A Scalable Mextram Model for Advanced Bipolar Circuit Design

Hsien-Chang Wu

A Scalable Mextram Model for Advanced Bipolar Circuit Design

PROEFSCHRIFT

ter verkrijging van de graad van doctor
aan de Technische Universiteit Delft,
op gezag van de Rector Magnificus Prof. dr. ir. J. T. Fokkema,
voorzitter van het College voor Promoties,
in het openbaar te verdedigen

op maandag 24 september 2007 om 12:30 uur

door

Hsien-Chang WU

Master of Science, National Taiwan University

Dit proefschrift is goedgekeurd door de promotor:
Prof. Dr.-Ing. J. N. Burghartz

Samenstelling promotiecommissie:

Rector Magnificus, voorzitter
Prof. Dr. J. N. Burghartz, Technische Universiteit Delft, promotor
Prof. Dr. L. K. Nanver, Technische Universiteit Delft
Prof. Dr. J. W. Slotboom, Technische Universiteit Delft
Prof. Dr. P. J. French, Technische Universiteit Delft
Prof. Dr. T. Zimmer, University of Bordeaux I (France)
Dr. H. C. de Graaff, NXP Semiconductors
Dr. S. Mijalkovic, Silvaco Technology Center (United Kingdom)

The work in this thesis was supported by NXP Semiconductors.

Hsien-Chang Wu,
A Scalable Mextram Model for Advanced Bipolar Circuit Design,
Ph.D. Thesis Delft University of Technology,
with summary in Dutch.

Keywords: Mextram model, scalable model, SiGe HBT, bipolar device, parameter extraction.

ISBN: 978-90-8559-146-7

Printed by Optima Grafische Communicatie, Rotterdam, The Netherlands.

Contents

Chapter 1

Introduction

1.1 Evolution of bipolar transistors

The first transistor was the point contact transistor having two metal point contacts on a germanium substrate. It was invented in 1947 by W. Brattain, J. Bardeen and W. Shockley. Since then, bipolar junction transistors (BJTs) have been evolving rapidly. Nowadays, high-performance bipolar transistors are not only used in high-speed circuits for optical communication systems [1] but also to enhance the high-frequency performance of BiCMOS circuits building blocks [2] used in wireless communication systems.

Fig. 1.1 shows the cut-off frequency (f_T), the stages of technology development and the applications of bipolar transistors over time. Prior to 1970, the bipolar transistor was based on a diffused p-n junction structure, which was large in both lateral dimension and vertical doping profile. Due to large parasitic capacitance and high forward transit time (τ_F) for the diffused transistor, its maximum f_T in this period of time was just around 1 GHz. The main applications in this period were analog amplifiers to replace the power hungry vacuum tubes and switches in telephone relays [4,5]. Later in the 70's, a bipolar transistor was formed by using ion-implantation, which has a steep vertical doping profile to reduce its τ_F. As a result, its maximum f_T was increased to 5 GHz. Though f_T of the conventional bipolar transistor was still limited by its parasitic capacitance resulting from p^+ junction isolation and extrinsic base-collector junction as shown in Fig. 1.2, it was believed that the bipolar technology was mature at that time. However, in the 80's, advanced bipolar transistors using a double polysilicon (poly-emitter and poly-base con-

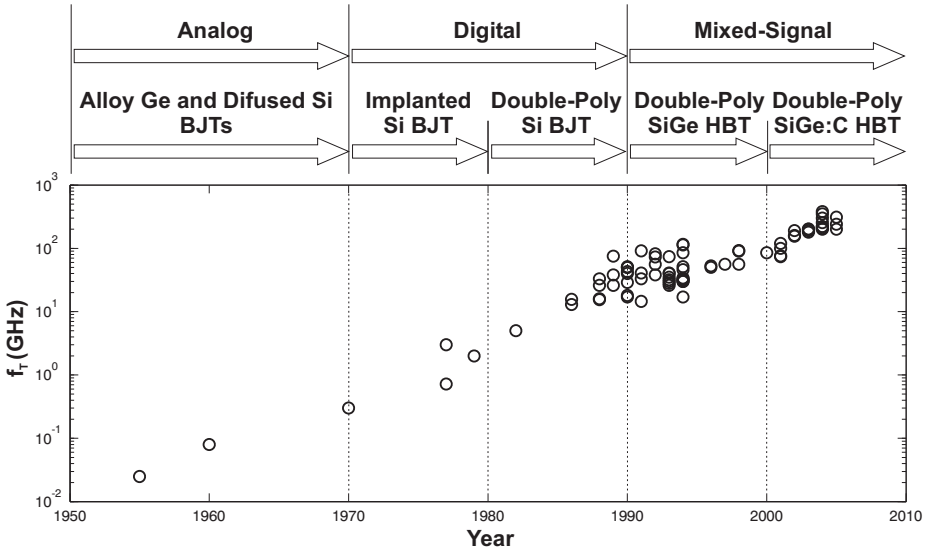


Figure 1.1: Cut-off frequency, the evolution of bipolar devices and applications vs. the year reported [3].

tacts) structure [6,7] were built to prevent from current-gain degradation from the self-aligned shallow emitter and to reduce the parasitic base-collector capacitance. Additional features such as an n^+ local collector implant and deep trench isolation were also used to enhance quasi-saturation [8] immunity and to reduce collector-substrate capacitance. With the new double-poly self-aligned device structure, the f_T of the bipolar transistor was increased to 50 GHz and an ECL gate delay of 73 ps [9] had been reported. The bipolar transistors in the 70s and 80s were mainly used for high-speed digital applications, such as emitter-coupled logic (ECL) for main-framed computer and SRAM, transistor-transistor logic (TTL) for microprocessors.

In the late 80's, SiGe-base heterojunction bipolar transistors (HBTs) [10] were invented to further improve bipolar transistor performance. Using the "MBE" or "UHV/CVD" [11] epitaxy technique, an in-situ-doped SiGe-base becomes steeper than a boron ion-implant base. In addition to the narrower SiGe-base, the Ge-induced bandgap narrowing in the graded SiGe-base region leads to a reducing base transit time and a decreasing Gummel number [12] of a SiGe HBT as explained by Kroemer [13]. As a result of the smaller Gummel number, doping concentration in the thin SiGe-base can be increased to avoid punch-through from emitter to collector. Therefore, the current gain can be

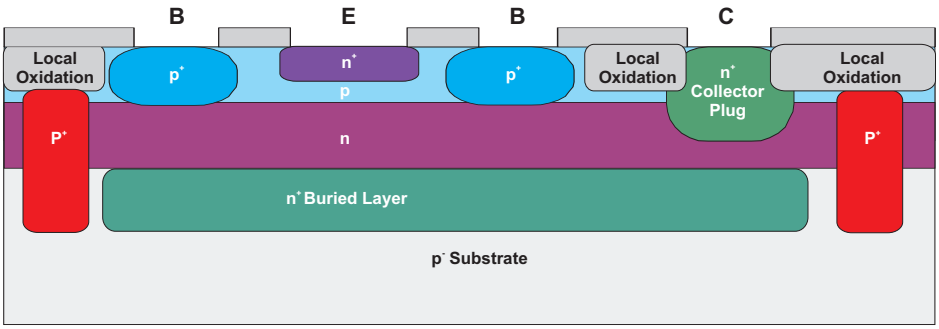


Figure 1.2: Cross-section of a conventional bipolar transistor.

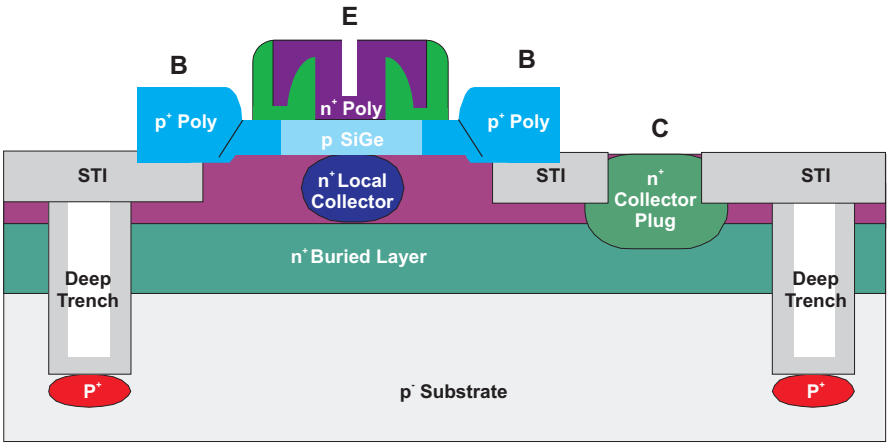


Figure 1.3: Cross-section of an advanced SiGe HBTs.

maintained or be even higher in spite of the high doping density in the base. On the other hand, a high doping density in the base can also reduce the base resistance. In the mean time, the SiGe HBT inherits the device structure from the advanced Si bipolar transistor. As shown in Fig. 1.3, an advanced SiGe HBT has an n^+ poly-emitter contact, a p^+ poly-base contact, an n^+ local collector implant, a buried n^+ sub-collector and deep trench isolation. Therefore, the SiGe-base HBT has superior performance compared to its Si-base counter part. Later in the late 90's, carbon incorporated in the SiGe base to suppress boron out-diffusion [14] from thermal process or transient enhanced diffusion is introduced to form a steeper base doping profile to further improve the performance of a SiGe HBT. With all the advanced features mentioned above, the most advanced SiGe:C HBT has pushed the peak f_T up to 380 GHz and a ECL gate delay down to merely 3.2 ps [15].

For the sake of higher power consumption and lower integration density of pure bipolar circuits compared to the CMOS circuits, bipolar transistors started to retreat from digital applications in the 70's and 80's. From the 90's up to now, it is the BiCMOS technology, which combines high-speed bipolar devices for analog applications with low-standby power CMOS devices for digital applications, to dominate in mixed-signal applications. Especially with the superior RF performance [16] of SiGe HBTs for a given lithographic generation compared to that of CMOS devices, SiGe HBTs are widely used in wireless transceiver circuits such as low-noise amplifiers (LNAs), mixers, voltage-controlled oscillators (VCOs) and power amplifiers (PAs). With the booming market for the wireless mobile communication devices, the need for low-power and high-speed BiCMOS circuits are expect to continue to grow in the future.

1.2 Modeling bipolar transistors

A compact semiconductor model is a mathematical description of the physical behavior of a semiconductor device. The first and simplest bipolar model is the Ebers-Moll (EM1) model, which considers two back-to-back PN diodes with only four model parameters to describe the large-signal behavior of a 3-terminal NPN/PNP bipolar transistor during transient analysis. This simple model was invented by J. J. Ebers and J. L. Moll [17] in 1954. The EM1 model is valid for 4 different regions of operation depending on the base-emitter and base-collector junction biases as shown in Fig. 1.4, where the reverse-active region is considered as the forward-active region of the common-

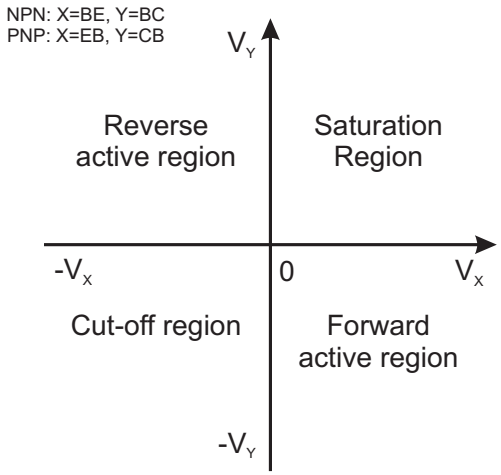


Figure 1.4: Regions of operation for a bipolar transistor.

collector configuration in the original EM1 model.

However, the bipolar circuit design was still a trial and error procedure on a breadboard with a lot of manual calculations for circuit analysis until the invention of the "simulation program with integrated circuit emphasis" (SPICE) [18, 19] in the 70's. The SPICE includes compact models for semiconductor devices and linear passive components (resistors, capacitors, inductors etc.) so that designers can perform circuit simulations in a computer aided design (CAD) environment. The original EM1 model is derived from a common-base configuration, which is later transformed to a common-emitter configuration with two base currents (I_{bf}, I_{br}) and a collector to emitter transfer current (I_N) as shown in Fig. 1.5. This is because the common-emitter configuration is better suited for the description of a switch or an amplifier. With limited device physics and regions modeled, the EM1 model was later extended to have more physical effects, including: 1. Charge storage in the junction and base regions and 2. parasitic resistances. As a result, an extended EM1 model with three non-linear depletion capacitances to model depletion charges (Q_{tE}, Q_{tC} and Q_{tS}), two non-linear diffusion capacitances to model minority charge storage in the neutral base (Q_{BE} and Q_{BC}) and three parasitic resistances (R_E, R_B and R_C) to model the bulk semiconductors in emitter, base and collector is shown in Fig. 1.5. The extended EM1 model is generally recognized as the second generation bipolar device model (EM2). Though the extrinsic region is larger than the intrinsic region in the conventional bipolar

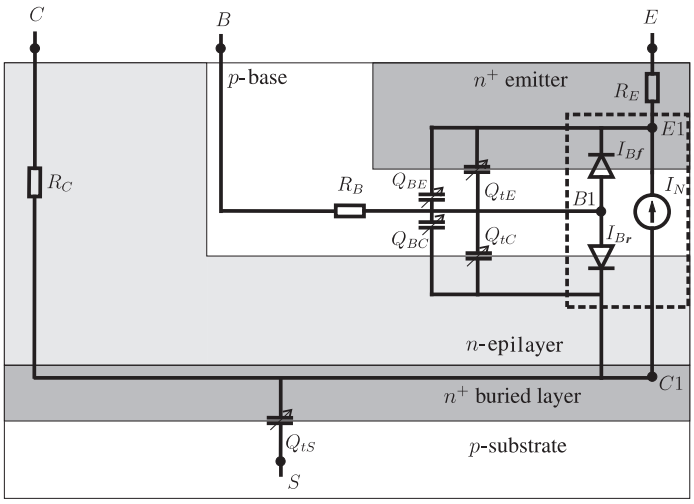


Figure 1.5: A simplified cross-section of a bipolar transistor with the Ebers-Moll model (in dashed box) and extended Ebers-Moll model (in the complete equivalent circuit).

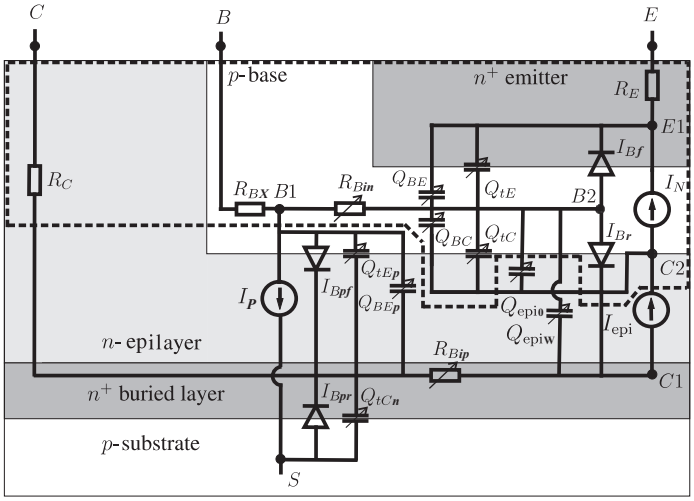


Figure 1.6: A simplified cross-section of a bipolar transistor with the Gummel-Poon model (in dashed box) and the extended Gummel-Poon model (in the complete equivalent circuit).

transistor [20], only the intrinsic region is modeled by EM1 and EM2 as can be seen in Fig. 1.5.

After the EM2 model, Getreu [21] classified a further extension of the EM model as a 3rd generation model (EM3) [22] and the Gummel-Poon (GP) model [23] from the 70's as a 4th generation model. The main difference between the EM3 and GP models is that there is an unified description of Early effect [24] and high-current injection in the base by a normalized base charge in the GP model. But both EM3 and GP models are still limited in modeling the intrinsic region of the device. The quasi-saturation effect is sill not addressed. With the advances of bipolar technology and higher integration density on a single chip in the 80's, intrinsic device delay from high current effects start to dominate the total delay time. Therefore, de Graaff [25] classifies the GP model as a 3rd generation bipolar model and Kull's extended GP (Ext. GP) model [26] as a 4th generation bipolar model. The Ext. GP model as shown in Fig 1.6 incorporates a bias dependent current source (I_{epi}) and two injected epi-layer charges (Q_{epi0} and Q_{epiw}) to model the quasi-saturation effect in the epi-collector and a parasitic PNP model to model the extrinsic region. The Mextram model [27], which will be described in the next chapter, is viewed as 5th generation model. Moreover, HICUM [28] and VBIC [29] are also classified as 5th generation bipolar models [30]. All three 5th generation models include new physical effects such as avalanche breakdown, self-heating, non-linear bias dependent Early effect etc. to model advanced Si/SiGe bipolar transistors. The main difference among the three 5th generations models is the regions where the integral charge control relationship (ICCR) [23] for the transfer current is applied as shown in Fig. 1.7. Earlier versions of Mextram [27] tried to calculate the transfer current from the drift-diffusion current equation to incorporate a non-uniform base doping induced electric field at low injection. The integral base charge at both low and high injections is evaluated directly from the junction law at quasi-neutral base (QNB) boundaries (x_{BE} to x_{BC}). In the later versions [31] , it came back to the linear transfer current and charge relationship as used in Ext. GP and VBIC models. As a result, Ext. GP, Mextram and VBIC all applied ICCR from x_{BE} to x_{BC} and they also have a similar voltage-dependent current source I_{epi} to model the epi-collector region. However, HICUM integrates charge for ICCR from edge of emitter-base depletion region ($E1$) to the edge of epi-collector ($C1$). Since the epi-collector is included in ICCR, no I_{epi} is needed for HICUM. The transfer current to integral charge relationship, however, becomes more complex to cover transistor operations from low injection to high injection.

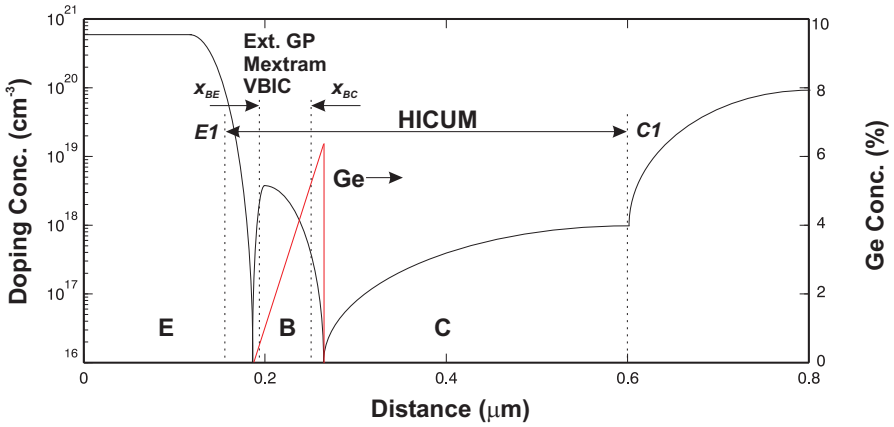


Figure 1.7: The 1-D SiGe HBT doping profile and integral charge control intervals employed in the evaluation of the transfer current [32].

After the development of the SiGe HBT technology in the early 90's, a Gummel number including the non-uniformed bandgap narrowing effect was later taken into account in HICUM [33] and Mextram [34] to model SiGe HBTs. When we look back to the bipolar model development history, we can conclude that compact modeling always lagged behind the device technology development. Therefore, numerical table-based models [35] or empirical models [36] are sometimes used to fill the gap between the compact modelling needs for the new technology generation and available compact device models.

1.3 Scope and outline

In this chapter, the evolution of bipolar transistor's technology and the main applications in each period of time since its invention has been described. Then, it is the discussion about how historically the compact model development has been driven by the technology progress. For the rest of the thesis, the outline is as follows.

In Chapter 2, the detailed equivalent circuit and physics of the Mextram model, on which the thesis is focused, is described.

In Chapter 3, a "short-cut" for compact model implementation with a high-level language VHDL Verilog-A to avoid time consuming C-code implementation and model verification after the implementation is discussed. Moreover, two additional features (tunneling current and behavioral substrate resistance)

for the Mextram model are implemented in Verilog-A to demonstrate that it is suitable for practicing new modeling ideas .

In Chapter 4, the bipolar transistor measurements (including DC, CV, S-parameters and 1/f noise), the test structures for use in device measurements and the pad parasitic de-embedding are discussed. Following the device measurements and test structure design, different measurement setups with their relevant Mextram model parameters are used to demonstrate parameter extraction from a SiGe HBT.

In Chapter 5, a reference based geometry scalable model and its parameter extraction procedure is proposed. The scaling rules for the scalable Mextram model are based on device physics and then re-formulated in terms of reference parameters, geometry parameters and drawn dimensions of the device layout for geometry scaling. As a result, the parameters can be systematically extracted from the measured data of devices with different geometries. The new approach is tested with SiGe HBTs within an IC-CAP model file, which is extended from the single device parameter extraction procedure.

In Chapter 6, the extension of the geometry scalable Mextram model to the configuration scalable model including different emitter numbers and collector contacts for high-power applications is described. The configuration scalable model uses almost the same set of model parameters as the geometry scalable model except additional geometry parameters representing the mutual heating effect and collector resistance between two adjacent fingers and additional instance parameters representing different layout configurations. The configuration model nicely predicts the f_T-degradation measured from SiGe HBTs due to an increase of the mutual heating and an increase the collector delay time with increasing emitter numbers.

Finally, Chapter 7 presents the conclusion and suggestions for the future directions in bipolar device compact modeling.

Chapter 2

Physics of the Mextram model

2.1 Introduction

The worldwide interest in SiGe heterojunction bipolar transistors (HBTs) as a commercial IC technology is growing rapidly [37]. The corresponding circuit design activities essentially depend on accurate compact models of SiGe HBTs in all relevant modes of the transistor operation. This chapter gives an overview of the vertical bipolar transistor compact model Mextram (the acronym of the "most exquisite transistor model") and its capabilities to meet SiGe HBT circuit design challenges. The Mextram has been introduced by De Graaff and Kloosterman at *Royal Philips Electronics* in 1985 [27]. The first Mextram release was introduced as Level 501 but the "real" 5th generation bipolar transistor model (after the previous four generation described in Chapter 1) is believed to be Level 502 in 1986 [38] with complete epi-layer model. Following the requirements of the continuous technology development, Mextram has appeared later in several update releases: Level 503 in 1994 [39] and Level 504 in 2000 [40].

