese, M. Deng, M. Potereau, C. Ayela, K. Aufinger, T. Zimmer, *et al.*, "On-wafer characterization of silicon transistors up to 500 ghz and analysis of measurement discontinuities between the frequency bands," *IEEE Transactions on Microwave Theory and Techniques*, vol. 66, no. 7, pp. 3332–3341, 2018.

[51] B. Ardouin, M. Schroter, T. Zimmer, *et al.*, "Compact model validation strategies based on dedicated and benchmark circuit blocks for the mm-wave frequency range," in *2015 IEEE Compound Semiconductor Integrated Circuit Symposium (CSICS)*, IEEE, 2015, pp. 1–4.

[52] E. Lourandakis, *On-wafer Microwave Measurements and De-embedding*. Artech House, 2016.

[53] K. Yau, E. Dacquay, I. Sarkas, and S. P. Voinigescu, "Device and ic characterization above 100 ghz," *IEEE Microwave Magazine*, vol. 13, no. 1, pp. 30–54, 2012.

[54] W. Deal, X. Mei, K. M. Leong, V. Radisic, S Sarkozy, and R. Lai, "Thz monolithic integrated circuits using inp high electron mobility transistors," *IEEE Transactions on Terahertz Science and Technology*, vol. 1, no. 1, pp. 25–32, 2011.

[55] S. Fregonese, M. De Matos, M. Deng, D. Céli, N. Derrier, and T. Zimmer, "Importance of probe choice for extracting figures of merit of advanced mmw transistors," *IEEE Transactions on Electron Devices*, vol. 68, no. 12, pp. 6007–6014, 2021.

[56] D. F. Williams, P. Corson, J. Sharma, *et al.*, "Calibrations for millimeter-wave silicon transistor characterization," *IEEE transactions on Microwave Theory and Techniques*, vol. 62, no. 3, pp. 658–668, 2014.

[57] S. P. Voinigescu, E. Dacquay, V. Adinolfi, *et al.*, "Characterization and modeling of an sige hbt technology for transceiver applications in the 100–300-ghz range," *IEEE Transactions on Microwave Theory and Techniques*, vol. 60, no. 12, pp. 4024–4034, 2012.

[58] L. Galatro, A. Pawlak, M. Schroter, and M. Spirito, "Capacitively loaded inverted cpws for distributed trl-based de-embedding at (sub) mm-waves," *IEEE Transactions on Microwave Theory and Techniques*, vol. 65, no. 12, pp. 4914–4924, 2017.

[59] M. Deng, T. Quémerais, S. Bouvot, *et al.*, "Small-signal characterization and modelling of 55 nm sige bicmos hbt up to 325 ghz," *Solid-State Electronics*, vol. 129, pp. 150–156, 2017.

[60] M. Cabbia, "(sub)-millimeter wave on-wafer calibration and device characterization," Ph.D. dissertation, Université de Bordeaux, 2021.

[61] S. Fregonese, M. Deng, and e. a. De Matos Magali, "Comparison of on-wafer trl calibration to iss solt calibration with open-short de-embedding up to 500 ghz," *IEEE Transactions on Terahertz Science and Technology*, vol. 9, no. 1, pp. 89–97, 2019.

[62] L. Galatro and M. Spirito, "Millimeter-wave on-wafer trl calibration employing 3-d em simulation-based characteristic impedance extraction," *IEEE Transactions on Microwave Theory and Techniques*, vol. 65, no. 4, pp. 1315–1323, 2017.

[63] M. Seelmann-Eggebert, M. Ohlrogge, R. Weber, *et al.*, "On the accurate measurement and calibration of s-parameters for millimeter wavelengths and beyond," *IEEE Transactions on Microwave Theory and Techniques*, vol. 63, no. 7, pp. 2335–2342, 2015.

[64] C. Yadav, M. Deng, M. Cabbia, M. De Matos, B. Plano, T. Zimmer, *et al.*, "Importance and requirement of frequency band specific rf probes em models in sub-thz and thz measurements up to 500 ghz," *IEEE Transactions on Terahertz Science and Technology*, 2020.