Fig. 2.1 shows the equivalent circuit of the Mextram model as it is specified in its latest release (Level 504). The branches representing model currents and charges are schematically associated with different physical regions of a bipolar transistor separated by the base-emitter (BE), base-collector (BC) and substrate-collector (SC) junctions. All current and charge branches in Mextram are given as explicit functions of external and internal nodal potentials and there are no implicit modeling variables that require internal iterations. The governing Mextram equations are formulated having in mind NPN tran-

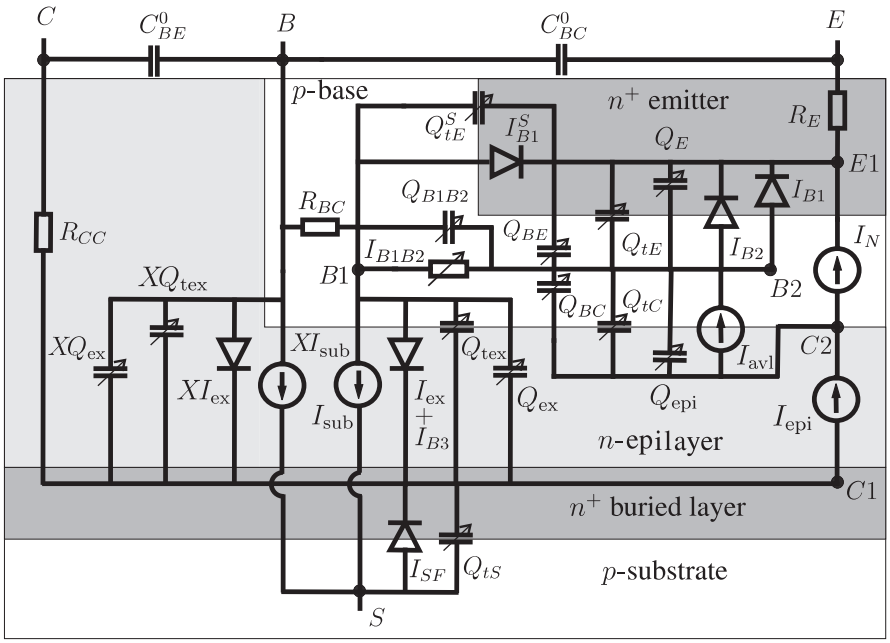


Figure 2.1: The equivalent circuit of Mextram model.

sistors, but the model can be equally well used for PNP transistors with the same structure by simply change the current and charge polarities.

The main transfer current I_N in Mextram, as in the Gummel-Poon model, is evaluated in the quasi-neutral base (QNB) by the ICCR. However, the mathematical appearance of its equations governing the normalized QNB charge is slightly different in comparison to the Gummel-Poon model. The normalized QNB charge in Mextram is represented as a product of a normalized base charge without injection charges and a normalized base charge with injection charges. In the Gummel-Poon model, it is represented as a normalized sum of zero bias base charge, modulated depletion charges and injection charges. Moreover, the effects of graded Ge profile in QNB [41] are physically addressed in the transfer current description. A distinguished feature of the Mextram model is the description of the epilayer transfer current I_{epi}. It is employed for intensive physical modelling of the quasi-saturation phenomena including base widening [8], hot-carrier behavior in the epilayer and advanced modeling of distortion effects [42].

The diode-like injection currents I_{B1}, I^S_{B1}, I_{B2}, I_{B3}, I_{ex} and XI_{ex} in the Mextram equivalent circuit describe various recombination currents in the

quasi-neutral and depletion transistor regions. The recombination in the modulated QNB, which is sometimes important for SiGe HBT applications [43], is also included. The effect of a distributed hole injection across the BE junction is described by an additional current branch I_{B1B2}. Mextram provides also a sophisticated model for the weak avalanche current in the branch I_{avl}. The contribution of the parasitic PNP transistor transfer current to the substrate current, represented by the current sources I_{sub} and XI_{sub}, is implemented by using a simplified Gummel-Poon integral charge control relationship.

The dynamic currents are accounted for by the depletion capacitances (charges) Q_{tE}, Q_{tE}^{S}, Q_{tC}, Q_{tex}, XQ_{tex} and Q_{tS} as well as diffusion charges (capacitances) Q_{BE}, Q_{BC}, Q_{E}, Q_{epi}, Q_{ex} and XQ_{ex} in the intrinsic and extrinsic transistor regions as shown in Fig. 2.1. The BC depletion capacitance is particularly equipped to account for the full depletion of the epilayer and modulation of the depletion charge by current in the BC depletion region. An additional charge branch Q_{B1B2} accounts for the distributed high frequency effects in the intrinsic transistor base.

The Mextram model is equipped with physical temperature scaling rules. Moreover, thermal phenomena are addressed in Mextram by a simple thermal impedance (thermal resistance and capacitance in parallel) and a power dissipation source but an elaborate thermal impedance network can be connected to the thermal node to model more complex self-heating and mutual heating effects. A set of temperature scaling parameters is extracted along with the corresponding electrical model parameters in a straightforward procedure [44]. The physical background of the Mextram parameters provides also an excellent framework for the geometrical [45] and configurational scaling [46].

Mextram has been already used excessively in various SiGe HBTs applications including high performance SiGe HBT bipolar transistor integration [47], low noise amplifiers [48,49], power amplifiers [50,51] and SiGe HBT phototransistors [52]. As an additional illustration of Mextram capabilities, some results of the Mextram parameter extraction from an IBM SiGe HBT test sample device are demonstrated through the Compact Modeling Council's [30] bipolar model standardization procedure.

2.2 Transfer current

The transfer current is maintained by the electron flow in vertical direction from the emitter contact to the collector buried layer and further up to the collector contact. The pure ohmic regions on that way are represented by

the constant *emitter resistance* `RE` and *collector resistance* `RCC`. The electron current density J_n in QNB and epilayer is assumed to obey the one-dimensional drift-diffusion equation

$$J_n = -q\mu_n n(x)\frac{d\phi_n(x)}{dx} \tag{2.1}$$

where n is the electron concentration, ϕ_n is the electron quasi-Fermi potential, μ_n is the electron mobility, q is the elementary charge and the x-axis is directed from the emitter towards the epilayer region. Moreover, in the model derivation the electron and hole concentrations are related by the pn product

$$p(x)n(x) = n_{ie}^2(x)\exp\left(\frac{\phi_{pB}-\phi_n(x)}{V_T}\right) \tag{2.2}$$

where p is the hole concentration, n_{ie} is the position dependent effective intrinsic carrier concentration, ϕ_{pB} is the constant hole quasi-Fermi level in the QNB, and V_T is the thermal voltage.

2.2.1 Quasi-Neutral Base (QNB)

The electron current density (2.1), and transfer current I_N, can be expressed in QNB with the help of the pn product (2.2) as

$$J_n = -\frac{I_N}{A_E} = qV_T\mu_n(x)\frac{n_{ie}^2(x)}{p(x)}\cdot\frac{d}{dx}\exp\left(\frac{\phi_{pB}-\phi_n(x)}{V_T}\right), \tag{2.3}$$

where A_E is the effective emitter area. Integrating (2.3) for constant J_n in the interval (x_{BE}, x_{BC}), where x_{BE} and x_{BC} denote the edges of the QNB, we have

$$I_N = \frac{qn_i^2A_E}{G\left(p; x_{BE}, x_{BC}\right)}\left[\exp\left(\frac{\phi_{pB}-\phi_n(x_{BE})}{V_T}\right) - \exp\left(\frac{\phi_{pB}-\phi_n(x_{BC})}{V_T}\right)\right], \tag{2.4}$$

where n_i is the intrinsic carrier concentration and the functional

$$G\left(p; x_1, x_2\right) = \int_{x_1}^{x_2}\frac{p(x)}{\mu_n V_T}\left(\frac{n_i}{n_{ie}(x)}\right)^2 dx \tag{2.5}$$

is well known as the Gummel number [12]. The transfer current (2.4) is implemented in Mextram as

$$I_N = \frac{\texttt{IS}}{g_B}\left[\exp\left(\frac{V_{B2}-V_{E1}}{V_T}\right) - \exp\left(\frac{V_{B2}-V_{C2}^*}{V_T}\right)\right], \tag{2.6}$$

where

$$g_B = \frac{G(p; x_{BE}, x_{BC})}{G(N_A; x^0_{BE}, x^0_{BC})} \tag{2.7}$$

is the relative change of the Gummel number in QNB with BE and BC junction biases, x^0_{BE} and x^0_{BC} denote the edges of the QNB at zero bias, N_A is the acceptor doping concentration in the base and $\texttt{IS} = qn_i^2 A_E / G(N_A; x^0_{BE}, x^0_{BC})$ is *transfer saturation current*. The quasi-Fermi potentials in (2.4) are replaced in (2.6) by the corresponding Mextram nodal biases. Note, that the value of the electron quasi-Fermi potential $\phi_n(x_{BC})$ in (2.4) is not directly associated with the nodal bias V_{C2}. Instead an effective value V^*_{C2} is used in Mextram (Level 504) as provided by the epilayer model description.

The relative Gummel number g_B in (2.7) is further expressed in Mextram as a product $g_B = g_{BW} \cdot g_{BH}$ where

$$g_{BW} = 1 + \frac{G(N_A; x_{BE}, x^0_{BE})}{G(N_A; x^0_{BE}, x^0_{BC})} + \frac{G(N_A; x^0_{BC}, x_{BC})}{G(N_A; x^0_{BE}, x^0_{BC})} \tag{2.8}$$

and

$$g_{BH} = 1 + \frac{G(n; x_{BE}, x_{BC})}{G(N_A; x_{BE}, x_{BC})}. \tag{2.9}$$

The role of g_{BW} is to describe the relative change of the Gummel number in QNB with the modulation of BE and BC depletion widths (Early effect) at low injection. On the other hand, the term g_{BH} accounts for the relative change of the Gummel number in the QNB due to the minority electron concentration at high injection currents.

Assuming a linear graded Ge profile (and band-gap E_g) in the QNB, as shown in Fig. 2.2, for the spatial dependence of $(n_i/n_{ie})^2$ in (2.5) holds

$$\left(\frac{n_i}{n_{ie}}\right)^2 \propto \exp\left(-\frac{\texttt{DEG}}{qV_T \cdot w_{B0}}\, x\right), \tag{2.10}$$

where $w_{B0} = x^0_{BC} - x^0_{BE}$ is the width of the QNB and $\texttt{DEG} = E_g(x^0_{BE}) - E_g(x^0_{BC})$ is the *band-gap difference over the base* (in eV), both evaluated at zero bias. Integrating Gummel numbers in (2.8) for constant N_A and μ_n using (2.10) the term g_{BW} becomes [34]

$$g_{BW} = 1 + \frac{\exp\left(\frac{\texttt{DEG}}{V_T}\frac{w_{BE}}{w_{B0}}\right) - 1}{1 - \exp\left(-\frac{\texttt{DEG}}{V_T}\right)} + \frac{\exp\left(-\frac{\texttt{DEG}}{V_T}\frac{w_{BC}}{w_{B0}}\right) - 1}{1 - \exp\left(\frac{\texttt{DEG}}{V_T}\right)}, \tag{2.11}$$

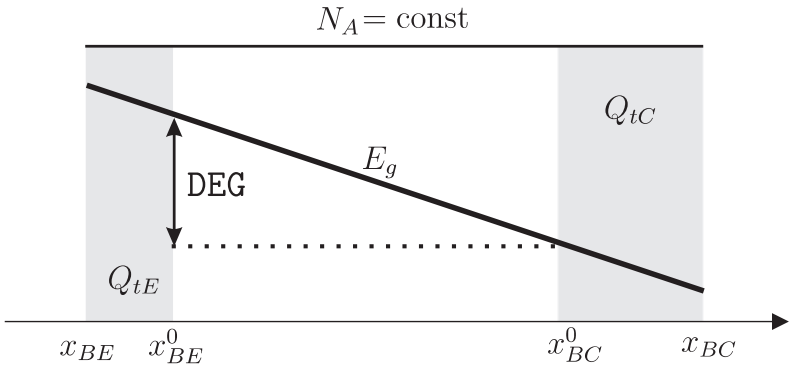


Figure 2.2: Doping and band-gap distribution in the QNB.

where $w_{BE} = x^0_{BE} - x_{BE}$ and $w_{BC} = x_{BC} - x^0_{BC}$ define the modulation of BE and BC depletion regions. Notice that in the limiting case $\texttt{DEG} = 0$, the term g_{BW} is reduced to

$$g^0_{BW} = 1 + \frac{w_{BE}}{w_{B0}} + \frac{w_{BC}}{w_{B0}} \tag{2.12}$$

representing the relative change of QNB width or the Early factor [21]. The relative variation of the depletion widths in (2.12) are implemented in Mextram in terms of the BE and BC depletion charges Q_{tE} and Q_{tC} as

$$\frac{w_{BE(C)}}{w_{B0}} = \frac{Q_{tE(C)}}{\texttt{VER(F)} \cdot \texttt{CJE(C)}}, \tag{2.13}$$

where `CJE` and `CJC` are *zero-bias BE and BC depletion capacitances* while `VER` and `VEF` are *reverse and forward Early voltages*. In that way, g^0_{BW} becomes identical to the Early factor proposed in [53] for silicon bipolar transistors. The model parameters `VER` and `VEF` may be interpreted for SiGe HBTs as the effective Early voltages that define the modulation rate of the QNB width instead of the Gummel number. For an alternative way to account for the Early effect in SiGe HBTs see [54].

The high injection term g_{BH} is implemented in Mextram in terms of minority electron concentration as

$$g_{BH} = 1 + \frac{1}{2}\left(\frac{n(x_{BE})}{N_A} + \frac{n(x_{BC})}{N_A}\right), \tag{2.14}$$

which for simplicity assumes uniform n_{ie} and linear distribution of n in QNB. The required normalized electron concentrations in (2.14) are evaluated solving

the system of the electro-neutrality equation $p = n + N_A$ and the pn product (2.2) at x_{BE} and x_{BC} as

$$\begin{aligned} \frac{n(x)}{N_A} &= F\left(\frac{n_{ie}^2}{N_A^2}\exp\left(\frac{V_{B2}-\phi_n(x)}{V_T}\right)\right) \\ F(z) &= \frac{2z}{1+\sqrt{1+4z}} \end{aligned} \tag{2.15}$$

where $\phi_n(x_{BE}) = V_{E1}$ and $\phi_n(x_{BC}) = V_{C2}^*$. This is also the place where the *forward knee current* IK is introduced in Mextram using

$$\frac{n_{ie}^2}{N_A^2} = \frac{\mathtt{IS}}{\mathtt{IK}} \tag{2.16}$$

but in principle it has the same physical role in the description of the high current effects as the corresponding knee current parameter of the Gummel-Poon model [31].

2.2.2 Epilayer

The drop of the electron quasi-Fermi potential over epilayer is essential for the correct evaluation of $\phi_n(x_{BC})$ and the transfer current I_N if the transistor is operating in quasi-saturation [55]. In that case, the epilayer region, between 0 and $w_{\rm epi}$ having uniform donor doping concentration N_{epi}, can be split into quasi-neutral injection region $0 < x < x_i$ and drift region $x_i < x < w_{\rm epi}$ as shown in Fig. 2.3.

From the pn product (2.2) we have

$$n(x)\frac{d\phi_n(x)}{dx} = -\frac{V_T}{p(x)}\frac{d\left(p(x)n(x)\right)}{dx} \tag{2.17}$$

which together with the quasi-neutrality condition $n = N_{epi} + p$ allows to express the electron current density (2.1) in epilayer as [26]

$$J_n = -\frac{I_{\rm epi}}{A_E} = q\mu_n N_{epi}\cdot V_T\left(\frac{2}{N_{epi}}+\frac{1}{p}\right)\frac{dp}{dx}. \tag{2.18}$$

Integrating (2.18) in the injection region for constant J_n and introducing the *epilayer resistance* $\mathtt{RCV} = w_{\rm epi}/(q\mu_n N_{epi} A_E)$ we have

$$\frac{x_i}{w_{\rm epi}}\mathtt{RCV}\cdot I_{\rm epi} = V_T\left(2\frac{p(0)-p(x_i)}{N_{epi}}+\ln\frac{p(0)}{p(x_i)}\right) \tag{2.19}$$

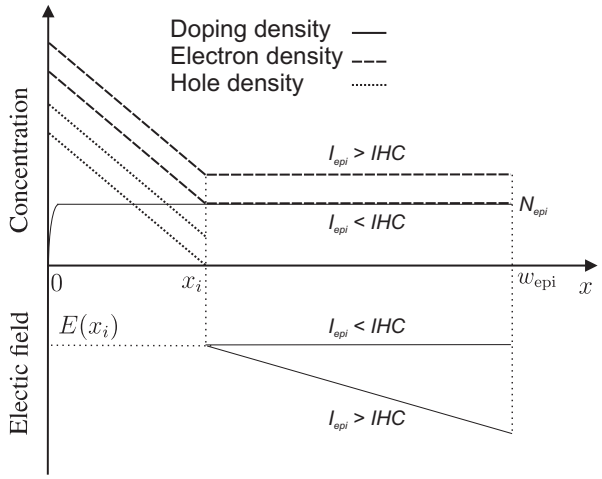


Figure 2.3: Distribution of the carrier concentrations and electric field in the epilayer when quasi-saturation occurs.

as an integral control relationship for the electron transport through the injection epilayer region. The normalized minority hole concentrations $p(0)/N_{epi}$ and $p(x_i)/N_{epi}$ in (2.19) are obtained similar to the electron normalized concentrations in (2.15) as

$$\frac{p(x)}{N_D} = F\left(\frac{n_{ie}^2}{N_{epi}^2}\exp\left(\frac{V_{B2}-\phi_n(x)}{V_T}\right)\right) \tag{2.20}$$

but this time the doping related term n_{ie}^2/N_{epi}^2 is represented with the *BC built-in voltage* VDC as

$$\frac{n_{ie}^2}{N_{epi}^2} = \exp\left(-\frac{\mathtt{VDC}}{V_T}\right). \tag{2.21}$$

The drift epilayer region is neutral and has ohmic behavior as long as $I_{\rm epi}$ is significantly smaller than the *critical current for velocity saturation* $\mathtt{IHC} = qA_E N_{epi} v_{\rm sat}$, where $v_{\rm sat}$ is the electron saturation velocity. Otherwise, the electron concentration becomes current dependent as $n = N_{epi} I_{\rm epi}/\mathtt{IHC}$ and the electric field $E(x)$ has slope defined by the Poisson equation

$$\frac{dE(x)}{dx} = \frac{qN_{epi}}{\epsilon}\left(1-\frac{I_{\rm epi}}{\mathtt{IHC}}\right), \tag{2.22}$$

where ϵ is the permittivity of the epilayer. It should be emphasized that for the drift electron transport holds $E(x) = -d\phi_n(x)/dx$ which allows to directly

evaluate the drop of ϕ_n over the drift epilayer region integrating the Poisson equation (2.22) twice in the interval $(x_i, w_{\rm epi})$. The result is

$$V_{C1} - \phi_n(x_i) = -E(x_i) w_{\rm epi} d_R + \frac{qN_{epi} w_{\rm epi}^2}{2\epsilon} \left(\frac{I_{\rm epi}}{\texttt{IHC}} - 1 \right) d_R^2, \tag{2.23}$$

where $d_R = 1 - x_i/w_{\rm epi}$ is the relative depth of the drift epilayer region. Using the interface condition

$$E(x_i) = -\frac{v_s}{\mu_n} = -\frac{\texttt{IHC} \cdot \texttt{RCV}}{w_{\rm epi}} \tag{2.24}$$

and introducing the *epilayer space charge resistance* $\texttt{SCRCV} = qN_{epi} w_{\rm epi}^2/(2\epsilon \cdot \texttt{IHC})$, (2.23) becomes

$$V_{C1} - \phi_n(x_i) = \texttt{IHC} \cdot \texttt{RCV} \cdot d_R + (I_{\rm epi} - \texttt{IHC}) \cdot \texttt{SCRCV} \cdot d_R^2 \tag{2.25}$$

and represents the governing equation for the electron transport through the drift epilayer region if $I_{\rm epi} > \texttt{IHC}$. The validity of (2.25) could be extended to the ohmic case after transformation

$$\texttt{IHC} \to \frac{\texttt{IHC} \cdot I_\Omega}{\texttt{IHC} + I_\Omega} \tag{2.26}$$

where

$$I_\Omega = \frac{V_{C1} - \phi_n(x_i)}{\texttt{RCV} \cdot d_R} \tag{2.27}$$

is the ohmic current in the drift epilayer region.

The integral relationships (2.19) and (2.25) for the injection and drift epilayer regions still require an additional condition to close the system of equations for unknown $I_{\rm epi}$, $x_i/w_{\rm epi}$ and $\phi_n(x_i)$. It may be the continuation of the electric field as proposed in [55]. In order to avoid implicit model variables and to improve smoothness of the modeling equations the present Mextram release (Level 504) implements the above epilayer physics in a qualitatively different way. Namely, the epilayer current $I_{\rm epi}$, as seen from the circuit simulator, is evaluated only from (2.19) applied to the whole epilayer. The governing equations (2.19) and (2.25) are then employed in the evaluation of the effective electron quasi-Fermi potential $V_{C2}^* = \phi_n(x_{BC})$ that substitutes the nodal bias V_{C2} in all subsequent calculations. The detailed implementation procedure is given in [31, 56].

2.3 Recombination currents

The recombination currents in the intrinsic transistor region, due to the hole injection into quasi-neutral emitter (QNE) and BE space-charge region as well as recombination in QNB, actually serve to model the transistor forward base current. The ohmic part of the base region is represented by the *constant base resistance* `RBC`. The electron and hole injection across the BC junction, in the reverse transistor operation, are addressed in Mextram in the extrinsic transistor area. The carrier generation due to the weak avalanche effects is introduced separately as a controlled current source.

2.3.1 Intrinsic transistor region

The carrier recombination in the intrinsic transistor region is indirectly evaluated in terms of the currents injected into quasi-neutral or space-charge regions. These currents are typically given in the form of a diode-like characteristics

$$I_D(V_j, I_0, m) = I_0 \left[\exp\left(\frac{V_j}{m V_T} \right) - 1 \right], \tag{2.28}$$

where V_j is the junction voltage, I_0 is the diode saturation current and m is the non-ideality factor.

The most important contribution to the static base current is the recombination (injection) into QNE. It is expressed in terms of the diode current (2.28) as

$$I_{BE} = I_D \left(V_{B2E1}, \frac{\mathtt{IS}}{\mathtt{BF}}, 1 \right), \tag{2.29}$$

where `BF` is the *ideal forward current gain*.