[65] R. Sakamaki and M. Horibe, "Uncertainty analysis method including influence of probe alignment on on-wafer calibration process," *IEEE Transactions on Instrumentation and Measurement*, vol. 68, no. 6, pp. 1748–1755, 2019.

[66] D. F. Williams, F.-J. Schmückle, R. Doerner, G. N. Phung, U. Arz, and W. Heinrich, "Crosstalk corrections for coplanar-waveguide scattering-parameter calibrations," *IEEE transactions on microwave theory and techniques*, vol. 62, no. 8, pp. 1748–1761, 2014.

[67] G. N. Phung, F. J. Schmückle, R. Doerner, *et al.*, "Influence of microwave probes on calibrated on-wafer measurements," *IEEE Transactions on Microwave Theory and Techniques*, vol. 67, no. 5, pp. 1892–1900, 2019.

[68] S. Fregonese, M. Deng, M. Cabbia, C. Yadav, M. De Matos, and T. Zimmer, "Thz characterization and modeling of sige hbts: Review," *IEEE Journal of the Electron Devices Society*, vol. 8, pp. 1363–1372, 2020.

[69] M Cabbia, M. Deng, S Fregonese, M De Matos, D Céli, and T Zimmer, "In-situ calibration and de-embedding test structure design for sige hbt on-wafer characterization up to 500 ghz," in *2020 94th ARFTG Microwave Measurement Symposium (ARFTG)*, IEEE, 2020, pp. 1–4.

[70] S. R. Panda, S. Fregonese, A. Chakravorty, and T. Zimmer, "Tcad simulation and assessment of anomalous deflection in measured s-parameters of sige hbts in thz range," in *2019 IEEE BiCMOS and Compound semiconductor Integrated Circuits and Technology Symposium (BCICTS)*, IEEE, 2019, pp. 1–4.

[71] S. Fregonese, M. Cabbia, C. Yadav, *et al.*, "Analysis of high-frequency measurement of transistors along with electromagnetic and spice cosimulation," *IEEE Transactions on Electron Devices*, vol. 67, no. 11, pp. 4770–4776, 2020.

[72] T. Sentaurus, "Manuals, synopsys inc," *Mountain View, CA*, vol. 94043, 2009.

[73] F. M. Bufler, *Full-band Monte Carlo simulation of electrons and holes in strained Si and SiGe*. Herbert Utz Verlag, 1998.

[74] G. Wedel and M Schröter, "Hydrodynamic simulations for advanced sige hbts," in *2010 IEEE Bipolar/BiCMOS Circuits and Technology Meeting (BCTM)*, IEEE, 2010, pp. 237–244.

[75] G Sasso, "Transport models and advanced numerical simulation of silicon-germanium heterojunction bipolar transistors," *Universita degli Studi di Napoli" Federico II*, 2010.

[76] T. Sentaurus, "Version l-2016.," *Synopsys Inc., Mountain View, CA, USA*, 2016.

[77] S. Raman, P. Sharma, T. G. Neogi, M. R. LeRoy, R. Clarke, and J. F. McDonald, "On the performance of lateral sige heterojunction bipolar transistors with partially depleted base," *IEEE Transactions on Electron Devices*, vol. 62, no. 8, pp. 2377–2383, 2015.

[78] V. Van-Tuan, D. Celi, T. Zimmer, S. Fregonese, and P. Chevalier, "Tcad calibration of high-speed si/sige hbts in 55-nm bicmos," *ECS Transactions*, vol. 75, no. 8, pp. 113–119, 2016.

[79] T. Sentaurus, *Sdevice user guide, version g-2012.06*, 2012.

[80] A. Chakravorty *et al.*, *Compact hierarchical bipolar transistor modeling with HICUM*. World Scientific, 2010.

[81] J. Van Den Biesen, "A simple regional analysis of transit times in bipolar transistors," *Solid-state electronics*, vol. 29, no. 5, pp. 529–534, 1986.

[82] B. M. Haugerud, M. M. Pratapgarhwala, J. P. Comeau, *et al.*, "Proton and gamma radiation effects in a new first-generation sige hbt technology," *Solid-state electronics*, vol. 50, no. 2, pp. 181–190, 2006.

[83] Y. Taur and T. H. Ning, *Fundamentals of modern VLSI devices*. Cambridge university press, 2013.

[84] B. Saha, S. Fregonese, A. Chakravorty, S. R. Panda, and T. Zimmer, "Sub-thz and thz sige hbt electrical compact modeling," *Electronics*, vol. 10, no. 12, p. 1397, 2021.

[85] A Pawlak, M Schroter, J Krause, G Wedel, and C Jungemann, "On the feasibility of 500 ghz silicon-germanium hbts," in *2009 International Conference on Simulation of Semiconductor Processes and Devices*, IEEE, 2009, pp. 1–4.

[86] T Lacave, P Chevalier, Y Campidelli, *et al.*, "Vertical profile optimization for+ 400 ghz fmax si/sige: C hbts," in *2010 IEEE Bipolar/BiCMOS Circuits and Technology Meeting (BCTM)*, IEEE, 2010, pp. 49–52.

[87] M. Tonouchi, "Cutting-edge terahertz technology," *Nature photonics*, vol. 1, no. 2, pp. 97–105, 2007.

[88] D. L. Woolard, J. O. Jensen, and R. J. Hwu, *Terahertz science and technology for military and security applications*. world scientific, 2007, vol. 46.

[89] T. Ellermeyer, R. Schmid, A. Bielik, J. Rupeter, and M. Möller, "Da and ad converters in sige technology: Speed and resolution for ultra high data rate applications," in *36th European Conference and Exhibition on Optical Communication*, IEEE, 2010, pp. 1–6.

[90] B Hoeneisen and C. Mead, "Limitations in microelectronics—ii. bipolar technology," *Solid-State Electronics*, vol. 15, no. 8, pp. 891–897, 1972.

[91] S. P. Gaur, "Performance limitations of silicon bipolar transistors," *IEEE Journal of Solid-State Circuits*, vol. 14, no. 2, pp. 337–343, 1979.

[92] Y. Shi and G. Niu, "2-d analysis of device parasitics for 800/1000 ghz f/sub t//f/sub max/sige hbt," in *Proceedings of the Bipolar/BiCMOS Circuits and Technology Meeting, 2005.*, IEEE, 2005, pp. 252–255.

[93] M. Schroter, G. Wedel, B. Heinemann, *et al.*, "Physical and electrical performance limits of high-speed sigec hbts—part i: Vertical scaling," *IEEE Transactions on Electron Devices*, vol. 58, no. 11, pp. 3687–3696, 2011.

[94] Y. Guerfi and G. Larrieu, "Vertical silicon nanowire field effect transistors with nanoscale gate-all-around," *Nanoscale research letters*, vol. 11, no. 1, pp. 1–7, 2016.

[95] C. Thelander, P. Agarwal, S Brongersma, *et al.*, "Nanowire-based one-dimensional electronics," *Materials today*, vol. 9, no. 10, pp. 28–35, 2006.

[96] B. McCarthy and S Ponedal, "Ibm unveils world's first 2 nanometer chip technology, opening a new frontier for semiconductors," *IBM News Room.*, pp. 6–8, 2021.

[97] N Loubet, T Hook, P Montanini, *et al.*, "Stacked nanosheet gate-all-around transistor to enable scaling beyond finfet," in *2017 Symposium on VLSI Technology*, IEEE, 2017, T230–T231.

[98] G.-J. Zhang, A. Agarwal, K. D. Buddharaju, N. Singh, and Z. Gao, "Highly sensitive sensors for alkali metal ions based on complementary-metal-oxide-semiconductor-compatible silicon nanowires," *Applied physics letters*, vol. 90, no. 23, p. 233 903, 2007.

[99] Y. Xia, P. Yang, Y. Sun, *et al.*, "One-dimensional nanostructures: Synthesis, characterization, and applications," *Advanced materials*, vol. 15, no. 5, pp. 353–389, 2003.

[100] N. A. Melosh, A. Boukai, F. Diana, *et al.*, "Ultrahigh-density nanowire lattices and circuits," *Science*, vol. 300, no. 5616, pp. 112–115, 2003.