The recombination in the QNB may produce significant impact on the performance of SiGe HBTs [43]. It could be physically evaluated as an additional current component

$$I_{BB} = q A_E \int_{x_{BE}}^{x_{BC}} \frac{\Delta n}{\tau_n} \, dx, \tag{2.30}$$

where Δn is the excess minority carrier concentration and τ_n is the minority carrier lifetime in the QNB. Assuming that the excess minority concentrations

at the QNB boundaries are proportional to the injection currents as [57]

$$\Delta n(x_{BE}) \propto I_D\left(V_{B2E1}, \frac{\texttt{IS}}{\texttt{BF}}, 1\right) \quad \text{and} \quad \Delta n(x_{BC}) \propto I_D\left(V^*_{B2C2}, \frac{\texttt{IS}}{\texttt{BF}}, 1\right), \tag{2.31}$$

the QNB recombination current (2.30) is implemented in Mextram as [34]

$$I_{BB} = \texttt{XREC} \cdot \left[I_D\left(V_{B2E1}, \frac{\texttt{IS}}{\texttt{BF}}, 1\right) + I_D\left(V^*_{B2C2}, \frac{\texttt{IS}}{\texttt{BF}}, 1\right)\right]\left(1 + \frac{w_{BC}}{w_0}\right), \tag{2.32}$$

where XREC is the *EB recombination current prefactor*. Note, that the width modulation of QNB, especially that due to the base-collector depletion capacitance, produces an Early-like effect in the forward base current.

The hole injection into QNE is essentially a two-dimensional phenomenon. Namely, a fraction of the holes is injected along the side walls of the BE junction. Moreover, the hole injection is laterally nonuniform along the intrinsic base below the emitter due to the variations of the internal base-emitter junction bias. The sidewall base current component is introduced by splitting the injection current into the area

$$I_{B1} = (1 - \texttt{XIBI})\left[(1 - \texttt{XREC})\, I_{BE} + I_{BB}\right] \tag{2.33}$$

and sidewall

$$I^S_{B1} = \texttt{XIBI} \cdot I_D\left(V_{B1E1}, \frac{\texttt{IS}}{\texttt{BF}}, 1\right) \tag{2.34}$$

components by using the *diode partition factor* XIBI. The distributed injection along the pinched transistor base below the emitter is emulated using a nonlinear current branch [58, 59]

$$I_{B1B2} = \frac{g^0_{BW}}{3 \cdot \texttt{RBV}}\left\{V_{B1B2} + 2V_T\left[\exp\left(\frac{V_{B1B2}}{V_T}\right) - 1\right]\right\}, \tag{2.35}$$

where RBV is the *resistance of the pinched base under the emitter* at low injection. It may be approximated as

$$\texttt{RBV} = \frac{\rho_\square W_E}{3L_E} \tag{2.36}$$

for the given pinched base sheet resistance $\rho_\square$ as well as emitter width and length W_E and L_E, respectively.

The Shockley-Read-Hall (SRH) recombination [60, 61] in the BE space-charge region is implemented in Mextram as a non-ideal diode current

$$I_{B2} = I_D\left(V_{B2E1}, \texttt{IBF}, \texttt{MLF}\right) + G_{\min} V_{B2E1}, \tag{2.37}$$

where IBF is the *saturation current* and MLF is the *emission coefficient* of the BE leakage diode. A small conductance $G_{\min} = 10^{-13}\,\Omega^{-1}$ in (2.37) is introduced for numerical stability.

2.3.2 Extrinsic quasi-neutral regions

The current $I_{\rm ex}$ describes the recombination of carriers injected into the quasi-neutral regions of the extrinsic BC junction. It is evaluated by

$$I_{\rm ex} = \frac{\texttt{IS}}{\texttt{BRI}} \frac{\exp\left(\dfrac{V_{B1C1}}{V_T}\right) - 1}{1 + \dfrac{1}{2}\dfrac{n(x_{BC}^{\rm ext})}{N_A}}, \tag{2.38}$$

where BRI is the *ideal reverse current gain* and

$$\frac{n(x_{BC}^{\rm ext})}{N_A} = F\left(\frac{\texttt{IS}}{\texttt{IK}} \exp\left(\frac{V_{B1C1}}{V_T}\right)\right) \tag{2.39}$$

is the electron concentration at the edge of extrinsic QNB which is obtained in the same way as $n(x_{BE})$ but for the bias V_{B1C1}. In principle $I_{\rm ex}$ represents the ideal component of the reverse base current taking also high injection effects into account. An additional extrinsic current component $XI_{\rm ex}$, similar to (2.38) but evaluated for internal bias V_{BC1}, is introduced if the model flag EXMOD is set to 1.

It is also assumed that SRH recombination in the BC space-charge region considerably contributes to the extrinsic base current component. This current component is derived from maximum value of the net SRH recombination in the space charge region as [25]

$$I_{B3} = \texttt{IBR} \frac{\exp\left(\dfrac{V_{B1C1}}{V_T}\right) - 1}{\exp\left(\dfrac{V_{B1C1}}{V_T}\right) + \exp\left(\dfrac{\texttt{VLR}}{2V_T}\right)} + G_{\min} V_{B1C1}, \tag{2.40}$$

where IBR is *saturation current* and VLR is the *crossover voltage* of the BC leakage diode.

2.3.3 Weak avalanche current

The avalanche current I_{avl} is evaluated in Mextram as [62, 63]

$$I_{\text{avl}} = I_{\text{epi}} \int_0^{\texttt{WAVL}} \alpha_n \left(E(x)\right)\, dx, \tag{2.41}$$

where WAVL the *effective width of the epilayer*. The ionization rate α_n is given by [64]

$$\alpha_n(E(x)) = A_n \exp\left(-\frac{B_n}{E(x)}\right) \tag{2.42}$$

and A_n and B_n are predefined model constants. Assuming the linear electric field distribution

$$E(x) = E_M\left(1 - \frac{x}{\lambda_D}\right) \approx \frac{E_M}{1 + x/\lambda_D} \tag{2.43}$$

with the slope $-1/\lambda_D$ around the maximum electric field value E_M, eqn.(2.41) may be integrated or express the weak avalanche current as

$$I_{\text{avl}} = I_{\text{epi}} \frac{A_n}{B_n} \lambda_D E_M \left[\exp\left(-\frac{B_n}{E_M}\right) - \exp\left(-\frac{B_n}{E_M}\left(1 + \frac{\texttt{WAVL}}{\lambda_D}\right)\right)\right]. \tag{2.44}$$

The maximum electric field value E_M and λ_D are obtained from the Poisson equation (2.22) in the epilayer. To this end, it is rewritten as

$$\frac{dE(x)}{dx} = 2\frac{\texttt{VAVL}}{\texttt{WAVL}^2}\left(1 - \frac{I_{epi}}{\texttt{IHC}}\right), \tag{2.45}$$

where $\texttt{VAVL} = qN_D\texttt{WAVL}^2/(2\epsilon)$ is the *avalanche curvature voltage*. If the model flag EXAVL is set to 1 the weak avalanche model handles also the electric field distribution in quasi-saturation due to the Kirk effect [31].

2.4 Substrate currents

The substrate current is implemented in Mextram using a simplified Gummel-Poon integral charge control relationship for the parasitic PNP transistor:

$$I_{\text{sub}} = \frac{2 \cdot \texttt{ISS}\left[\exp\left(\frac{V_{B1C1}}{V_T}\right) - 1\right]}{1 + \sqrt{1 + 4\frac{\texttt{IS}}{\texttt{IKS}}\exp\left(\frac{V_{B1C1}}{V_T}\right)}}, \tag{2.46}$$

where ISS is PNP *transistor saturation current* and IKS is the *substrate knee current*. Note, that the effects of the base width modulation by depletion capacitances is neglected and $V_{SC1} = 0$ is assumed. Moreover, the high-injection effects are expressed in terms of IS/IKS instead of ISS/IKS to simplify parameter extraction. An additional extrinsic substrate current component $XI_{\rm sub}$, similar to (2.46), but evaluated for an internal bias V_{BC1}, is introduced if the model flag EXMOD is set to 1. A diode-like current I_{SF} is added between the substrate and collector nodes, S and $C1$, to serve as an indicator of a falsely polarized SC junction.

It should be emphasized that the Mextram equivalent circuit is deliberately left without any circuit elements connecting the intrinsic substrate node S to the substrate contact. In that way, an external substrate network of arbitrary complexity could easily be attached to the intrinsic substrate node [65, 66].

2.5 Charges and capacitances

The temporal variation of the electric field results in the displacement current components across the space charge regions. These dynamic current components are represented by the depletion capacitances (charges). On the other hand, the temporal variations of the compensated (diffusion) charges, produce an effective *dynamic recombination* current along the transistor transfer current flow which is implemented by the diffusion charges (capacitances). Mextram (Level 504) takes also into account the *BE and BC parasitic overlap capacitances* CBEO and CBCO as shown in Fig. 2.1.

2.5.1 Depletion capacitances

The bias dependence of the depletion capacitances is in Mextram generally considered as

$$C_t(V_j) = (1 - X_p)\,\frac{C_0 f_I}{\left(1 - \dfrac{V_j^{\rm eff}(V_j)}{V_{\rm bi}}\right)^P} + X_p C_0, \tag{2.47}$$

where C_0 is the zero bias depletion capacitance, V_j is the internal P-N junction bias, $V_{\rm bi}$ is the junction built-in voltage and P is the grading coefficient. The expression (2.47) is inspired by the simple empirical description for depletion capacitance of abrupt (or linear) P-N junctions but also enhanced, with

quantities X_p and f_I as well as function $V_j^{\text{eff}}(V_j)$, to increase the physical and computational range of the model validity.

In order to avoid singular capacitance behavior at the forward bias, an effective junction bias $V_j^{\text{eff}}(V_j)$ is employed in the denominator of (2.47). It is related to the real junction bias V_j as

$$V_j^{\text{eff}}(V_j) = V_j - V_{\text{ch}} \ln\left[1 + \exp\left(\frac{V_j - V_F}{V_{\text{ch}}}\right)\right], \tag{2.48}$$

where the control voltage

$$V_F = V_{\text{bi}} \cdot \left(1 - a_j^{-1/P}\right) \tag{2.49}$$

forces the capacitance to asymptotically approach the constant value $a_j C_0$ for $V_j > V_F$ (see Fig. 2.4). The smoothness of this transition is defined by V_{ch}. The quantity X_p in (2.47) limits the decrease of the capacitance under the reverse bias. It is of particular importance for BC depletion capacitance having a fully depleted epilayer region. Finally, the term f_I accounts for the modulation of the BC depletion capacitance by the transfer epilayer current. It is defined as

$$f_I = \left(1 - \frac{I_{\text{epi}}}{I_{\text{epi}} + \texttt{IHC}}\right)^{\texttt{MC}} \tag{2.50}$$

where `MC` is the *current-modulation coefficient*. Table 2.1 provides the corresponding model parameters in Mextram for the BE, BC and SC junctions.

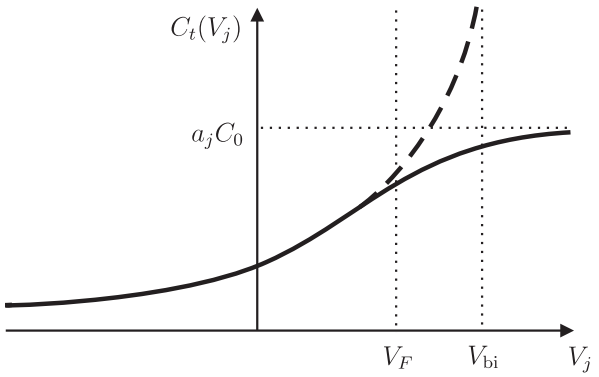


Figure 2.4: Implementation of the Mextram depletion capacitances.

Table 2.1: Parameters for Depletion capacitances.

	$V_{\rm bi}$	C_0	P	X_p	f_I
BE	VDE	CJE	PE	-	-
BC	VDC	CJC	PC	XP	MC
SC	VDS	CJS	PS	-	-

For compact modeling purposes, it is much better to consider corresponding depletion charges

$$Q_t(V_j) = \int_0^{V_j} C_t\left(V_j^{\rm eff}(V_j)\right)\, dV, \tag{2.51}$$

instead of the depletion capacitances. The Mextram depletion charges (see Fig. 2.1) are implemented as

$$Q_{tE}(V_{B2E1}) = (1 - \texttt{XCJE}) \cdot Q_t\left(V_{B2E1}\right), \tag{2.52}$$
$$Q_{tE}^S(V_{B1E1}) = \texttt{XCJE} \cdot Q_t\left(V_{B1E1}\right), \tag{2.53}$$
$$Q_{tC}(V_{B2C2}) = \texttt{XCJC} \cdot Q_t\left(V_{jC}(V_{B2C2})\right), \tag{2.54}$$
$$Q_{\rm tex}(V_{B1C1}) = (1 - \texttt{XCJC})\,(1 - \texttt{XEXT})\, Q_t\left(V_{B1C1}\right), \tag{2.55}$$
$$XQ_{\rm tex}(V_{BC1}) = (1 - \texttt{XCJC}) \cdot \texttt{XEXT} \cdot Q_t\left(V_{BC1}\right), \tag{2.56}$$
$$Q_{tS}(V_{SC1}) = Q_t\left(V_{SC1}\right), \tag{2.57}$$

where XCJE, XCJC and XEXT are *geometry partitioning factors* splitting the depletion capacitances into area and sidewall components as well as between intrinsic and extrinsic parts of the transistor. Note, that the internal BC junction bias V_{jC} in (2.54) has to be evaluated implicitly as a function of the bias V_{B2C2} in order to correctly take into account the effects of quasi-saturation in the epilayer [31].

2.5.2 Diffusion charges

The diffusion charges are evaluated in Mextram independently for the QNB, QNE and epilayer as well as for the extrinsic transistor region.

For the linear distribution of electrons in QNB, the total base diffusion charge is

$$Q_B = Q_{B0} \cdot g_{BW}^0 \cdot \frac{1}{2}\left(\frac{n(x_{BE})}{N_A} + \frac{n(x_{BC})}{N_A}\right), \tag{2.58}$$

where $Q_{B0} = qA_E W_{B0} N_A$. In the quasi-static approximation, the base diffusion charge (2.58) is split into BE and BC components Q_{BE} and Q_{BC} as

$$Q_{BE(C)} = \frac{1}{2} \cdot \texttt{TAUB} \cdot \texttt{IK} \cdot g^0_{BW} \cdot \frac{n(x_{BE(C)})}{N_A}, \tag{2.59}$$

associated with Mextram nodes $E1$ and $C2$. The model parameter $\texttt{TAUB} = Q_{B0}/\texttt{IK}$ is introduced as *base transit time*.

The diffusion charge in the QNE (including the compensated charge in the BE space-charge region is expressed as

$$Q_E = Q_{E0} \left[\exp\left(\frac{V_{B2E1}}{\texttt{MTAU} \cdot V_T} \right) - 1 \right], \tag{2.60}$$

where MTAU is the *emitter diffusion charge coefficient*. The emitter transit time can approximately be expressed from (2.60) as

$$\tau_E(I_N) \approx \frac{Q_{E0}}{I_N} \left(\frac{I_N}{\texttt{IS}} \right)^{1/\texttt{MTAU}}. \tag{2.61}$$

Introducing an *emitter transit time* as $\texttt{TAUE} = \tau_E(\texttt{IK})$, the prefactor Q_{E0} in (2.60) becomes

$$Q_{E0} = \texttt{TAUE} \cdot \texttt{IK} \left(\frac{\texttt{IS}}{\texttt{IK}} \right)^{1/\texttt{MTAU}} \tag{2.62}$$

as it is implemented in Mextram. The epilayer diffusion charge

$$Q_{\text{epi}} = qA_E \int_0^{x_i} p(x)\, dx \tag{2.63}$$

actually represent the hole (minority carrier) charge in the injection epilayer region. This charge can be related to the epilayer current by the Gummel integral charge relationship

$$I_{\text{epi}} = \frac{q^2 n_i^2 A_E^2 D_n}{Q_{\text{epi}}} \left[\exp\left(\frac{V^*_{B2C2}}{V_T} \right) - \exp\left(\frac{V_{B2} - \phi_n(x_i)}{V_T} \right) \right]. \tag{2.64}$$

Introducing the *epilayer transit time* $\texttt{TEPI} = W^2_{\text{epi}}/(4D_n)$ and with the help of pn product (2.2), the epilayer diffusion charge can be expressed also as

$$Q_{epi} = \frac{Q^2_{epi0}}{4 \cdot \texttt{TEPI} \cdot I_{\text{epi}}} \left[\frac{p(0)}{N_D} \left(\frac{p(0)}{N_D} + 1 \right) - \frac{p(x_i)}{N_D} \left(\frac{p(x_i)}{N_D} + 1 \right) \right], \tag{2.65}$$

where

$$Q_{epi0} = qA_E W_{\text{epi}} N_D = \frac{4 \cdot \texttt{TEPI} \cdot V_T}{\texttt{RCV}}. \tag{2.66}$$

For practical implementation in Mextram, the expression for Q_{epi} is combined with the expression for the epilayer current I_{epi} and further simplified [31,67].

The extrinsic diffusion charge is evaluated combining the expression for the injection in QNB and the epilayer in the extrinsic part of the transistor:

$$Q_{\text{ex}} = \frac{\texttt{TAUR}}{\texttt{TAUB} + \texttt{TEPI}} \frac{1}{2} \left(Q_{B0} \frac{n(x_{BC}^{ext})}{N_A} + Q_{\text{epi}} \frac{p(x_W^{\text{ex}})}{N_D} \right), \tag{2.67}$$

where `TAUR` is the *reverse base transit time*,

$$\frac{p(x_W^{\text{ex}})}{N_D} = F \left(\exp \left(\frac{V_{B1C1} - \texttt{VDC}}{V_T} \right) \right), \tag{2.68}$$

while $n(x_{BC}^{ext})/N_A$ is given in (2.39). Further partition of the extrinsic diffusion charge, similar to that of the extrinsic injection currents, is possible if the model flag `EXMOD` is set to 1. In that case, a similar expression to (2.67) is used to evaluate the extrinsic charge XQ_{ext} in terms of junction bias V_{BC1}. The *BC diode partition factor* `XEXT` is used to split their contributions.

2.5.3 Distributed and non-quasi-static charges

In high-frequency and high-speed applications the quasi-static assumption is no longer valid. Moreover, the effects of the distributed capacitances along the BE junction should be taken into account. The high-frequency current-crowding effects are model by introducing an effective charge branch

$$Q_{B1B2} = \frac{1}{5} \frac{dQ_{B2E1}}{dV_{B2E1}} V_{B1B2}, \tag{2.69}$$

where $Q_{B2E1} = Q_{tE} + Q_{BE} + Q_E$.

The non-quasi-static effects in QNB base are accounted for in Mextram by introducing charge partitioning:

$$Q_{BC} \rightarrow \frac{1}{3} Q_{BE} + Q_{BC} \quad \text{and} \quad Q_{BE} \rightarrow \frac{2}{3} Q_{BE}. \tag{2.70}$$

Both modeling options require the flag `EXPHI` to be set to 1.

2.6 Thermal phenomena

The electrical characteristics of bipolar transistors are particularly susceptible to temperature variations due to the self-heating or thermal interaction with other devices. The electro-thermal effects are implemented in Mextram as a combination of the electrical model (current and charge branches), with temperature dependent parameters, and a thermal model that links the device average temperature to the dissipated electrical power. The temperature scaling parameters are based on strong physical background of the electrical parameters.

The default Mextram thermal model is a one-pole linear thermal impedance (parallel *thermal resistance* `RTH` and *thermal capacitance* `CTH` of the device surrounding) connected between the external thermal node dT and the zero bias (ambient) device temperature. If necessary, the default thermal impedance may be suspended and bypassed by a more advanced thermal network via the external thermal node. In order to correctly take into account the time delays of the internal biases, the total dissipated power in the device is calculated as a sum of the power dissipated in all nonreactive circuit elements.

Most of the Mextram current and charge modeling expressions have explicit temperature dependence. The actual device temperature is expressed as

$$T = T_A + \texttt{DTA} + V_{dT} + 273.15\ K, \tag{2.71}$$

where T_A is the ambient (simulation) temperature, parameter `DTA` specifies a constant temperature shift to the ambient temperature and V_{dT} is the temperature at the thermal node dT. The electrical potential at the thermal node dT actually represents the excess device temperature in K. The temperature at which the electrical parameters are extracted is the *reference temperature* `TREF` in ${}^0C$ or $T_{\rm ref} = \texttt{TREF} + 273.15\ K$.

The model parameters depend implicitly on temperature via the intrinsic carrier concentration and the carrier mobility temperature dependence. The temperature dependence of the intrinsic carrier concentration is

$$n_i^2 \propto t_N^3 \exp\left(-\frac{\texttt{VG*}}{V_{\Delta T}}\right), \tag{2.72}$$

where

$$t_N = \frac{T}{T_{\rm ref}}, \tag{2.73}$$

$$\frac{1}{V_{\Delta T}} = \frac{q}{k}\left(\frac{1}{T} - \frac{1}{T_{\rm ref}}\right), \tag{2.74}$$

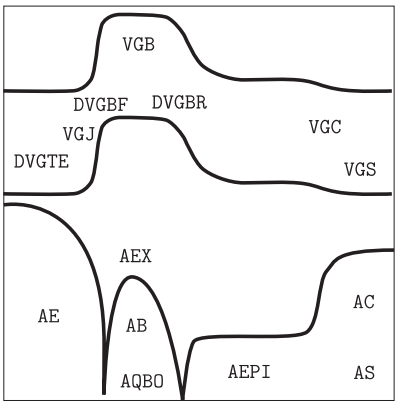


Figure 2.5: Temperature scaling parameters.

and k is the Boltzmann constant. The '$*$' may be `B`, `C`, `S` and `J` for the band-gap in base, collector, substrate and BE depletion region, respectively. This approach is particularly suitable for HBTs with varying band-gap across the device. The carrier mobility is scaled with temperature as

$$\mu \propto t_N^{-\mathrm{A}*}, \tag{2.75}$$

where '$*$' may be `E`, `B`, `EX EPI`, `C` and `S` for the emitter, base, extrinsic base, epilayer, collector and substrate region, respectively.

A few temperature scaling parameters, which could particularly be important for SiGe HBT applications, are introduced separately for certain electrical parameters. The forward and reverse current gain parameters `BF` and `BRI` depend on the difference of band-gaps at BE and BC junctions. Their temperature dependence is additionally expanded by

$$\mathtt{BF(RI)} \propto \exp\left(-\frac{\mathtt{DVGBF(R)}}{V_{\Delta T}}\right), \tag{2.76}$$

where `DVGBF` and `DVGBR` are the *band-gap voltage differences*. An additional temperature parameter `DAIS` for fine tuning of temperature dependence of collector-emitter saturation current is given as

$$\mathtt{IS} \propto t_N^{4-\mathtt{AB}-\mathtt{AQBO}+\mathtt{DAIS}} \exp\left(-\frac{\mathtt{VGB}}{V_{\Delta T}}\right). \tag{2.77}$$

It is well known that the compensated charge in the BE space-charge region could have significant effect on the dynamic performance of SiGe HBTs

[57]. This charge also has a different temperature dependence from the base diffusion charge. In Mextram it may easily be taken into account by the temperature dependence

$$\texttt{TAUE} \propto \exp\left(-\frac{\texttt{DVGTE}}{V_{\Delta T}}\right), \tag{2.78}$$

where DVGTE is the *emitter transit time band-gap voltage difference*. Finally, the base zero-bias charge Q_{B0} has a separate temperature scaling coefficient AQBO. Since it accounts also for the temperature dependence of w_{B0}, it is crucial for the temperature scaling of Early voltages VER and VEF but also for the Ge-related parameter DEG as

$$\texttt{DEG} \propto t_N^{\texttt{AQBO}}. \tag{2.79}$$

Fig. 2.5 shows the complete set of Mextram temperature scaling parameters associated with different band-gaps or doping areas along the transistor structure.

2.7 Noise

Noise is the statistical fluctuation of the terminal currents or voltages due to the quantization of charge, thermal motion of the carriers and generation or recombination process [68]. There are three different kinds of noise sources in the Mextram model, which are the thermal noise, shot noise and flicker noise. The thermal and shot noises are white noises, which are frequency independent, while the flicker noise is frequency dependent.

The thermal noise is generated from the random thermal motion of the carriers in a conducting material. As a result, a noise voltage source S_v in the unit of power spectral density (V^2/Hz) in the frequency band of 1 Hz is in series with a noiseless constant resistance R during noise analysis, expressed as

$$S_v = 4kTR. \tag{2.80}$$

The term S_v in series with R can be transformed to an equivalent noise current source S_i (A^2/Hz) in parallel with R as shown in Fig. 2.6 by

$$\begin{aligned} S_i &= \frac{S_v}{R^2}, \\ &= \frac{4kT}{R}. \end{aligned} \tag{2.81}$$

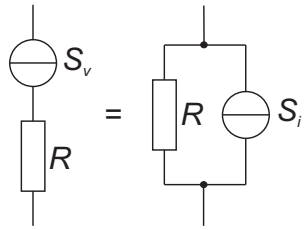


Figure 2.6: thermal noise equivalent circuit for a resistor.

In Mextram, a noise current source S_i is added in parallel with the temperature scaled resistances $\mathtt{RE}_\mathrm{T}$, $\mathtt{RBC}_\mathrm{T}$ and $\mathtt{RCC}_\mathrm{T}$, but the thermal noise from the variable base resistance due to current crowding is described as [69] :

$$S_{iR_{BV}} = 4kT\frac{g^0_{BW}}{\mathtt{RBV}_\mathrm{T}}\frac{4e^{V_{B1B2}/V_T}+5}{9}, \tag{2.82}$$

which is added to the nonlinear current branch I_{B1B2}.

The shot noise that arises from the dc current I flowing through a PN junction is described as

$$S_i = 2q|I|. \tag{2.83}$$

The shot noise source S_i is added in parallel with the branch current I_N, I_{B1}, I_{B2}, I^S_{B1}, I_{B3}, I_{sub}, XI_{sub} and I_{ex}. When $\mathtt{EXMOD}$=1, the shot noise contribution from the sidewall base current (XI_{ex}) of the parasitic PNP is also taken into account.

The flicker noise is a low frequency noise source due to the generation recombination process in the PN junction. It is empirically modeled as

$$S_i = \frac{\mathtt{KF}}{f}|I|^{\mathtt{AF}}, \tag{2.84}$$

where f is the measured frequency. $\mathtt{AF}$ and $\mathtt{KF}$ are fitting parameters. The flicker noise of the branch current I_{B3} and I_{ex} is modeled by (2.84). As a result of splitting the base currents of the NPN and the parasitic PNP in the Mextram model, the flicker noise from the area contribution of I_{B1} is modeled as

$$S_i = \frac{\mathtt{KF}}{f}(1-\mathtt{XIBI})\left(\frac{|I_{B1}|}{1-\mathtt{XIBI}}\right)^{\mathtt{AF}}, \tag{2.85}$$

and the flicker noise from the sidewall contribution (I^S_{B1}) is modeled as

$$S_i = \frac{\mathtt{KF}}{f}\mathtt{XIBI}\left(\frac{|I^S_{B1}|}{\mathtt{XIBI}}\right)^{\mathtt{AF}}. \tag{2.86}$$

However, the flicker noise from I_{B2} [70] is slightly different from the ideal one. It is modeled as

$$S_i = \frac{\texttt{KFN}}{f}|I_{B2}|^{2(\texttt{MLF}-1)+\texttt{AF}(2-\texttt{MLF})}\triangle f \tag{2.87}$$

to take MLF into account. In addition, when extended modeling of reverse current gain (EXMOD=1) is considered, the flicker noise from I_{ex} and XI_{ex} are modeled as

$$S_i = \frac{\texttt{KF}}{f}(1-\texttt{XEXT})\left(\frac{|I_{ex}|}{1-\texttt{XEXT}}\right)^{\texttt{AF}}, \tag{2.88}$$

and

$$S_i = \frac{\texttt{KF}}{f}\texttt{XEXT}\left(\frac{|XI_{ex}|}{\texttt{XEXT}}\right)^{\texttt{AF}}. \tag{2.89}$$

There is also excess noise due to avalanche effect [71, 72] in the Mextram model. The extra noise sources resulting from the avalanche effect are split and added to I_{B1} and I_N. So, the noise currents in parallel with I_{B1} and I_N are increased and correlated. The I_N excess noise spectral density($S_{i_{Nex}}$), the I_{B1} excess noise spectral density($S_{i_{B1ex}}$) and their correlated noise spectral density ($S_{i_{B1ex}i_{Nex}}$) are represented as [31]:

$$S_{i_{Nex}} = \texttt{KAVL}\cdot 2qI_{avl}(3+2G_{EM}), \tag{2.90}$$

$$S_{i_{B1ex}} = \texttt{KAVL}\cdot 2qI_{avl}(1+2G_{EM}), \tag{2.91}$$

$$S_{i_{B1ex}i_{Nex}} = -\texttt{KAVL}\cdot 2qI_{avl}(2+2G_{EM}). \tag{2.92}$$

Where KAVL is the model parameter to switch on/off the excess noise, I_{avl} is the avalanche current and "G_{EM}" is the multiplication factor minus one (M-1).

2.8 Conclusion

Mextram model covers all the important effects in modern bipolar devices including:

1. Bias-dependent Early effect;
2. Low-level non-ideal base currents;
3. High-injection effects;
4. Ohmic resistance of epilayer;
5. Velocity saturation effects on the resistance of the epilayer;
6. Hard and quasi-saturation (including non-ohmic quasi-saturation effect);
7. Weak avalanche (optionally including snap-back behavior);
8. Charge storage effects;
9. Split base-collector and base-emitter depletion capacitance;
10. Substrate effects and parasitic PNP;
11. Explicit modelling of inactive regions;
12. Current crowding and conductivity modulation of the base resistance;
13. First order approximation of the distributed high frequency effects in the intrinsic base (high-frequency current crowding and excess phase-shift);
14. Recombination in the base (meant for SiGe transistors);
15. Early effect in the case of graded bandgap (meant for SiGe transistors);
16. Temperature scaling;
17. Self-heating;
18. Thermal noise, shot noise and $1/f$ noise,

which meet various circuit design (ECL, mixer, VCO, LNA, PA...) requirements.

In addition, it is elected by the members of the Compact Modelling Council as the bipolar standard model along with the HICUM model. The latest Mextram model level 504.6 parameters are listed in Appendix A.

Chapter 3

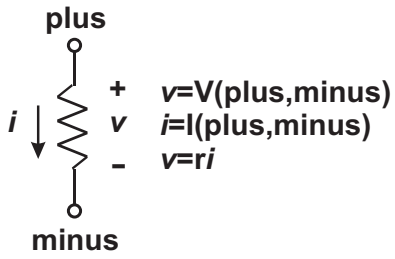
Model implementation and verification

3.1 Introduction

The path from compact model development to the implementation into a commercial circuit simulator is often time consuming. Moreover, it is not always straightforward how to implement behavioral models in SPICE-like simulators. In this chapter, the capability of the analog hardware description language (AHDL) Verilog-A [73] to handle a state-of-the-art compact bipolar transistor model mixed with a behavioral substrate coupling model is described. Verilog-A is a high-level language developed to describe the structure and behavior of analog systems and their components. It is an extension to the IEEE 1364 Verilog HDL specification for the digital design. The analog systems are described in Verilog-A in a modular way using hierarchy and different levels of model complexity. The motivation is to move into a higher level of abstraction in analog design and a combination with the digital design.

The basic programming unit for the structural and behavioral description of analog systems in Verilog-A is a **module**. The analog system structure is defined through the module's input and output signals and their connections. On the other hand, a sequence of mathematical equations is employed at the core of the module to describe its behavior. Here is an example of how a simple resistor is implemented in Verilog-A:

```
//Resistor model in Verilog-A
'include "disciplines.vams"
module resistor(plus,minus);
    parameter real r=0;
    inout plus,minus;
    electrical plus,minus;
    analog
        V(plus,minus)<+r*I(plus,minus);
endmodule
```



The first line started with "//" is a comment line. Followed by including the "disciplines.vams" which is a collection of related physical signal types. The model statement starts from the keyword **module** followed by the name of the module "resistor" and a list of its ports within the parentheses (plus, minus). It is possible to control the model equations by a set of parameters that can be passed to the module at the moment of its instantiation into the analog system. As a result, the resistance value is defined as **parameter** with a default value of zero. The direction of the ports are defined as bi-directional by **inout** and the types of ports are defined as **electrical**, which means the signals applied to the ports are expected to be voltage and current. The keyword **analog** indicates continuous time behavior. "V(plus, minus)" and "I(plus, minus)" are the voltage and current across the resistor. The voltage and the product of current and resistance "r" are forced to be equal, which forms a linear resistor model, through the contribution statement "< +". Finally, the description ends with the keyword **endmodule**. With the features introduced above and many others in the Verilog-A language, it represents an excellent environment for rapid development and verification of compact and behavioral modelling ideas in the commercial circuit simulators.

3.2 Compact model implementation and testing of Mextram

In order to practically verify capabilities of Verilog-A to serve as a framework for mixed compact and behavioral model developments we have implemented the bipolar transistor compact model Mextram Level 504 directly following the model description based on the NXP Semiconductors' documentation [40]. The Verilog-A Mextram 504 is implemented as follows:

```
‘include "frontdef.inc" //The useful functions, numerical and physical constants are included.

‘define SELFHEATING //The self-heating effect is set.
‘define SUBSTRATE //The substrate node is set.

module bjt504t_va (c, b, e, s, dt);
    inout c, b, e, s, dt; //The directions of external ports are defined.
    electrical c, b, e, s, dt; //The types of external ports are defined.
    electrical c1, c2, b1, b2, e1; //The internal nodes are defined.
    electrical noi; //The node for correlated noise is defined .
    ‘include "parameters.inc" //The model parameters are included.
    ‘include "variables.inc" //The variables are included.
    analog begin
        ‘include "initialize.inc" //The initial model constants are included.
        ‘include "tscaling.inc"//The temperature scaling rules are included.
        ‘include "evaluate.inc" //The model equations are included.
    end
endmodule
```

According to the implementation above, the model is separated into many small modules based on their functionality (Ex: model parameters are added in the module "parameters.inc") and there is a comment after each of them. So, it will be easier to update and maintain the model in the future. By including all the small modules, the description of the Mextram model in Verilog-A language is accomplished. The model has been tested using the ADS [74] circuit simulator equipped with the Tiburon's Verilog-A compiler [75]. Perhaps the most important question related to the Verilog-A implementation of Mextram 504 is about the achieved accuracy of the simulated electrical characteristics and the simulation speed. To this end, the C-coded NXP Semiconductors' SiMKit 2.3 implementation of Mextram 504 in the Agilent circuit simulator ADS has been used as a reference for comparison. As a measure of the discrepancy between the two model implementations, we have considered the percentage of relative error ((SiMKit-Verilog-A)/SiMKit*100%) of the simulated electrical characteristics. Since only the computational accuracy and the simulation speed have been analyzed, the comparison has been based on the default values of the model parameters [76] and the standard setups for

Mextram model parameter extraction listed below.
a) Junction capacitances (C_{BE} v.s. V_{BE}, C_{BC} v.s. V_{BC} and C_{SC} v.s. V_{SC});
b) Forward and reverse Early voltages (Vef v.s. V_{CB} and Ver v.s. V_{EB});
c) Forward and Reverse Gummel characteristics (I_C, I_B v.s. V_{BE} and I_E, I_B, I_S v.s. V_{BC});
d) Output characteristics (I_C v.s. V_{CE});
e) Cut-off frequency(f_T v.s. I_C);
f) Flicker noise (S_{I_C} v.s. Freq.);
g) Transient simulation with step input voltage in the base(I_C v.s. Time);
h) Harmonic Balance simulation matched at constant gain circle of 15 dB (Pout1, Pout3 and Pout5 v.s. Pin).

The model has been tested at temp. = -50, 25, 150 ^oC. Fig. 3.1 $\sim$ Fig. 3.9 show the comparisons of the various electrical characteristics obtained by SiMKit Mextram 504 (SiMKit solid lines) implementation, the Verilog-A Mextram 504 (VA dot lines) implementation, and the corresponding relative error at 25 ^oC. It has been observed that in most cases the relative error of the electrical characteristics used in our comparisons is quite low with the worst case of $\approx 0.1\%$. The source of $< 0.1\%$ discrepancies could be the numerical accuracy of variables and functions within the Verilog-A compiler (in comparison to C double precision variables and functions), and the order in which the expressions are executed (especially in the symbolic evaluation of Jacobian derivatives).

Table 3.1: CPU time for model simulation at different setup.

Setup	SiMKit (s)	Verilog-A (s)
a	0.88/0.86/0.82	0.85/0.85/0.85
b	0.76/0.76	0.80/0.77
c	0.77/0.76	0.78/0.77
d	0.84	0.82
e	2.06	2.36
f	0.83	0.85
g	0.85	0.86
h	1.05	1.05

The simulation time of each test setup from SiMKit and Verilog-A models is listed in Table 3.1. As can be seen from Table 3.1, the simulation time of each test setup from Verilog-A is compatible to SiMKit and some of the setups

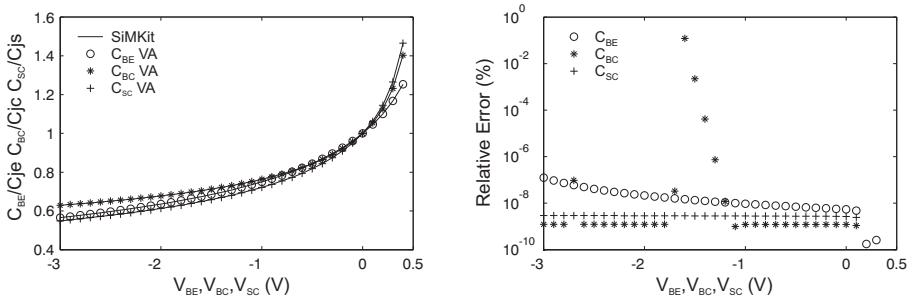


Figure 3.1: Junction capacitance normalized to its zero bias value and relative errors.

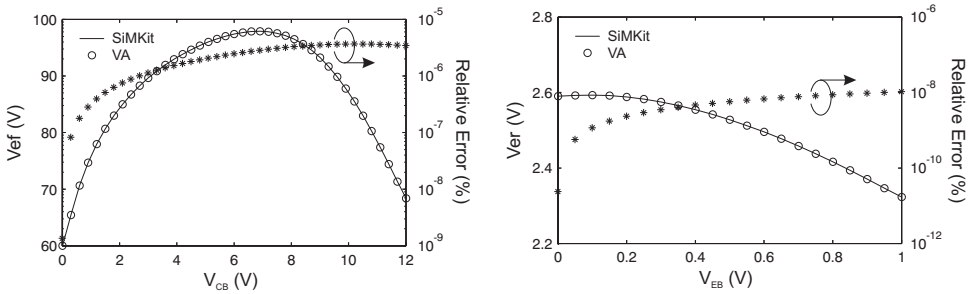


Figure 3.2: Forward and reverse Early voltage and relative errors.

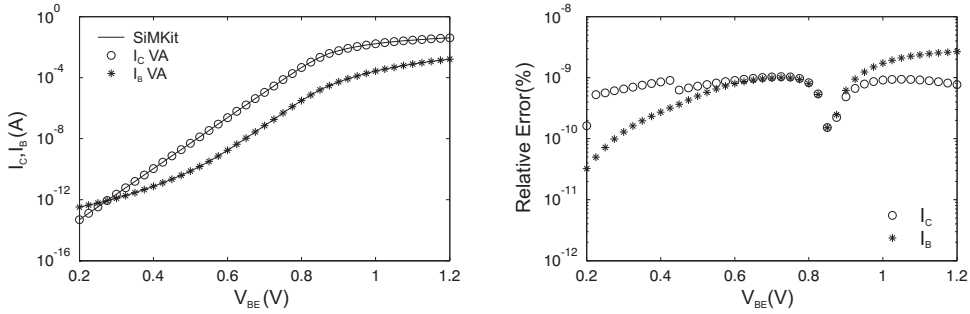


Figure 3.3: Forward Gummel characteristics and relative errors.

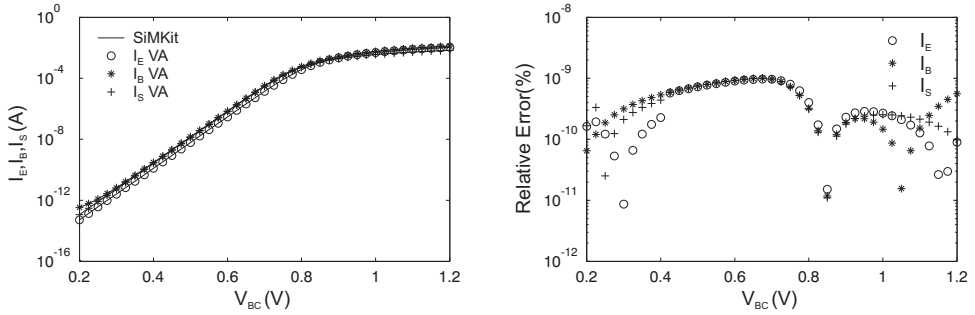


Figure 3.4: Reverse Gummel characteristics and relative errors.

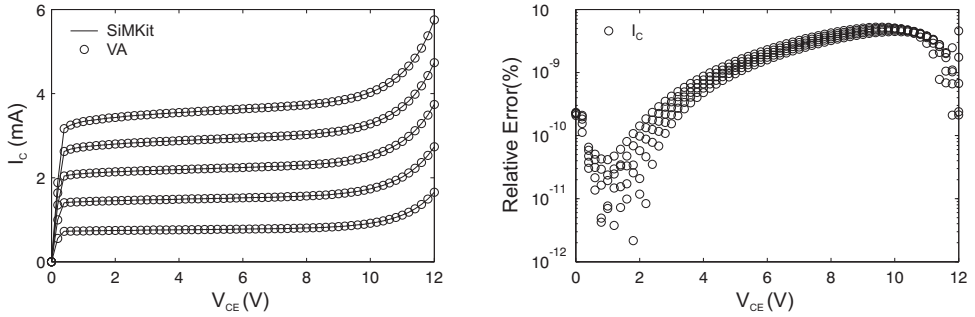


Figure 3.5: Output characteristics biased at I_B = 5, 10, 15, 20, 25 μA and relative errors.

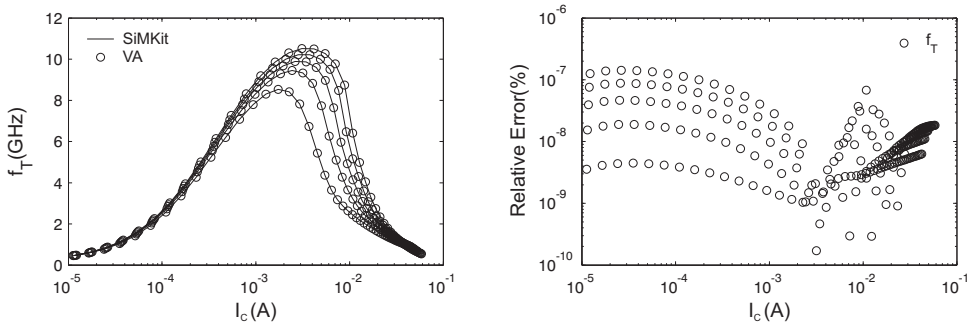


Figure 3.6: f_T vs. I_C biased at V_{CE} = 2, 4, 6, 8, 10 V and relative errors.

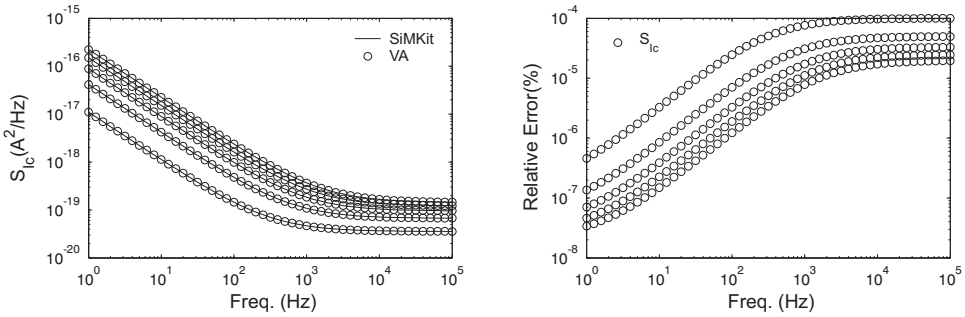


Figure 3.7: Flicker noise simulation and relative errors.