[101] D. Wang, B. A. Sheriff, and J. R. Heath, "Complementary symmetry silicon nanowire logic: Power-efficient inverters with gain," *Small*, vol. 2, no. 10, pp. 1153–1158, 2006.

[102] M. Law, J. Goldberger, and P. Yang, "Semiconductor nanowires and nanotubes," *Annu. Rev. Mater. Res.*, vol. 34, pp. 83–122, 2004.

[103] N. Singh, K. D. Buddharaju, S. K. Manhas, *et al.*, "Si, sige nanowire devices by top–down technology and their applications," *IEEE Transactions on Electron Devices*, vol. 55, no. 11, pp. 3107–3118, 2008.

[104] K. Tomioka, M. Yoshimura, and T. Fukui, "Sub 60 mv/decade switch using an inas nanowire–si heterojunction and turn-on voltage shift with a pulsed doping technique," *Nano letters*, vol. 13, no. 12, pp. 5822–5826, 2013.

[105] M. Egard, S. Johansson, A.-C. Johansson, *et al.*, "Vertical inas nanowire wrap gate transistors with ft> 7 ghz and f max> 20 ghz," *Nano letters*, vol. 10, no. 3, pp. 809–812, 2010.

[106] E. Lind and L.-E. Wernersson, "Design of rf properties for vertical nanowire mosfets," *IEEE transactions on nanotechnology*, vol. 10, no. 4, pp. 668–673, 2010.

[107] L. Liu, Q. Han, S. Makovejev, *et al.*, "Analog and rf analysis of gate all around silicon nanowire mosfets," in *2017 Joint International EUROSOI Workshop and International Conference on Ultimate Integration on Silicon (EUROSOI-ULIS)*, IEEE, 2017, pp. 176–179.

[108] F. Lindelöw, N. S. Garigapati, L. Södergren, M. Borg, and E. Lind, "Iii–v nanowire mosfets with novel self-limiting Λ-ridge spacers for rf applications," *Semiconductor Science and Technology*, vol. 35, no. 6, p. 065 015, 2020.

[109] O.-P. Kilpi, M. Hellenbrand, J. Svensson, E. Lind, and L.-E. Wernersson, "Vertical nanowire iii–v mosfets with improved high-frequency gain," *Electronics Letters*, vol. 56, no. 13, pp. 669–671, 2020.

[110] J. Hamel, Y. Tang, and K Osman, "Technological requirements for a lateral sige hbt technology including theoretical performance predictions relative to vertical sige hbts," *IEEE Transactions on Electron Devices*, vol. 49, no. 3, pp. 449–456, 2002.

[111] H. Rücker and B. Heinemann, "High-performance sige hbts for next generation bicmos technology," *Semiconductor Science and Technology*, vol. 33, no. 11, p. 114 003, 2018.

[112] M. J. Rodwell, M. Urteaga, T. Mathew, *et al.*, "Submicron scaling of hbts," *IEEE Transactions on Electron Devices*, vol. 48, no. 11, pp. 2606–2624, 2001.

[113] H Kroemer, "Two integral relations pertaining to the electron transport through a bipolar transistor with a nonuniform energy gap in the base region," *Solid-State Electronics*, vol. 28, no. 11, pp. 1101–1103, 1985.

[114] F. Stein, "Spice modeling of terahertz heterojunction bipolar transistors," Ph.D. dissertation, Université de Bordeaux, 2014.

[115] W. Kloosterman, J. Paasschens, and D. Klaassen, "Improved extraction of base and emitter resistance from small signal high frequency admittance measurements," in *Proceedings of the 1999 Bipolar/BiCMOS Circuits and Technology Meeting (Cat. No. 99CH37024)*, IEEE, 1999, pp. 93–96.

[116] M. Schröter and A. Chakravorty, *Compact hierarchical bipolar transistor modeling with HICUM*. World Scientific, 2010.

[117] H.-M. Rein, "Edge injection and pinch-in effect of transistors with cylindrical geometry," *Electronics Letters*, vol. 4, no. 25, pp. 553–554, 1968.

[118] I.-S. Sun, W. T. Ng, K. Kanekiyo, *et al.*, "Lateral high-speed bipolar transistors on soi for rf soc applications," *IEEE transactions on Electron Devices*, vol. 52, no. 7, pp. 1376–1383, 2005.