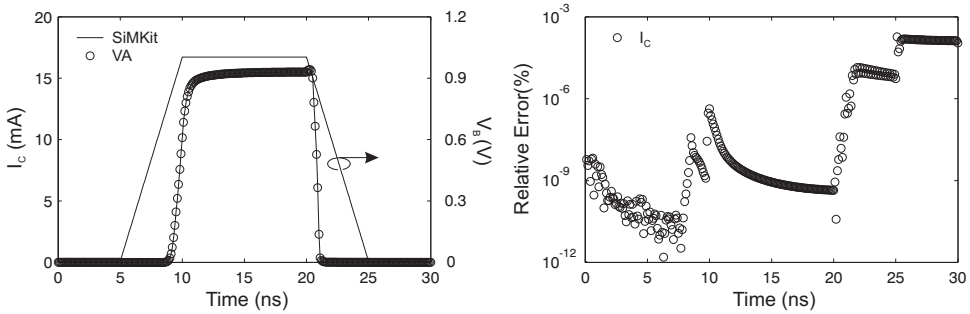


Figure 3.8: Transient simulation and relative errors.

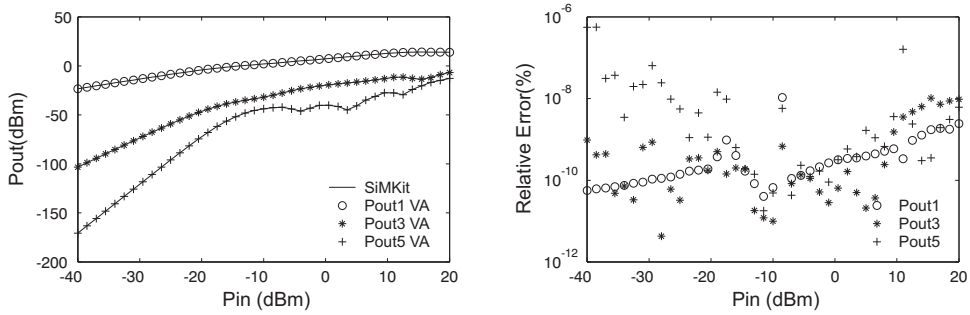


Figure 3.9: Harmonic Balance simulation and relative errors.

are even faster than SiMKit. It is different from what we had assumed, i.e. the C-coded model should be faster than the Verilog-A model.

3.3 Base-emitter tunneling current

With the progress of the bipolar process, a very thin SiGe Epi-base is feasible in the advanced bipolar device. With that the base doping concentration had to be increased to reduce the base resistance and to avoid punch-through from the BC to the BE junction. As a result of the high base and emitter doping concentrations, the Fermi level in the P-doped base and in the N-doped emitter are close to the valance and conduction bands, respectively. When the BE junction is reverse-biased, the BE tunneling current becomes significant due to the overlap between valance band in the base and conduction band in the emitter. The BE tunneling current model is not in the official release of the Mextram model. It is currently proposed as an extension of the model and as such under review within the Compact Modeling Council.

3.3.1 Physical model

The tunneling of an electron through a forbidden band is formally the same as tunneling through a potential barrier, which can be evaluated analytically using the classical Wentzel-Kramers-Brillouin (WKB) approximation. Taking into account the conservation of the perpendicular electron momentum and finite occupancy of the states in valence and conduction band, the tunneling current can be expressed as [77]

$$I_{BEt}(V_j) = A_E \alpha q D(V_j) \left(E_g^{-1/2} F_m(V_j) \right)^{\sigma} \exp\left(-\beta \frac{E_g^{3/2}}{F_m(V_j)} \right), \tag{3.1}$$

where V_j is the applied junction voltage, A_E is the effective PN junction area, α and β are material constants (considered here as temperature independent), E_g is the semiconductor energy band-gap, $F_m(V_j)$ is the bias dependent maximum electric field intensity at the PN junction, σ is the model constant having values $\sigma = 1$ for direct and $\sigma = 3/2$ for indirect tunneling mechanism and $D(V_j)$ is the bias dependent state occupancy function.

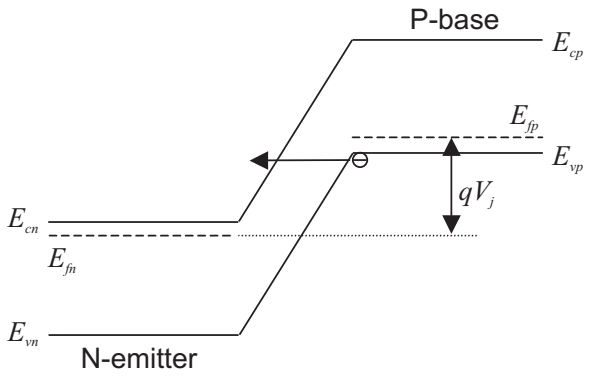


Figure 3.10: Band diagram of a PN junction at reverse bias

The state occupancy function is calculated as

$$D(V_j) = \begin{cases} \int_{E_{cn}}^{E_{vp}} \left[f\left(E, E_{fp}(V_j)\right) - f\left(E, E_{fn}(V_j)\right) \right] dE & \text{if } E_{vp} > E_{cn}, \\ 0 & \text{if } E_{vp} \leq E_{cn} \end{cases} \tag{3.2}$$

where E_{cn} and E_{vp} are energies corresponding to the bottom of the conduction on the N-side and the top of the valence band on the P-side of the PN junction, respectively. E_{fp} and E_{fn} are the corresponding quasi-Fermi levels as shown in Fig. 3.10. The term

$$f(E, E_f) = \frac{1}{\exp[(E - E_f)/kT] + 1} \tag{3.3}$$

is the Fermi-Dirac distribution function. In (3.2) it is assumed that the occupancy function and the tunneling current do not exist if there is no overlap between the P-valence and N-conduction bands. In principle this assumption should be softened for the indirect, phonon assisted tunneling but for simplicity it is kept in the current description and implementation of the tunneling current model. Note from (3.2) that for degenerate semiconductors, with $E_{fn} > E_{cn}$ or $E_{fp} < E_{vp}$, the tunneling current could be present also in the forward biased PN junction (when $E_{fn} > E_{fp}$). On the other hand, if both sides of the PN junction are nondegenerate, there would be essentially no tunneling current until the absolute value of the reverse applied voltage V_j is equal to sum of distances from the Fermi level to the nearest band edge in P and N region.

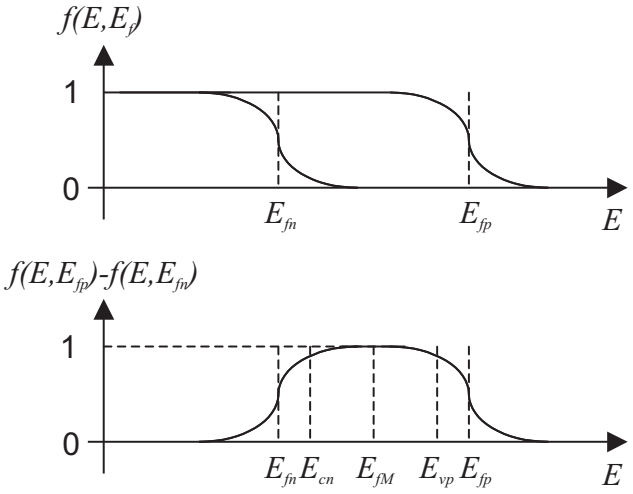


Figure 3.11: Fermi-Dirac distribution of N and P types semiconductors and their difference

The integrand of (3.2) for $V_j < 0$ has a maximum (or minimum if $V_j > 0$) at energy

$$E_{fM} = \frac{E_{fp}(V_j) + E_{fn}(V_j)}{2} \tag{3.4}$$

in Fig. 3.11. The eqn.(3.2) can be approximated as the product of the integral interval and the integrand evaluated at $E = E_{fM}$, resulting in

$$D(V_j) = \max\left(E_{vp} - E_{cn}, 0\right)\left[f\left(E_{fM}, E_{fp}(V_j)\right) - f\left(E_{fM}, E_{fn}(V_j)\right)\right]. \tag{3.5}$$

By introducing

$$E_{vp} - E_{cn} = E_{fp} - E_{fn} + q\Delta V_{jt}, \tag{3.6}$$

$$E_{fp} - E_{fn} = -qV_j, \tag{3.7}$$

the state occupancy function (3.5) can be expressed as

$$D(V_j) = q\min\left(V_j - \Delta V_{jt}, 0\right) \cdot \frac{1 - \exp\left(\dfrac{V_j}{2V_T}\right)}{1 + \exp\left(\dfrac{V_j}{2V_T}\right)}, \tag{3.8}$$

where $V_T = kT/q$ is the thermal voltage. Note, that if we approximate the Fermi functions by the discontinuous step function, which is equivalent to

evaluating (3.8) for $T = 0$, and the positions of the Fermi levels E_{fn} and E_{fp} coincide with the valence and conduction bands, respectively, (3.8) reduces to

$$D_0(V_j) = E_{fp}(V_j) - E_{fn}(V_j) = q \min(V_j, 0). \tag{3.9}$$

This term is commonly used in compact modeling the PN junction tunneling current [78, 79]. Although the step discontinuous Fermi function could be a reasonable assumption as long as $|V_j| \gg V_T$, it introduces a discontinuous tunneling conductance $(\partial I_{BEt}/\partial V_j)$ at $V_j = 0$ since $I_{BEt} = 0$ for $V_j \geq 0$. Moreover, it does not provide physical background for the forward PN junction tunneling effect.

The maximum junction electric field

The base-emitter tunneling current (3.1) essentially depends on the bias dependent maximum intensity of the electric field $F_m(V_j)$ across the PN junction. From the Poisson equation, it can indirectly be expressed in terms of the total junction depletion charge per unit area $Q_d(V_j)$ as

$$F_m(V_j) = \frac{Q_d(V_j)}{\epsilon_s}, \tag{3.10}$$

where ϵ_s is the permittivity of the semiconductor. The total junction depletion charge per unit area can be expressed as

$$Q_d(V_j) = Q_d(0) - \int_0^{V_j} C_j(V)\, dV, \tag{3.11}$$

where $C_j(V_j)$ is the junction depletion capacitance per unit area. The analytical solution of the Poisson equation for abrupt and linear PN junctions under depletion approximation for the zero bias total depletion charge and depletion capacitance are

$$Q_d(0) = \frac{C_{j0} V_d}{1-p} \quad \text{and} \quad C_j(V_j) = \frac{C_{j0}}{\left(1 - \dfrac{V_j}{V_d}\right)^p}, \tag{3.12}$$

where C_{j0} is the zero-bias value of the depletion capacitance per unit area and V_d is the built-in voltage. The grading coefficient p has the values 1/2 and 1/3 for abrupt and linear junction, respectively. Since abrupt and linear PN

junctions are the two extreme cases, it is reasonable to assume that expression (3.12) could be applicable for arbitrary PN junctions. One should be aware that due to the total depletion approximation the validity of (3.12) is basically limited to the reverse (and weakly positive) junction biases.

From (3.11) and (3.12) the total junction depletion charge is

$$Q_d(V_j) = \frac{C_{j0}V_d}{1-p}\left(1-\frac{V_j}{V_d}\right)^{1-p} \tag{3.13}$$

and the maximum electric field $F_m(V_j)$ in (3.10) becomes

$$F_m(V_j) = \frac{C_{j0}V_q(V_j)}{\epsilon_s}, \tag{3.14}$$

where

$$V_q(V_j) = \frac{V_d}{1-p}\left(1-\frac{V_j}{V_d}\right)^{1-p} \tag{3.15}$$

is *the effective bias of the total junction depletion charge*.

In bipolar transistor compact modeling, the depletion charges are also important for the description of the Early effect. However, to this end depletion charges are introduced relative to the zero bias total depletion charge as

$$Q_t(V_j) = Q_d(0) - Q_d(V_j) = C_{j0}V_t(V_j), \tag{3.16}$$

where

$$V_t(V_j) = \frac{V_d}{1-p}\left[1-\left(1-\frac{V_j}{V_d}\right)^{1-p}\right] \tag{3.17}$$

is *the effective junction bias of the relative depletion charge*. The total and relative depletion charge effective junction biases are related as

$$V_q(V_j) = \frac{V_d}{1-p} - V_t(V_j). \tag{3.18}$$

Since the effective junction bias of the relative depletion charge V_t is already available in the Mextram modeling implementation [40], it is better to employ (3.18) instead of (3.15) since it would require only one additional substraction (considering that the term $V_d/(1-p)$ is also already evaluated and available).

3.3.2 Temperature dependence

The tunneling nature of the measured currents can be verified by measuring its temperature dependence. Tunneling currents are typically identified by a small positive temperature dependence. It is mainly due to the temperature dependence of the bandgap E_g as well as temperature dependence of the maximum intensity of the junction electric field F_m.

The variation of the bandgap with temperature can be expressed approximately by the universal function

$$E_g(T) = E_g(0) - \frac{aT^2}{T+b}, \tag{3.19}$$

where a and b are material constants.

The temperature dependence of the maximum intensity of the junction electric field $F_m(T)$ is indirectly defined via temperature dependence of the diffusion voltage $V_d(T)$ and zero bias depletion capacitance $C_{j0}(T)$, which are readily available in standard model formulations.

3.3.3 Model implementation

The implementation of the BE tunneling current model is based on (3.1), (3.8) and (3.18) with $\Delta V_{jt} = 0$ and $\sigma = 1$. Therefore, it is implemented in the Mextram model as a current source between nodes $B2$ and $E1$ in Fig. 2.1 as

$$I_{BEt} = \begin{cases} \mathrm{BTJE_{TM}} \cdot V_{jE}^{\mathrm{eff}} \cdot V_{qE} \exp\left(-\dfrac{\mathrm{VTJE_T}}{V_{qE}}\right) & \text{if } V_{B2E1} < 0 \\ 0 & \text{if } V_{B2,E1} \geq 0 \end{cases} \tag{3.20}$$

with

$$V_{qE} = \frac{\mathrm{VDE_T}}{1-\mathrm{PE}} - V_{tE} \tag{3.21}$$

and

$$V_{jE}^{\mathrm{eff}} = \frac{1-c}{1+c} V_{B2E1}, \tag{3.22}$$

where

$$c = \exp\left(\frac{V_{B2E1}}{2V_T}\right). \tag{3.23}$$

Note, that the evaluation of V_{tE}, $\mathrm{VDE_T}/(1-\mathrm{PE})$ and $\exp(V_{B2E1}/V_T)$ is already part of the Mextram model evaluation procedure. The additional effort in

Table 3.2: The additional effort for the tunneling current model evaluation

+(−)	/	∗	exp	sqrt	if
3	2	5	1	1	1

terms of arithmetic operations and elementary functions for the evaluation of the tunneling current is given in Table 3.2. The scaling of the parameters is performed as

$$\mathrm{BTJE_T} = \frac{q}{\epsilon_s} A_E \alpha C_{j0}(T) E_g(T)^{-1/2}, \tag{3.24}$$

$$\mathrm{VTJE_T} = \beta \epsilon_s \frac{E_g(T)^{3/2}}{C_{j0}(T)}, \tag{3.25}$$

$$\mathrm{BTJE_{TM}} = \mathrm{BTJE_T} \cdot \mathrm{MULT}. \tag{3.26}$$

Where the quantities $\mathrm{BTJE_{TM}}$ and $\mathrm{VTJE_T}$ represent temperature and `MULT` scaled parameters. Introducing the model parameters `BTJE` and `VTJE` as $\mathrm{BTJE_T}$ and $\mathrm{VTJE_T}$ evaluated at $T = T_{\mathrm{ref}}$ we have

$$\mathrm{BTJE_T} = \frac{C_{j0}(T)}{C_{j0}(T_{\mathrm{ref}})} \left(\frac{E_g(T)}{E_g(T_{\mathrm{ref}})} \right)^{-1/2} \cdot \texttt{BTJE} \tag{3.27}$$

$$\mathrm{VTJE_T} = \left(\frac{C_{j0}(T)}{C_{j0}(T_{\mathrm{ref}})} \right)^{-1} \left(\frac{E_g(T)}{E_g(T_{\mathrm{ref}})} \right)^{3/2} \cdot \texttt{VTJE} \tag{3.28}$$

As a result of small temperature dependence, the band-gap temperature dependence is linearized around reference temperature T_{ref} and it is implemented as

$$\left(\frac{E_g(T)}{E_g(T_{\mathrm{ref}})} \right)^{-1/2} \simeq 1 + \texttt{ABTJE} \left(\frac{T}{T_{ref}} - 1 \right), \tag{3.29}$$

$$\left(\frac{E_g(T)}{E_g(T_{\mathrm{ref}})} \right)^{3/2} \simeq 1 + \texttt{AVTJE} \left(\frac{T}{T_{ref}} - 1 \right), \tag{3.30}$$

where `ABTJE` and `AVTJE` are the temperature parameters for `BTJE` and `VTJE`.

The tunneling current is also added to the device power dissipation as

$$P_{diss} = ... + I_{BEt} V_{B2E1} + ... \tag{3.31}$$

Finally, the description, units, default values and range of the new model parameters are given in Table 3.3. The Verilog-A model implementation is listed in Appendix B.

Table 3.3: Model parameter for Base-Emitter Band-to-Band Tunneling Current

Name	Description	Unit	Default	Range
BTJE	Pre-factor of the base-emitter tunneling current	[A/V^2]	0.0	$[0.0, \infty)$
VTJE	Exponential coefficient for the base-emitter tunneling current	[V]	0.0	$[0.0, \infty)$
ABTJE	Temperature coefficient of the BTJE parameter		0.0	
AVTJE	Temperature coefficient of the VTJE parameter		0.0	

3.3.4 Tunneling current parameter extraction

The parameters BTJE and VTJE can be extracted from the base current dependence with the reverse biasing of the BE junction either in standard reverse-Early measurements or separately performed off-state measurements. In the case of standard reverse-Early measurements it is important to first de-embed the part of the base current that is injected at the base-collector junction.

It should be also emphasized that in the lower bias range the SRH recombination current could be still dominant in comparison to the tunneling current and this region should not be used for parameter extraction. One way to avoid the low-bias region in the parameter extraction is to employ the functional relationship

$$\log\left(\frac{I_{BEt}}{V_{jE}V_{qE}}\right) = \log\left(\mathtt{BTJE}\right) - \frac{\mathtt{VTJE}}{V_{qE}} \tag{3.32}$$

with model parameters corresponding to the reference temperature and MULT = 1. It implies linear relationship of $I_{BEt}/(V_{jE}V_{qE})$ on $1/V_{qE}$ in the logarithmic scale. If the SRH recombination current is present for low V_{EB} it will be easily observed as deviation from the this linear relationship.

Fig. 3.12 shows an example of the parameter extraction procedure. The symbols in Figure 3.12 (a) are measured $I_{BEt}/(V_{jE}V_{qE})$ vs. $1/V_{qE}$ relationships for 9 different temperatures (-40, -20, 0, 25, 50, 75, 100, 125 and 150 °C).

The model parameters are directly, without nonlinear optimization procedures, extracted from the straight line segments in the area of higher reverse

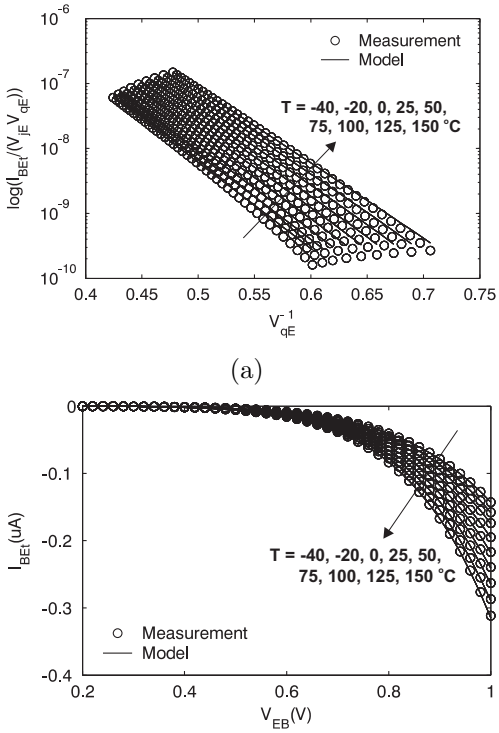


Figure 3.12: (a)parameter extraction and (b) model fit of the tunneling current.

BE biasing. The resulting plot of measurement and modelled tunneling current from different temperature is shown in Figure 3.12 (b).

3.4 Behavior modeling of substrate effect

For most of the bipolar transistor compact models, such as Mextram, substrate effects are not included in the model, or, such as in HICUM [28], it uses a simple R-C network to describe the substrate characteristic. In reality, bipolar transistors are made on top of the silicon substrate. When a bipolar device operates at radio frequencies (RF), the lossy Si substrate becomes a *distributed substrate network* (Z_{sub}) connected to the intrinsic device as shown in Fig. 3.13a. The Z_{sub} will be changed if the substrate contact of the bipolar device is changed. As a matter of fact, the substrate contact in the

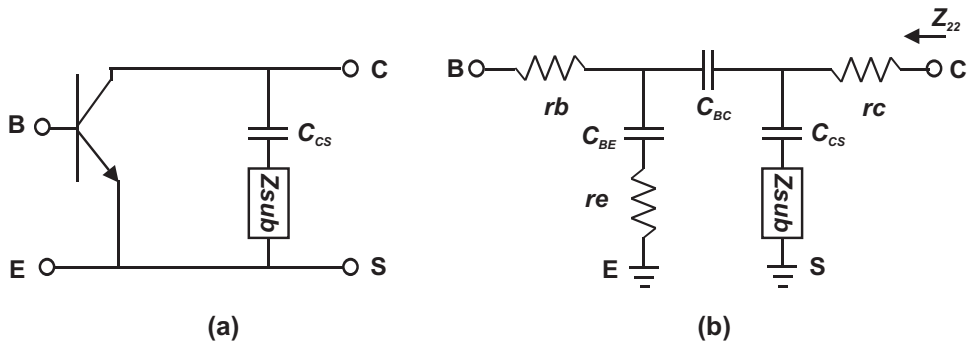


Figure 3.13: (a) A bipolar transistor with an embedded substrate network (Z_{sub}) and (b) its off-state equivalent circuit

final circuit design may be different from the one used for the test devices, from which the model parameter is extracted. That is the reason why the Mextram model doesn't include substrate as a part of the device model. But for some bipolar processes, the transistor's substrate contact is well defined within the p-cell. As a result, a lumped elements substrate network [80, 81] extracted from the 3D device simulation can be added to the intrinsic model as a sub-circuit to reduce designer's burden in determining the complicated Z_{sub}. However, it is not generally valid for all the bipolar processes. In fact, it depends on the topology of that specific process.