[119] J. Barth, D. Plass, E. Nelson, *et al.*, "A 45 nm soi embedded dram macro for the power™ processor 32 mbyte on-chip l3 cache," *IEEE Journal of Solid-State Circuits*, vol. 46, no. 1, pp. 64–75, 2010.

[120] STM, *Applications*, 2021. [Online]. Available: https://www.st.com/en/applications/.

[121] I.-S. Sun, W. T. Ng, H. Mochizuki, *et al.*, "A rf lateral bjt on soi for realization of rf soi-bicmos technology," in *Digest of Papers. 2006 Topical Meeting on Silicon Monolithic Integrated Circuits in RF Systems*, IEEE, 2006, 4–pp.

[122] Q. C. Ouyang, J. Cai, T. Ning, P. Oldiges, and J. B. Johnson, "A simulation study on thin soi bipolar transistors with fully or partially depleted collector," in *Proceedings of the Bipolar/BiCMOS Circuits and Technology Meeting*, IEEE, 2002, pp. 28–31.

[123] S. Fregonese, G. Avenier, C. Maneux, A Chantre, and T. Zimmer, "Thin film soi hbt: A study of the effect of substrate bias on the electrical characteristics," *Solid-state electronics*, vol. 50, no. 11-12, pp. 1673–1676, 2006.

[124] J. C. Sturm, J. P. McVittie, J. F. Gibbons, and L Pfeiffer, "A lateral silicon-on-insulator bipolar transistor with a self-aligned base contact," *IEEE electron device letters*, vol. 8, no. 3, pp. 104–106, 1987.

[125] G. Shahidi, D. Tang, B Davari, *et al.*, "A novel high-performance lateral bipolar on soi," in *International Electron Devices Meeting 1991 [Technical Digest]*, IEEE, 1991, pp. 663–666.

[126] M Rodder and D. Antoniadis, "Silicon-on-insulator bipolar transistors," *IEEE electron device letters*, vol. 4, no. 6, pp. 193–195, 1983.

[127] R Dekker, W. van der Einden, and H. Maas, "An ultra low power lateral bipolar polysilicon emitter technology on soi," in *Proceedings of IEEE International Electron Devices Meeting*, IEEE, 1993, pp. 75–78.

[128] T Shino, K Inoh, T Yamada, *et al.*, "A 31 ghz f/sub max/lateral bjt on soi using self-aligned external base formation technology," in *International Electron Devices Meeting 1998. Technical Digest (Cat. No. 98CH36217)*, IEEE, 1998, pp. 953–956.

[129] H. Nii, T. Yamada, K. Inoh, *et al.*, "A novel lateral bipolar transistor with 67 ghz f/sub max/on thin-film soi for rf analog applications," *IEEE Transactions on electron devices*, vol. 47, no. 7, pp. 1536–1541, 2000.

[130] T. Suligoj, M. Koricic, P. Biljanovic, and K. L. Wang, "Fabrication of horizontal current bipolar transistor (hcbt)," *IEEE Transactions on Electron Devices*, vol. 50, no. 7, pp. 1645–1651, 2003.

[131] M. J. Kumar and D Venkateshrao, "A new lateral sige-base pnm schottky collector bipolar transistor on soi for non-saturating vlsi logic design," in *16th International Conference on VLSI Design, 2003. Proceedings.*, IEEE, 2003, pp. 489–492.

[132] T. H. Ning and J. Cai, "On the performance and scaling of symmetric lateral bipolar transistors on soi," *IEEE Journal of the Electron Devices Society*, vol. 1, no. 1, pp. 21–27, 2013.

[133] A. Derrickson, A. H. Peterson, K. English, A. Haslam, S. Nath, and J. F. McDonald, "Assessment of thz performance for a lateral sige hbt on soi with a laterally graded base," *IEEE Transactions on Electron Devices*, vol. 65, no. 11, pp. 4747–4754, 2018.

[134] J. Cai, T. H. Ning, C. D'Emic, *et al.*, "Complementary thin-base symmetric lateral bipolar transistors on soi," in *2011 International Electron Devices Meeting*, IEEE, 2011, pp. 16–3.