In order to account for the substrate effect on the single transistor characteristics at ultra wide frequency region, the compact model can be used for the intrinsic transistor and a behavior model can be used for the Z_{sub} as a generic substrate modeling approach. Therefore, accurate modeling of the Z_{sub} will be of great value to high frequency circuit design. Since the present Mextram model release does not include a substrate model, it is added here in the form of a behavioral model based on the Laplace transfer function.

Combining the off-state S-parameter measurement [82] of the device as shown in Fig. 3.13a and its off-state small-signal equivalent circuit as shown in Fig. 3.13b, the Z_{sub} as a function of frequency in Fig. 3.13b can be obtained from the measured Z_{22} over frequency converted from the measured off-state S-parameters (See Appendix C for the conversion table). As a result, the measured Z_{sub} extracted from the measured Z_{22} and Fig. 3.13b will be:

$$Z_{sub}(s) = \frac{(Z_{22} - rc)(1 + sC_{bc}Z_1)}{1 + sC_{bc}(Z_1 - Z_{22} + rc)} - \frac{1}{sC_{cs}}, \tag{3.33}$$

where $Z_1(s) = re + \frac{1}{sC_{be}}$. For the purpose of behavioral modeling, the Z_{sub} is considered as a 1-port functional block associated with potential and current. In the Laplace domain, the relationship between the input excitation and the corresponding response can be described by the rational complex function as

$$Z(s) = \frac{\sum_{k=1}^{n} N_{n-k} s^{n-k}}{s^n + \sum_{k=1}^{n} D_{n-k} s^{n-k}}. \tag{3.34}$$

Where "N_i" and "D_i" are model parameters. For a given set of frequency response data, $2n$ model parameters $((N_i, D_i), i = 0, \ldots, n-1)$ for the behavioral model are obtained by solving the linear least square problem for both real (Re) and imaginary (Im) parts of Z_{sub} and Z.

$$S(N_i, D_i) = \sum_{l=1}^{N_w} (Re, Im[Z_{sub}(jw_l)] - Re, Im[Z(jw_l)])^2, \tag{3.35}$$

where $S(N_i, D_i)$ is the sum of the squares of the deviations and N_w is the number of frequency samples. Fig. 3.14 shows the measured Z_{sub} of a SiGe HBT [83] from 100 MHz $\sim$ 100 GHz and the behavior model as a function of frequency after the least square curve fitting at n=6. In this particular case, the rational complex function fits the measured magnitude and phase very well. In Verilog-A, we can also easily implement this rational complex function with the Laplace transform analog operator "laplace_nd" [73] listed in Appendix D.

Since Y_{22} is the admittance parameter that is most sensitive to the substrate effect, Fig. 3.15 shows measured and simulated Y_{22} of the same SiGe HBT at four different bias conditions. As can be seen from Fig. 3.15, with the behavior model of substrate added to a intrinsic SiGe bipolar model, which has been extracted based on the standard Mextram parameter extraction procedures [76], the simulated Y_{22} fits the measured Y_{22} better than without it. It is because of the Verilog-A language that we can model such a complex distributed network in a very efficient way.

3.5 Discussion

One of the first observations and potential benefits of using Verilog-A in compact model development is that Verilog-A facilitates a quite "compact" representation of the compact models. The size of the Verilog-A Mextram 504

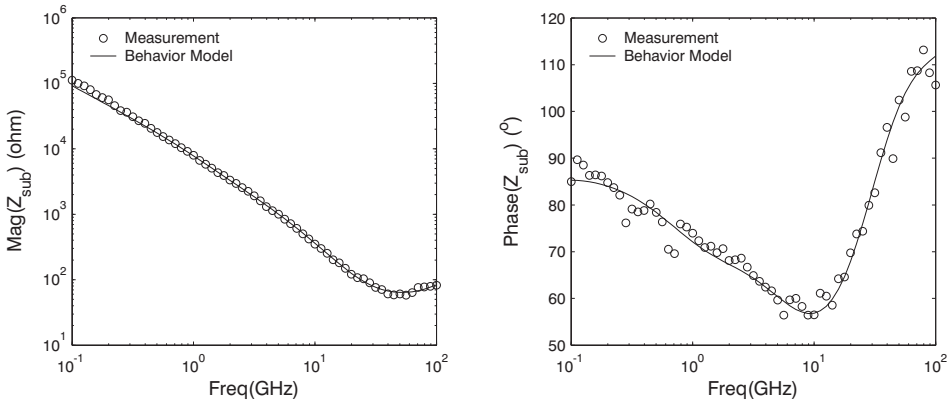


Figure 3.14: Magnitude and phase of substrate impedance vs. frequency.

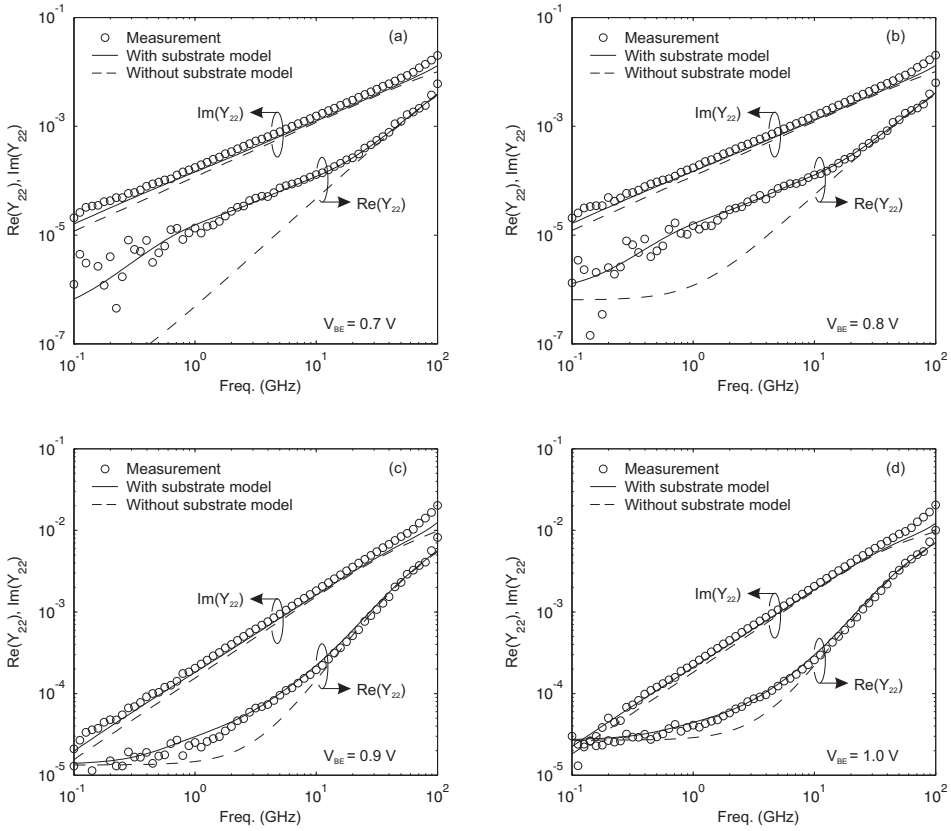


Figure 3.15: Measured and simulated Y_{22} biased at V_{BE} = (a) 0.7 V, (b) 0.8 V, (c) 0.9 V, d) 1.0 V and V_{CE} = 2.0 V v.s. frequency.

code is around 1200 lines compared to around 12000 lines in the source code of the SiMKit. It is mainly due to the fact that there is no need to program interfaces to the simulator [84] (concern of the particular circuit simulator with Verilog-A interface) and the derivatives of the electrical quantities (These are evaluated using symbolic derivation of the equation describing electrical signals in Verilog-A). The second one is that a Verilog-A model is easy to incorporated new features in it for evaluating new compact modeling ideas. However, the writer of the Verilog-A code is still fully responsible for the numerical stability of the governing model equations (smooth transitions, range of functions, etc.). A good programming practice and experiences from the C-code implementations could be very useful in achieving an effective Verilog-A code.

3.6 Conclusions

In this chapter, Mextram 504 model implementation in Verilog-A is introduced. It has been tested and shown to be equivalent to the SiMiKit 2.3 implementation from NXP Semiconductors. Furthermore, the CPU time of single device Verilog-A and SiMKit models simulated in ADS has been found to be compatible. However, implementation of Verilog-A has big advantages for the model developer such as no need to implement derivatives and interfaces to the simulator. Therefore, the general conclusion could be that Verilog-A indeed represents an effective environment for the fast evaluation and exchange of new compact modeling ideas such as BE tunneling, substrate coupling, thermal coupling network etc.

Chapter 4

Data acquisition and model parameter extraction

4.1 Introduction

An accurate compact device model, which can represent the semiconductor device characteristics in the CAD environment, depends on accurate measurements of DC, capacitance (CV), S-parameters and noise data for parameter extraction. This requires the design of suitable test structures for the measurement of DC, CV, S-parameters and noise data including dummy devices for pad parasitic removal (de-embedding) during high-frequency (HF) measurement.

4.2 On-wafer measurement for model parameter extraction

On-wafer measurement is preferred for semiconductor device modeling because of much lower parasitic resistances, capacitances and inductances comparing to measurement on a packaged device. It is usually done in a micro-chamber on top of a probe station since a micro-chamber provides an enclosed environment to shield the device-under-test (DUT) from environmental lights and noises. There are mechanical manipulators outside the micro-champer to manually move the chuck under the wafer in x-y-z directions to accurately position probes on a test device. It is also possible to automatically measured many

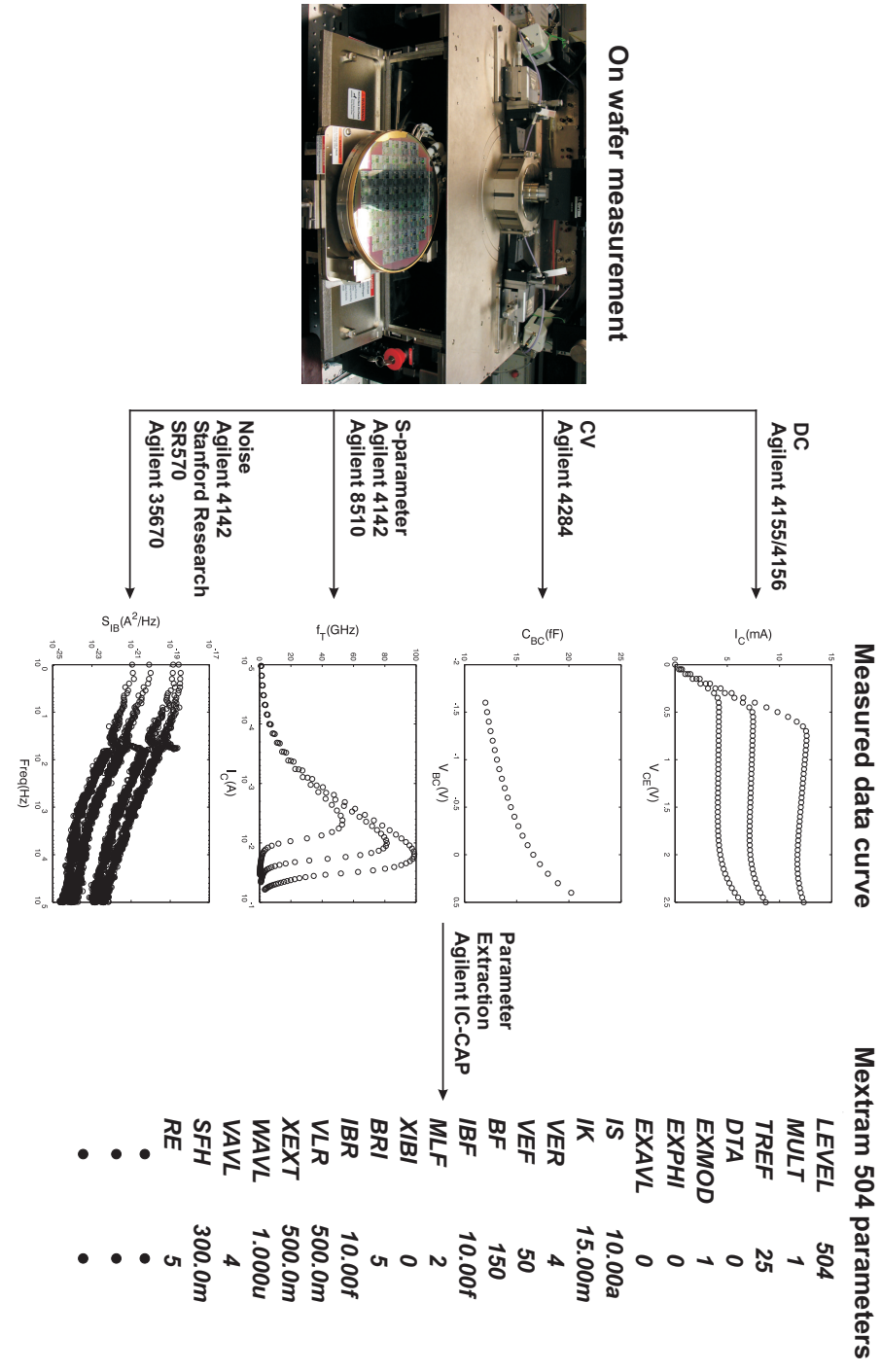


Figure 4.1: On-wafer measurement for model parameter extraction.

devices in a single die or in different dies of a wafer by setting indexes in a computer program, which controls the chuck movement through servomotors. The wafer chuck in the micro-champer serves for both fixing the position of the wafer and providing a controllable temperature. It is usually made of metal over an insulating material to have both high thermal conductivity and low leakage current down to pA level. The chuck temperature, which offers an ambient temperature to the wafer, is set by a controller allowing measurement from -40 to 200 °C. The complete flow from on-wafer measurement to model parameter extraction is illustrated in Fig. 4.1.

4.2.1 DC measurement

The DUT is exposed to a series of DC measurement at different biasing conditions to detect its DC characteristics. It is common practise to use voltages and currents testing and then monitor currents and voltages in characterization of bipolar transistors. The DC measurement reference plane is the DC analyzer [85] itself having sense and force features in each stimulus/measurement units (SMU). By using a Kelvin probe [86] equipped with a sense and a force probe tips, the reference plane can be moved from the DC analyzer to the contacted pad, i.e. closest to the DUT. This allows for excluding the cable parasitic and pad contact resistances, which can partly be larger than the device resistances. As a result, the measured resistance data can be very close to those of the intrinsic device.

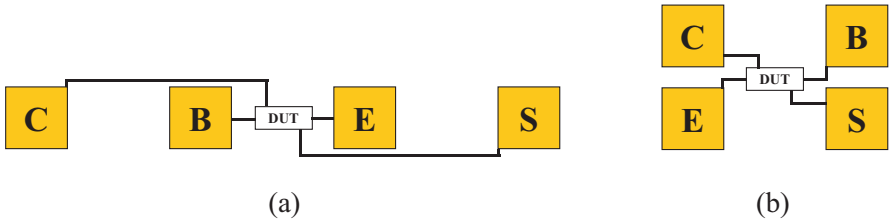


Figure 4.2: DC pads layout in (a) one testline and (b) two testlines.

The DUT on wafer is connected to DC pads with dimensions ranging typically between $50 \times 50 \sim 100 \times 100\ \mu m^2$ to ensure a reliable alignment of probes to device pads. The DC pads of a DUT can be arranged either in (a) a single test line or (b) in two test lines as shown in Fig. 4.2. Since the base and emitter resistances are usually more influential than collector and substrate resistances in the forward operating regime, it is better to have base

and emitter contacts as close to the pads as possible in order to minimize the parasitic resistances in the single test line layout. The DC pads arranged in two test lines layout have the advantage that microscopic alignment at high magnification will be easier. In addition, there is more space for putting many small transistors in parallel in two test lines if necessary. Therefore, two test lines layout style is recommended.

4.2.2 Capacitance measurement

CV measurement of the BE, BC and SC junction capacitances can be done by both using a low-frequency LCR meter [87] and DC pads or a high-frequency (HF) S-parameters measurement setup with an appropriate HF pad (see next subsection). The LCR meter test has the advantage of faster sweeping speed and smoother data output comparing to the S-parameters measurement at low frequency. Especially, when measuring a large amount of CV data in a production line, the throughput of a LCR meter cannot be matched by a S-parameters measurement.

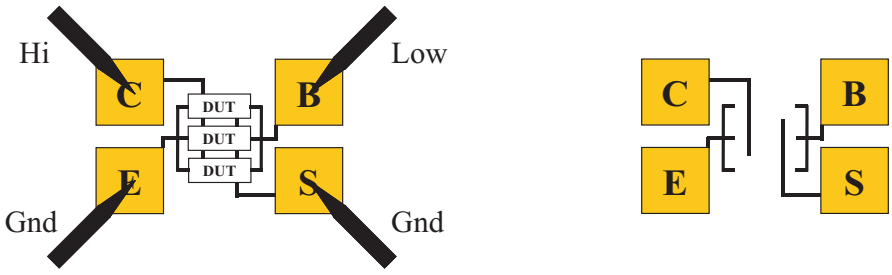


Figure 4.3: (a) 4-probe methods for measuring base-collector capacitance of transistors in parallel and its (b) dummy structure.

The typical resolution for the advanced LCR meter is in the order of 1 fF but it takes longer integration time to get smooth data when the measurement level is approaching the resolution limit. Therefore, it is common practise to measure many transistors in parallel as shown in Fig. 4.3a. This leads to an increased capacitance value and thus to better accuracy. Also, the capacitance of an open dummy pad shown in Fig. 4.3b for de-embedding the parasitic capacitance is then small in comparison to the measured capacitances of the transistor array in the DUT. The measured intrinsic junction

capacitance (C_{INT}) can be obtained from

$$C_{INT} = \frac{1}{N_{DUT}}(C_{DUT} - C_{open}), \tag{4.1}$$

where N_{DUT} is the number of devices in parallel, C_{DUT} is the measured capacitance from transistors in parallel and C_{open} is the measured capacitance from the open dummy structure.

It should be mentioned though that CV measurements can also be arranged by using a 4-point probe setup method [88] as also shown in Fig. 4.3a. The capacitances of the BE and BC junction capacitances can then be measured without the need of de-embedding the parasitic pad capacitances. Still required, however, is the de-embedding of the pads in the SC junction capacitance (C_{SC}) measurement to remove collector pad capacitance that couples to the substrate. As a result, it is advisable to provide dummy pad structures together with the DUTs on wafer in order to achieve accurate CV measurements.

4.2.3 S-parameters measurement

S-parameters (scattering parameters) are used in HF measurements for the following reasons [89]:

1. It is relatively easy to obtain high frequency test data from measuring traveling voltage waves with a Vector Network Analyzer (VNA) [90]. There is no need for open/short terminations that may cause the active device to oscillate or self-destruct;
2. The familiar measurements of power gain, loss and reflection coefficients in microwave applications can easily be applied;
3. S-parameters of multiple devices can be cascaded to predict the circuit performance;
4. S-parameters data can easily be transformed to H, Y, Z, ABCD-parameter (see Appendix C) if required.

Fig. 4.4 shows a schematic representation of S-parameters used in a 2-port measurement system with a characteristic impedance Z_0. The incident and reflected power to and from a DUT can be described by a linear 2-port as:

$$\begin{aligned} b_1 &= S_{11}a_1 + S_{12}a_2, \\ b_2 &= S_{21}a_1 + S_{22}a_2. \end{aligned} \tag{4.2}$$

Where a_1 and a_2 are incident voltage waves normalized by $\sqrt{Z_0}$ to port1 and port2, b_1 and b_2 are the reflected voltage waves normalized by $\sqrt{Z_0}$ from port1

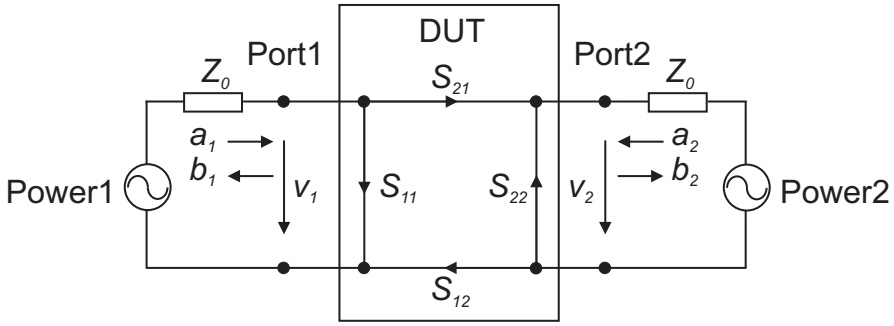
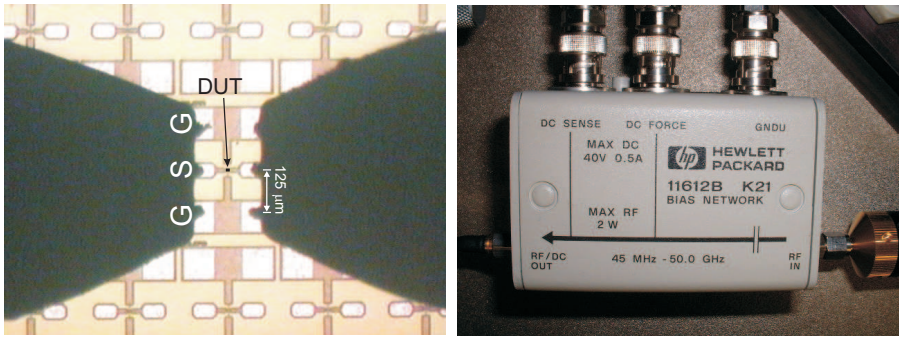


Figure 4.4: S-parameter in 2-port measurement.

and port2, respectively. The S-parameters in (4.2) can be determined by either port 1 ($a_1 = 0$) or port 2 ($a_2 = 0$) being terminated with Z_0.

On-wafer S-parameters measurements can be done by using HF Ground-Signal-Ground (GSG) or Ground-Signal (GS) probes [91] that are positioned onto the corresponding HF pads that are connected to the DUT (Fig. 4.5a). Since the field pattern in a GSG probe is symmetrical with respect to the electromagnetic field originating from the signal tip and terminating to the ground tip, this reduces undesirable fringing effects and resonances. With the GS probe, in contrast, the field pattern terminates into the ground tip on the one side, but terminates into the wafer ground plane or probe station wafer chuck on the other side. For that reason, the GSG probe has much higher bandwidth than the GS probe due to the symmetrical ground but it takes up more space on the wafer. For these reasons, GSG probes are used for HF measurement above 10 GHz when the area in a test chip is not of a major concern.