[135] T. H. Ning and J. Cai, "A perspective on symmetric lateral bipolar transistors on soi as a complementary bipolar logic technology," *IEEE Journal of the Electron Devices Society*, vol. 3, no. 1, pp. 24–36, 2014.

[136] O. E. O. K. H. R. T. M. K. M. H. T. Washio K Kondo M and H. T, "A0.2- lm self-aligned sige hbt featuring 107-ghz fmax and 6.7-ps ecl," in *IEEE IEDM Tech Dig*, IEEE, 1999, 557–60.

[137] J.-B. Yau, J. Cai, and T. H. Ning, "Substrate-voltage modulation of currents in symmetric soi lateral bipolar transistors," *IEEE Transactions on Electron Devices*, vol. 63, no. 5, pp. 1835–1839, 2016.

[138] T. Scaling, "Using temperature to explore the scaling limits of sige hbts," *Extreme Environment Electronics*, p. 211, 2012.

[139] J. W. Slotboom, "Analysis of bipolar transistors," *Ph. D. Thesis*, 1977.

[140] korn, *Device simulation of high-performance sige heterojunction bipolar transisors*, 2018. [Online]. Available: https://depositonce.tu-berlin.de/.

[141] M Schroter and H Tran, "Charge-storage related parameter calculation for si and sige bipolar transistors from device simulation," in *Proceedings of the NSTI Nanotechnology Conference and Trade Show*, 2006, pp. 735–740.

[142] P Agarwal, G. Hurkx, J. Donkers, and J. Slotboom, "On the delay times in vertically scaled sige hbts," in *Proceedings of the Bipolar/BiCMOS Circuits and Technology Meeting, 2005.*, IEEE, 2005, pp. 264–267.

[143] P. Palestri, M. Mastrapasqua, A. Pacelli, and C. A. King, "A drift-diffusion/monte carlo simulation methodology for si/sub 1-x/ge/sub x/hbt design," *IEEE Transactions on Electron Devices*, vol. 49, no. 7, pp. 1242–1249, 2002.

[144] Y. Shi and G. Niu, "Vertical profile design and transit time analysis of nano-scale sige hbts for terahertz f/sub t," in *Bipolar/BiCMOS Circuits and Technology, 2004. Proceedings of the 2004 Meeting*, IEEE, 2004, pp. 213–216.

[145] C. G. Van de Walle and R. M. Martin, "Theoretical study of band offsets at semiconductor interfaces," *Physical Review B*, vol. 35, no. 15, p. 8154, 1987.

[146] D. A. Neamen, "An introduction to semiconductor devices," 2006.

[147] A. Cathelin, "Fully depleted silicon on insulator devices cmos: The 28-nm node is the perfect technology for analog, rf, mmw, and mixed-signal system-on-chip integration," *IEEE Solid-State Circuits Magazine*, vol. 9, no. 4, pp. 18–26, 2017.

[148] R. C. Jaeger and T. N. Blalock, *Microelectronic circuit design*. McGraw-Hill New York, 1997.

[149] G. Gonzalez, *Microwave transistor amplifiers analysis and design*. Prentice-Hall, Inc., 1996.

[150] M. Golio, *RF and microwave semiconductor device handbook*. CRC press, 2017.

[151] T. C. De Albuquerque, F. Calmon, R. Clerc, *et al.*, "Integration of spad in 28nm fdsoi cmos technology," in *2018 48th European Solid-State Device Research Conference (ESSDERC)*, IEEE, 2018, pp. 82–85.

[152] V. Vu, D Celi, T Zimmer, S Fregonese, and P Chevalier, "Advanced si/sige hbt architecture for 28-nm fd-soi bicmos," in *2016 IEEE Bipolar/BiCMOS Circuits and Technology Meeting (BCTM)*, IEEE, 2016, pp. 64–67.

[153] G. Krieger, G. Spadini, P. P. Cuevas, and J. Schuur, "Shadowing effects due to tilted arsenic source/drain implant," *IEEE transactions on electron devices*, vol. 36, no. 11, pp. 2458–2461, 1989.