In order to measure S-parameters of a DUT at a certain DC bias, a bias network [92] (Fig. 4.5b) is placed between the HF probe and the VNA to combine the HF signal from VNA and the DC bias from the DC analyzer. The bias network behaves as a high-pass filter to allow the HF signal to pass through, while blocking the DC signal to the VNA. But the impedance of the combined port (probe-tip) of a bias network then becomes poorly matched at the crossover frequency of the bias network. High-gain transistors may become unstable and oscillate at the presence of these large reflections, particularly if they return with significant phase shifts. As a result, positioning the bias network close to probe head to reduce the length of cable between the bias network and the DUT will help to reduce unwanted bias oscillations if the



(a) (b)

Figure 4.5: (a) A GSG HF probe contact to a HF pad under microscope and (b) a bias network used in S-parameter measurement.

device is not unconditionally stable. Besides, some VNAs use internal bias networks, which are often more leaky. In such case, the behavior is equivalent to shunting a high resistivity resistor. It is thus helpful to use the less-leaky external bias network to avoid this problem, allowing for more accurate low-bias measurement of the transistor DC parameters through the DC analyzer.

De-embedding

One of major issues concerning the S-parameters measurement of the DUT is the de-embedding of the parasitic of a HF pad. As can be seen in Fig. 4.5a, the HF pad is much larger than the DUT. Calibration of the VNA on the well defined calibration standards [93, 94] moves the measurement reference plane from the VNA to the HF probe tips. To push the measurement reference plane from the calibrated probe tips further to the 2-port terminals of the DUT, a de-embedding of the HF pad parasitics is required. The first de-embedding technique introduced is the "Open" de-embedding method [95], which only takes the pad to ground capacitance (Y_1, Y_2) and pad to pad capacitance (Y_3) as shown in Fig. 4.6 into account. So, the Y-parameters of the DUT (Y^o_{dut}) obtained from the "Open" de-embedding method are

$$Y^o_{dut} = Y_{ext} - Y_{open}, \tag{4.3}$$

where Y_{ext} is the measured extrinsic Y-parameters and

$$Y_{open} = \begin{pmatrix} Y_1 + Y_3 & -Y_3 \\ -Y_3 & Y_2 + Y_3 \end{pmatrix}, \tag{4.4}$$

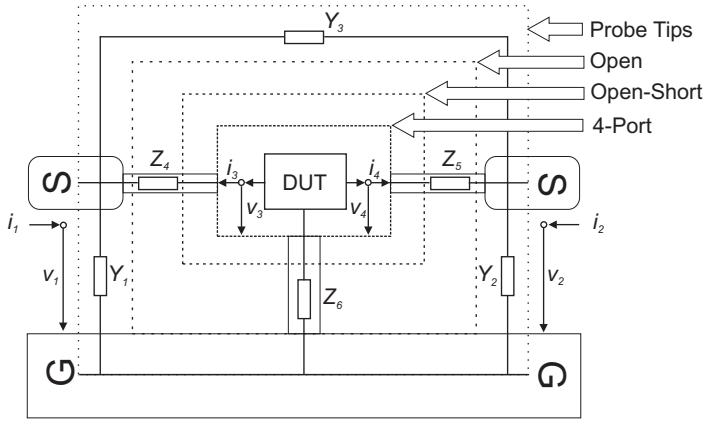


Figure 4.6: Moving of measurement reference plane from different de-embedding methods (Note: only half of the HF pad is shown in the figure due to symmetry of the pad).

which are the measured Y-parameters of an open dummy structure as shown in Fig. 4.7a. Later, an "Open-Short" de-embedding method [96], which assumes the HF pad to be an equivalent circuit of pad capacitances and series inductances (Z_4, Z_5 and Z_6) as shown in Fig. 4.6, was introduced. The Y-parameters of DUT (Y_{dut}^{os}) obtained from the "Open-Short" method become

$$Y_{dut}^{os} = \left((Y_{ext} - Y_{open})^{-1} - (Y_{short} - Y_{open})^{-1}\right)^{-1}, \tag{4.5}$$

where

$$Y_{short} = \begin{pmatrix} Z_4 + Z_6 & Z_6 \\ Z_6 & Z_5 + Z_6 \end{pmatrix}^{-1} + \begin{pmatrix} Y_1 + Y_3 & -Y_3 \\ -Y_3 & Y_2 + Y_3 \end{pmatrix}, \tag{4.6}$$

which are the measured Y-parameters of a short dummy structure in Fig. 4.7b. Though more involved lumped-element equivalent circuits had been developed for the HF pad and a higher number of dummy structures had been proposed for de-embedding at higher frequency [97–100], "Open-Short" is still the most widely accepted technique and has become the industry standard owing to its simplicity and accuracy at frequencies reaching up to 40 GHz.

However, there are weaknesses in using the above mentioned lumped-element equivalent circuit to model the HF pad. The de-embedding result depends on the assumed equivalent circuit, which may not be satisfactory at HF for each equivalent circuit. So, a more complicated equivalent circuit is

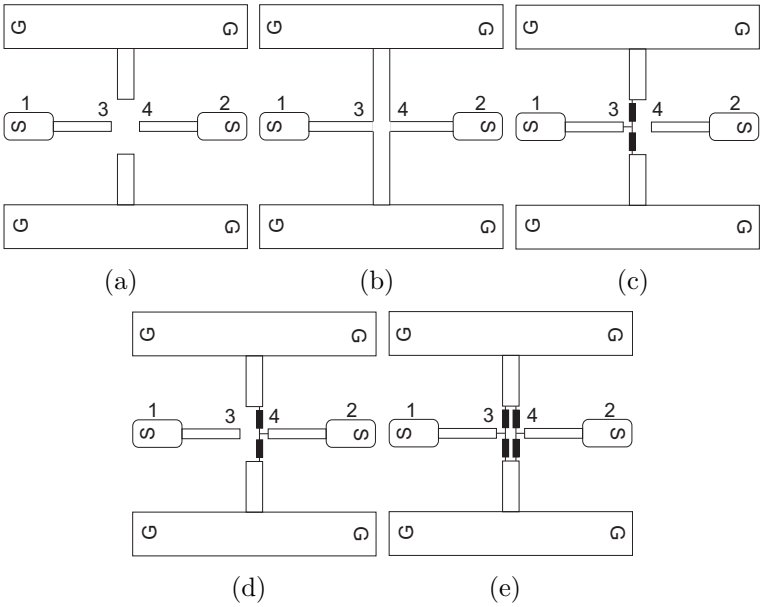


Figure 4.7: (a) open, (b) short, (c) load_3, (d) load_4 and (e) load_{34} dummy structures for HF pad de-embedding.

needed to take the HF distributed effect into account [100]. Actually, the HF pad can generally be modeled as a 4-port network [101–104] as shown in Fig. 4.6 with a 4×4 Y-matrix as

$$\begin{pmatrix} i_e \\ i_i \end{pmatrix} = \begin{pmatrix} Y_{ee} & Y_{ei} \\ Y_{ie} & Y_{ii} \end{pmatrix} \begin{pmatrix} v_e \\ v_i \end{pmatrix}, \tag{4.7}$$

where Y_{ee}, Y_{ei}, Y_{ie} and Y_{ii} are 2×2 Y-matrices.

$$\left(i_e\right) = \begin{pmatrix} i_1 \\ i_2 \end{pmatrix} \text{ and } \left(i_i\right) = \begin{pmatrix} i_3 \\ i_4 \end{pmatrix} \tag{4.8}$$

are external and internal current vectors.

$$\left(v_e\right) = \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} \text{ and } \left(v_i\right) = \begin{pmatrix} v_3 \\ v_4 \end{pmatrix} \tag{4.9}$$

are external and internal voltage vectors. If the 4×4 Y-matrix is known, the characteristic of the DUT is also known. There are two ways to find 16 unknowns in the 4×4 Y-matrix. For one there is Electromagnetic (EM) simulation [102]. But the question is how to verify the simulation result with the

measured data, which may become a major concern when the internal 2-port is connected to tiny metal lines that cannot be accessed by the HF probes. Even if the lines were accessible, the additional coupling and leakage between the two ports may influence the accuracy of the measurement. Therefore, calibrating the 4-port EM simulation data by comparing to the available external 2-port measurement results can be useful to increase the confidence level when using this method.

The other one is from the commonly used "open", "short", "through" and "loads" dummy structures [103]. It should be mentioned that a "through" is just a linear combination of an "open" and a "short" [100] so that it offers no new information in determining the 4-port parameters. It is important to uniquely define dummy structures in finding the 4-port parameters to save time in de-embedding and area on the test-chip. From 4-port terminal current and voltage relationships with the DUT, the open ($i_i = 0$) and the short ($v_i = 0$) placed in the internal 2-port, measured Y_{ext}, Y_{open} and Y_{short} from the external 2-port can be expressed as [103]

$$Y_{ext} = Y_{ee} - Y_{ei}(Y_{dut} + Y_{ii})^{-1}Y_{ie} \tag{4.10}$$

$$Y_{open} = Y_{ee} - Y_{ei}Y_{ii}^{-1}Y_{ie} \tag{4.11}$$

$$Y_{short} = Y_{ee} \tag{4.12}$$

Substituting (4.10), (4.11) and (4.12) into (4.5), the Open-Short de-embedding result of the DUT can be expressed in term of 2×2 Y-matrices as [104]

$$Y_{dut}^{os} = Y_{ei}Y_{ii}^{-1}Y_{dut}Y_{ii}^{-1}Y_{ie}, \tag{4.13}$$

while only Y_{ei}, Y_{ii} and Y_{ie} are left to find Y_{dut} after open-short de-embedding.

In fact, the 16 unknowns in the 4-port matrix can be reduced to 10 unknowns with the reciprocity theorem for the passive device, which allows Y_{ee}=Y_{ee}^T, Y_{ii}=Y_{ii}^T and Y_{ie}=Y_{ei}^T. By defining $A = Y_{ei}Y_{ii}^{-1}$ as in [104], the eqn. (4.13) becomes

$$Y_{dut}^{os} = AY_{dut}A^T. \tag{4.14}$$

There are now only 4 unknowns left in the matrix A to find the Y_{dut}, so that two additional well-defined dummy structures are needed. As a result, a well-defined load (load_3) at port 3 with port 4 open as in Fig. 4.7c can be chosen

as a third dummy structure. Its measured intrinsic Y-parameters are modeled as

$$Y_{load_3} = \begin{pmatrix} 1/Z(s) & 0 \\ 0 & 0 \end{pmatrix}, \tag{4.15}$$

where $Z(s)$ is the impedance of the load standard. Instead of providing Z(s) as a constant resistance [103] or a parallel R-C network [104] in (4.15), it is given as a frequency domain transfer function for using it at any frequency. The complexity and impedance of $Z(s)$ can be found by EM simulation with the proper process information. Substituting (4.15) into (4.14), the "Open-Short" de-embedding result of load$_3$ ($Y^{os}_{load_3}$) in terms of unknowns in matrix A and $Z(s)$ will be

$$Y^{os}_{load_3} Z(s) = \begin{pmatrix} a_{11}^2 & a_{11}a_{21} \\ a_{11}a_{21} & a_{21}^2 \end{pmatrix} = B. \tag{4.16}$$

Where B is a 2×2 matrix used to find a_{11} and a_{21}. To ensure the continuity of unknowns in the matrix B in the frequency domain, they can be described by the rational complex function as for Z_{sub} in Section 3.4 as

$$b_{ij}(s) = \frac{\sum_{k=1}^{n} N_{n-k} s^{n-k}}{s^n + \sum_{k=1}^{n} D_{n-k} s^{n-k}}, \tag{4.17}$$

where $(N_i, D_i)|_{i=0,\ldots,n-1}$ are 2n model parameters, which can be obtained by solving the linear least square problem for both real (Re) and imaginary (Im) parts of unknowns in the matrix B ($b_{ij}(s)$) and $Y^{os}_{load_3} Z(s)$ as

$$S(N_i, D_i) = \sum_{l=1}^{N_w} (Re, Im[b_{ij}(jw_l)] - Re, Im[Y^{os}_{load_{3_{ij}}} Z(jw_l)])^2, \tag{4.18}$$

where $S(N_i, D_i)$ is the sum of the squares of the deviations and N_w is the number of frequency samples. Since the "Open-Short" de-embedding is accurate enough at low frequency, it should be almost identical to the 4-port de-embedding method at low frequency. As a result, matrix A should be close to an identity matrix at low frequency. The boundary conditions of $|a_{11}| \simeq 1$ and $|a_{21}| \simeq 0$ at low frequency will help to find a unique solution for matrix B through the least square fitting to $Y^{os}_{load_3} Z(s)$. The a_{11} and a_{21} can then be solved from their corresponding b_{ij}.

Another load at port 4 with port3 open (load_4) as shown in Fig. 4.7d can be chosen as fourth dummy structure. Its measured Y-parameters are modeled as

$$Y_{load_4} = \begin{pmatrix} 0 & 0 \\ 0 & 1/Z(s) \end{pmatrix}. \tag{4.19}$$

The "Open-Short" de-embedding result of load_4 ($Y^{os}_{load_4}$) in terms of unknowns in matrix A and $Z(s)$ will be

$$Y^{os}_{load_4} Z(s) = \begin{pmatrix} a_{12}^2 & a_{12}a_{22} \\ a_{12}a_{22} & a_{22}^2 \end{pmatrix} = C. \tag{4.20}$$

Where a_{12} and a_{22} can be found by least square fitting of results in matrix C to $Y^{os}_{load_4} Z(s)$.

If the HF pad is symmetric as it is usually the case in a HF device layout, matrix A will be reciprocal and symmetric ($a_{11} = a_{22}$ and $a_{12} = a_{21}$). Then, only one more dummy structure is needed to find the A matrix. A load at port 3 and port 4 (load_{34}) as shown in Fig. 4.7e can be used as a third dummy structure. Its measured intrinsic Y-parameters are modeled as

$$Y_{load_{34}} = \begin{pmatrix} 1/Z(s) & 0 \\ 0 & 1/Z(s) \end{pmatrix} \tag{4.21}$$

Substituting (4.21) into (4.14), the "Open-Short" de-embedding result of load_{34} ($Y^{os}_{load_{34}}$) in terms of unknowns in matrix A and $Z(s)$ will become

$$Y^{os}_{load_{34}} Z(s) = \begin{pmatrix} a_{11}^2 + a_{21}^2 & 2a_{11}a_{21} \\ 2a_{11}a_{21} & a_{11}^2 + a_{21}^2 \end{pmatrix} = D \tag{4.22}$$

Again, a_{11} and a_{21} can be found by a least square fit of the transfer function represented unknowns in D to $Y^{os}_{load_{34}} Z(s)$. In both symmetric and asymmetric pads, Y_{dut} can be found after Y^{os}_{dut} and matrix A are determined.

From the above analysis a general 4-port de-embedding procedure can be achieved by measuring four dummy structures for asymmetric HF pads and three dummy structures for symmetric HF pads. Comparing to the other de-embedding methods [95–100], it is universally valid.

4.2.4 Noise measurement

Low frequency noise measurements are necessary for flicker ($1/f$) noise model parameter extraction. The $1/f$ noise spectral current density, as implies by

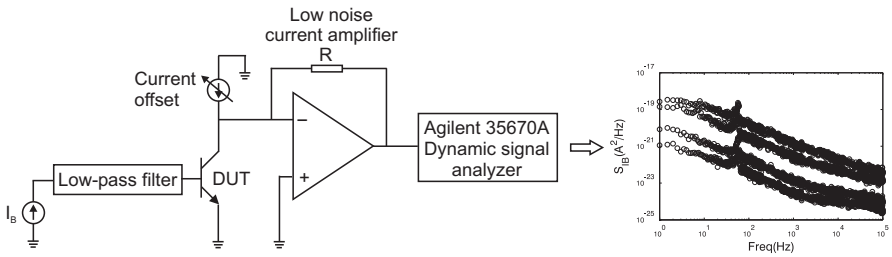


Figure 4.8: 1/f noise measurement setup.

its name, is reverse proportional to the measuring frequency and its corner frequency is usually in the KHz range. Basically, the measurement can be done on the DC test structure by using DC probes. Use of HF structures with HF probes, however, offers better shielding between the probes and is preferable over DC testing in order to avoid oscillations when measuring high gain transistors such as SiGe HBTs. Fig. 4.8 shows a schematic representation of $1/f$ noise measurement setup from [105]. The base terminal of a bipolar transistor is connected to the DC analyzer through a very low cut-off frequency low-pass filter to block the noise coming from the DC analyzer. The collector node of the device is connected to a low noise current amplifier (LNA) with a current offset to bias the transistor at the proper bias point without saturating the LNA. The amplified noise current is fed into a dynamic signal analyzer for analyzing the collector noise current spectral density (S_{IC}). Base noise current spectral density (S_{IB}) for $1/f$ noise parameter extraction can then be obtained from

$$S_{I_B} = \frac{S_{I_C}}{\beta^2}, \tag{4.23}$$

where β is the current gain of the DUT.

High frequency noise mainly results from the shot noise and the thermal noise, which arise from the DC current and parasitic resistances. The HF noise should be properly represented in the model if the DC current and parasitic resistance are modeled correctly; there are no particular Mextram model parameters for representing the high frequency noise. The high frequency noise data is mainly used for model verification, not for the purpose of Mextram parameter extraction. Therefore, it will not be discussed more details about the high frequency noise measurement here.

4.3 Single model parameter extraction

The accuracy of compact models in circuit simulation depends not only on the correct physical description of various physical phenomena in the device but also on a reliable, robust and unambiguous parameter extraction methodology. It should be emphasized that a general and fully automatic parameter extraction procedure is currently unavailable. As a practical example of the full Mextram (Level 504) parameter extraction, we present here results for the TSMC 0.18 μm double-poly high-speed SiGe BiCMOS technology [83]. The data has been offered by TSMC under the agreement that it is only for compact model evaluation and research purposes only.

4.3.1 Layout and technology data

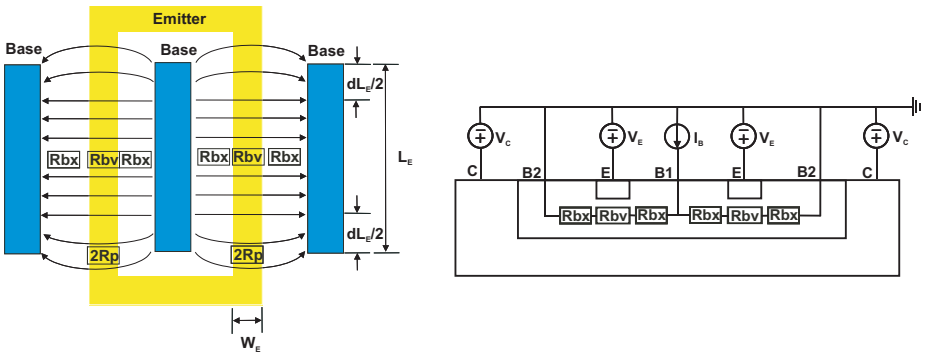
Prior to the parameter extraction, some layout and technology data in table 4.1 is needed for initial parameter calculation [76]. The device with W_E= 0.4 μm, L_E= 10.16 μm and N_B =2 is selected for parameter extraction, while W_{epi} and N_{epi} of the selected device can be obtained from the BC junction capacitance-voltage characteristic [106] or Secondary Ion Mass Spectroscopy (SIMS) profile. Among the process and technology data, $\rho_\square$ is the most critical parameter since it will be used for calculation of the Mextram parameter `RBV`, which is implicit in the device characteristics.

Table 4.1: Layout and technology data needed for parameter extraction.

Data	Description
W_E	Emitter width
W_E	Emitter length
N_B	Number of base contact
$\rho_\square$	Pinched base sheet resistance
W_{epi}	Collector epilayer thickness
N_{epi}	Collector epilayer doping level

Pinched base sheet resistance measurement

Although the base resistance can be extracted from HF small-signal equivalent circuits [107–110], it is difficult to calculate $\rho_\square$ out of the extracted base resistance as a result of base-width modulation and current crowding



(a) (b)

Figure 4.9: (a) A test structure and (b) setup for pinched base sheet resistance measurement.

effects [111, 112] at the device's operating point. So, $\rho_{\square}$ is measured and extracted by using a special emitter ring structure [113] as shown in Fig. 4.9(a). The bias voltages and currents of the measurement setup in Fig. 4.9(b) should be small to avoid base width modulation and current crowding effects. Therefore, I_B/W_E= 1∼20 μA/μm and V_{CB2}=V_{EB2}=0.1 V, which weakly reverse bias the BE and BC junctions to ensure current flowing between $B1$ and $B2$, are applied in the measurement setup. Based on the equivalent circuit of the test structure in Fig. 4.9(a), the measured base resistance (R_{B1B2}) including an internal base resistance (Rbv), an external base resistance (Rbx) and a fringing shunt resistance (Rp) is expressed as

$$\begin{aligned} R_{B1B2} &= \frac{V_{B1B2}}{I_B}, \\ &= (\frac{Rbv}{2} + Rbx)//Rp, \\ &= \left(\frac{\rho_{\square}}{2}\frac{W_E}{L_E - dL_E} + Rbxl\frac{1}{L_E - dL_E}\right)//Rp, \end{aligned} \tag{4.24}$$

where $Rbxl$ and dL_E are the external base resistance times length and a correction between drawn and intrinsic length of the intrinsic base resistance. Since R_{B1B2} comprises Rbv, Rbx and Rp, it is not possible to separate them from only a single W_E and L_E. Therefore, measurement of a series of different W_E and L_E test structures is required. From the measured R_{B1B2} at each different geometry as shown in Fig. 4.10, it is almost constant in the sweep-

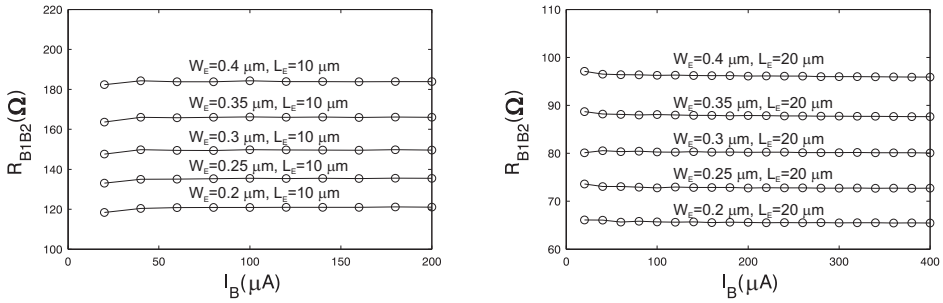


Figure 4.10: Measured R_{B1B2} from different geometries vs. forced base current.

ing I_B range. It proves that the measured resistance is not influenced by the current crowding and base-width modulation effects at such a low current and low bias level.

In order to remove influence of Rp from R_{B1B2}, R_{B1B2} is expressed in terms of its conductance (Y_{B1B2}) as

$$\begin{aligned} Y_{B1B2} &= \left(\frac{Rbv}{2} + Rbx\right)^{-1} + \frac{1}{Rp}, \\ &= \left(\frac{\rho_\square}{2}\frac{W_E}{L_E - dL_E} + Rbxl\frac{1}{L_E - dL_E}\right)^{-1} + \frac{1}{Rp}. \end{aligned} \tag{4.25}$$

By subtracting (4.25) for L_E=10 μm ($Y_{B1B2,10}$) from (4.25) for L_E=20 μm ($Y_{B1B2,20}$), the reciprocal of their difference becomes a linear equation as

$$(Y_{B1B2,20} - Y_{B1B2,10})^{-1} = \frac{\rho_\square}{2}\frac{W_E}{L_{E20} - L_{E10}} + Rbxl\frac{1}{L_{E20} - L_{E10}}. \tag{4.26}$$

Then, $\rho_\square$ and $Rbxl$ can be extracted from the slope and intersection at $W_E = 0$ of the $(Y_{B1B2,20} - Y_{B1B2,10})^{-1}$ v.s. W_E plot as shown in Fig. 4.11.

With the assumption that the external base resistance is described by a homogeneous low-resistivity value, it will behave like a contact surrounding the internal base. The Mextram parameter `RBV`, which is represented by the zero-bias value of the variable part of the base resistance, can be calculated from $\rho_\square$ for $N_B = 2$ as [114]

$$\texttt{RBV} = \rho_\square \frac{W_E}{L_E}\left(\frac{1}{12} - \left(\frac{1}{12} - \frac{1}{28.45}\right)\frac{W_E}{L_E}\right), \tag{4.27}$$

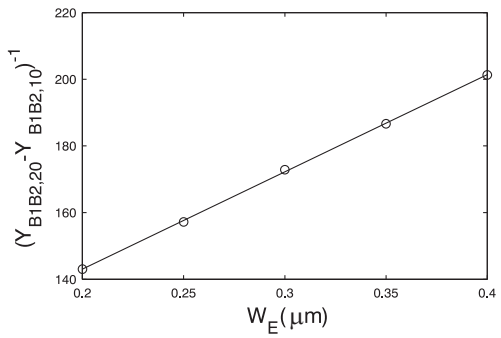


Figure 4.11: Reciprocal of conductance difference vs. emitter width.

where one of the coefficient 28.6 in [114] has been replaced with 28.45 in (4.27) suggested by [31]. The constant part of the external base resistance RBC is calculated from *Rbxl* divided by emitter periphery as

$$\texttt{RBC} = \frac{Rbxl}{2(W_E + L_E)}. \tag{4.28}$$

4.3.2 Parameters extracted from measured data

The implemented Verilog-A Mextram model is employed within IC-CAP [115] environment using Agilent ADS simulator [74] with a Verilog-A compiler for parameter extraction from measured data by fitting simulated to measured data. In order to minimize the correlation between electrical and temperature parameters of the Mextram model, the electrical parameters may be split into parameters extracted at low-current (not affected by the self-heating) and high-current related parameters. Moreover, in order to further reduce the parameter correlation, the electrical and temperature parameters could be further split into small groups that are extracted subsequently, starting from the measured data that is most sensitive to the particular parameters. An example of such a splitting for Mextram parameter extraction is given in Table 4.2 for low-current and in Table 4.3 for high-current measurement steps, respectively. The electrical parameters are extracted beginning with low-current parameters moving on to high-current parameters. In the mean time, the temperature parameters related to the low-current parameters are extracted prior to the high-current parameters in order to ensure accurate modeling the temperature rise as a result of self-heating in the high current region.

Table 4.2: Extraction of Low-Current Parameters

Measurement	Elec. parameters	Temp. parameters
BE depl. capacitance (C_{BE})	CJE[12a], PE[12a],VDE	VGB
BC depl. capacitance (C_{BC})	CJC[12b], PC[12b], VDC, XP[12b]	VGC[13c]
SC depl. capacitance (C_{SC})	CJS[12c], PS[12c], VDS	VGS
Forward-Early ($\frac{I_B}{I_{B0}}$)	WAVL[12d], VAVL[12d], XREC	
Reverse-Early ($\frac{I_E}{I_{E0}}$)	VER[12e]	AQBO[13d]
Forward-Early ($\frac{I_C}{I_{C0}}$)	VEF[12f]	AQBO
Forward-Gummel (I_C)	IS[12g]	VGB[13e], AB, DAIS
Forward-Gummel (H_{fe})	BF[12h], IBF[12h], MLF[12h]	DVGBF[13f], VGJ[13f]
Forward-Gummel (I_B)	RE[12g]	AE
Output-characteristic (I_C)	RCC[14b]	AC
Reverse-Gummel (I_{Sub})	ISS[13a], IKS[13a]	VGS[13g], AS
Reverse-Gummel (H_{fc})	BRI[13b], IBR[13b], VLR[13b]	DVGBR[13h], VGC

Low-current parameter extraction

Low current parameters that can be extracted from the measured data is listed in Table 4.2. Following the listed measurement setups in Table 4.2 from the top to the bottom, electrical parameters of depletion capacitances are extracted first. The zero bias depletion capacitances CJE, CJC, and CJS can be extracted from the measured C_{BE}, C_{BC} and C_{SC} around zero-bias. It is found that the build-in voltages VDE, VDC and VDS are co-related with the grading coefficients PE, PC and PS. Therefore, VDE, VDC and VDS are set to a reasonable value or calculated from the doping concentrations in the emitter, base, collector and substrate region. Only PE, PC, XP and PS are extracted from the measured C_{BE}, C_{BC} and C_{SC}, which is shown in Fig. 4.12a,b,c.

The low-current avalanche effect can be seen in the forward-Early measurement with decreasing I_B as a result of increasing V_{CB}. The measured and simulated I_B is normalized to its zero V_{CB} current (I_{B0}) as shown in Fig. 4.12d so there is no need to use additional parameters, which are used temporarily to adjust the simulated to the measured I_B at V_{CB}=0, to extract WAVL and VAVL. The XREC is set to zero due to a lack of evidence for existence

of the neutral base recombination current. The reverse Early voltage VER can be extracted from the measured emitter current normalized to its zero V_{EB} current (I_E/I_{E0}), while the forward VEF can be extracted from the measured collector current normalized to its zero V_{CB} current (I_C/I_{C0}) before tunneling and avalanche effects occur at high V_{EB} and V_{CB} as shown in Fig. 4.12e,f. The forward saturation current IS is extracted from the low I_C region of the forward-Gummel measurement in Fig. 4.12g while HFE, IBF and MLF are extracted from the forward current gain (H_{fe}=I_C/I_B) of the same measurement as shown in Fig. 4.12h. Parasitic resistance RE is derived from using the Ning-Tang method [116] based on the I_B deviation from its ideal value as indicated in Fig. 4.12g, while the RCC is extracted by applying the Forced-beta [21, 117] method to the output-characteristic for high current parameter extraction as an initial value and fit to the initial slope of the output-characteristic afterwards. The substrate current (I_{Sub}) in reverse-Gummel measurement at low V_{BC} and high V_{BC} in Fig. 4.13a are used for extraction of ISS and IKS, while BRI, IBR and VLR are extracted from the intrinsic reverse current gain (H_{fc}=$I_E/(|I_B| - |I_{Sub}|)$) as shown in Fig. 4.13b at the same measurement.

The low-current parameter extraction ends at reverse-Gummel measurement setup and each extracted parameter in Table 4.2 is linked to its fitting plot with a figure number at its superscript.

Temperature parameter extraction

The mobility temperature parameters AE, AB, AC and AEPI are determined from a database of mobility temperature parameters v.s. doping concentration [76, 118] at emitter, base, buried layer and epilayer and AS is equal to AC for a closed buried layer. In order to directly extract temperature parameters from measured data of various temperatures in IC-CAP, the measuring temperature of each DUT has to be added in the standard measured data files as an additional input parameter and then measured data files of different temperatures are combined into a single data fie. Since VGC is the corresponding temperature parameter to the measured C_{BC} and reverse Gummel setups in Table 4.2, it is first extracted from measured C_{BC} vs. temperature at 25 $\sim$ 150 °C in steps of 25 °C as shown in 4.13c. Following the VGC extraction is the temperature parameter of zero-bias base charge AQBO extraction from the reverse-Early measurement v.s. temperature as shown in Fig. 4.13d. The bandgap related temperature parameters VGB, VGJ, DVGBF, VGS, DVGBR are extracted from the measured forward-Gummel (I_C, H_{fe}), reverse-Gummel (I_{Sub}, H_{fc}) at the same temperature steps as shown in Fig. 4.13e, f, g, h, which is

Table 4.3: Extraction of High Current Parameters

Measurement	Elec. parameters	Temp. parameters
Output-characteristic (V_{BE}, I_C)	RTH[14a], IK[14b]	AB
Forward-Gummel (H_{fe})	RCV[14c], VDC	
Cut-off frequency (f_T)	TAUE[14d], TAUB[14d]	DVGTE[14e], AB, AQBO
	TEPI[14d], RCV[14d], VDC	AEPI, VGC[14g]
	IHC[14d], SCRCV[14d]	
	MC[14d], MTAU[14d], AXI[14d]	
S- and Y-Parameter (S_{12}, Y_{12})	XEXT[15,16]	
$1/f$ noise (S_{IB})	AF[14h], KF[14h]	

also labeled at each extracted temperature parameter in Table 4.2.

High-current parameter extraction

High current parameters that can be extracted from the measured data are listed in Table 4.3. The thermal resistance RTH is extracted from the V_{BE} drop-off in the output-characteristics measured at three different forced I_B=$\frac{1}{4}I_{Bset}$, $\frac{1}{2}I_{Bset}$ and I_{Bset} as shown in Fig. 4.14*a*, where I_{Bset} is the I_B in the forward-Gummel measurement when H_{fe} falls to 50% of its maximum value. The knee current IK is extracted from the spacing between different I_C of the same output-characteristic shown in Fig. 4.14*b* after RTH extraction. The resistance of the un-doped epilayer RCV can be extracted from the H_{fe} fall-off at high I_B in the forward-Gummel measurement shown in Fig. 4.14*c* or the following f_T measurement at low V_{CB} shown in Fig. 4.14*d*. The extraction of TAUE, TAUB, TEPI, RCV, IHC, SCRCV, MC, MTAU and AXI from different segments of f_T is shown in Fig.4.14d, while the temperature parameter of the emitter transit time DVGTE can be extracted from the f_T drop-off at high temperature as shown in Fig. 4.14*e*, *f*, *g* but care must be taken to see if the high temperature f_T at low V_{CB} drop-off in Fig. 4.14*g* is too excessive due to an over-estimated VGC. The measured S-parameters and its transformed Y-parameters v.s. frequency biased at around peak f_T of the device as shown in Fig. 4.15 and 4.16 are used for the partition parameter XEXT extraction by the optimal fit of the simulated to measured S_{12} and Y_{12}, while the good fit of the simulated to measured Y_{11} and Y_{21} verifies RBV, RBC and RE extraction at low-current.

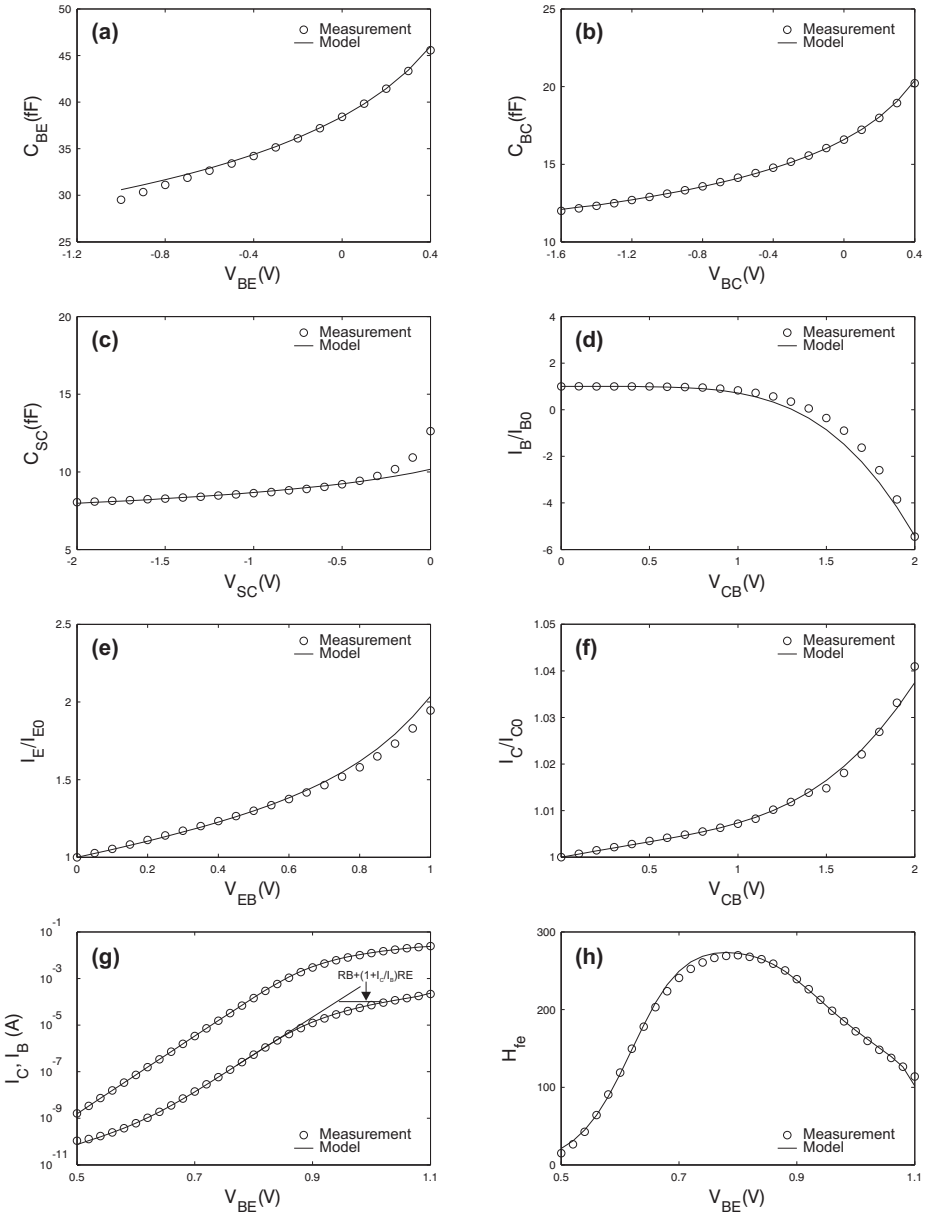


Figure 4.12: Fitting plots for the low-current parameter extraction.

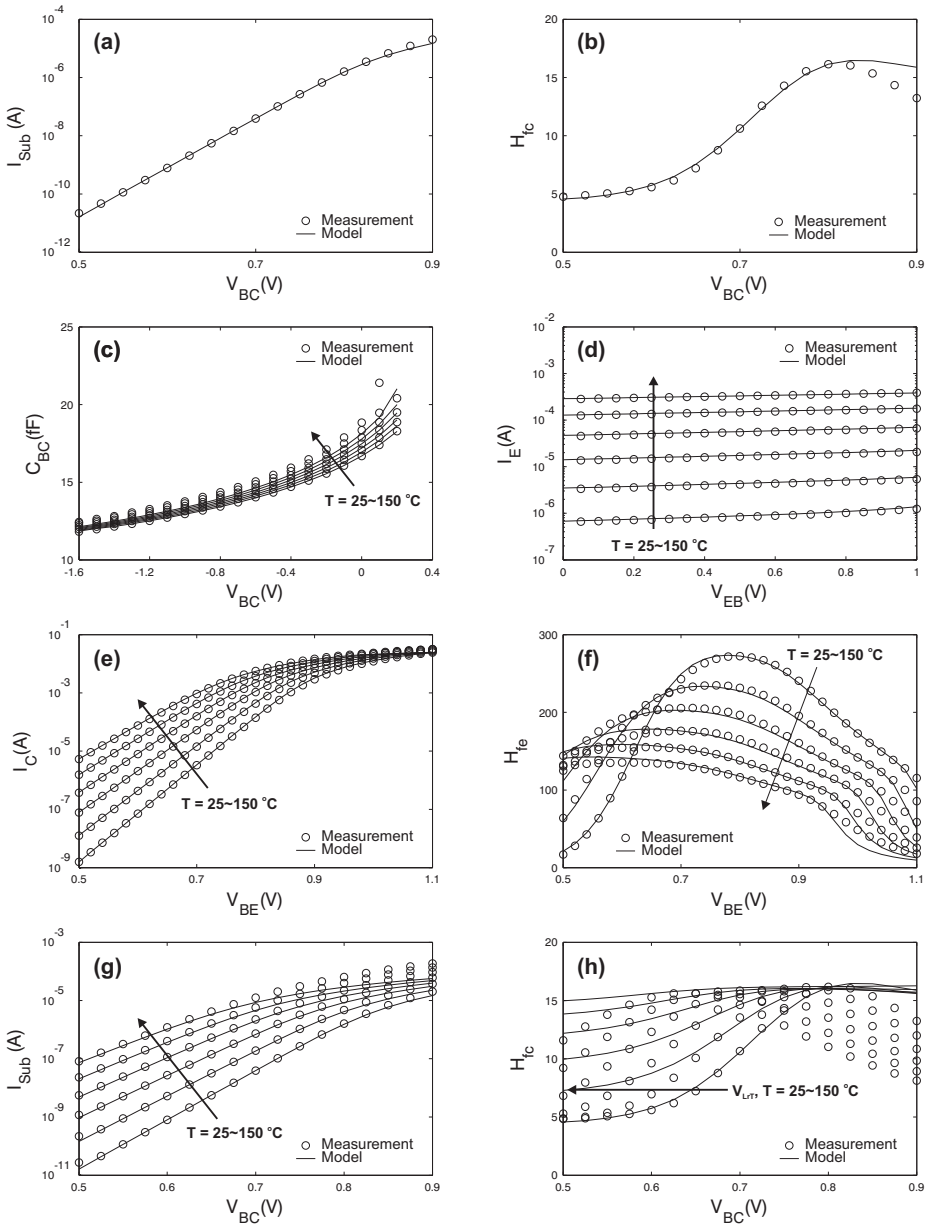


Figure 4.13: Fitting plots for low-current and temperature parameter extraction.

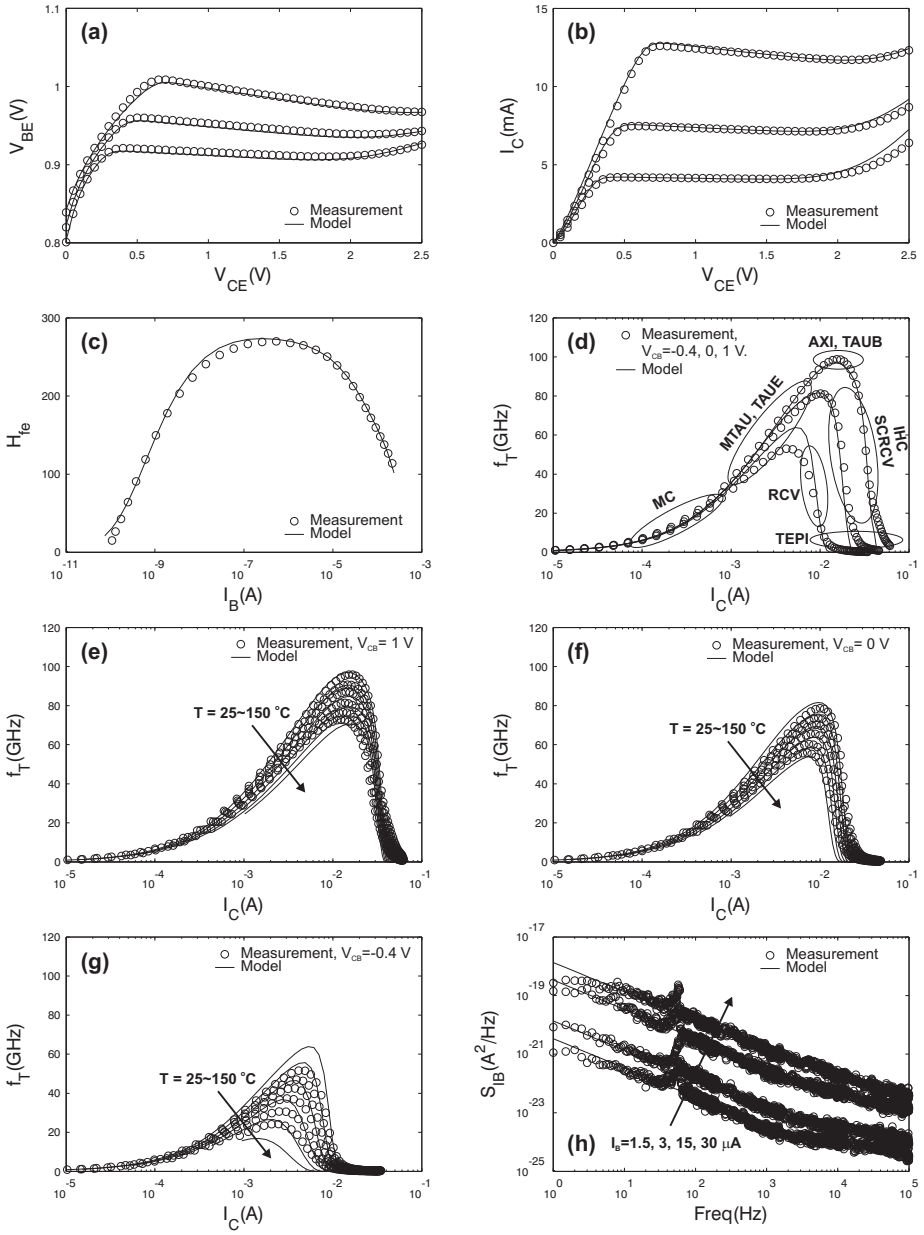


Figure 4.14: Fitting plots for high-current, `DVGTE` and noise parameter extraction.

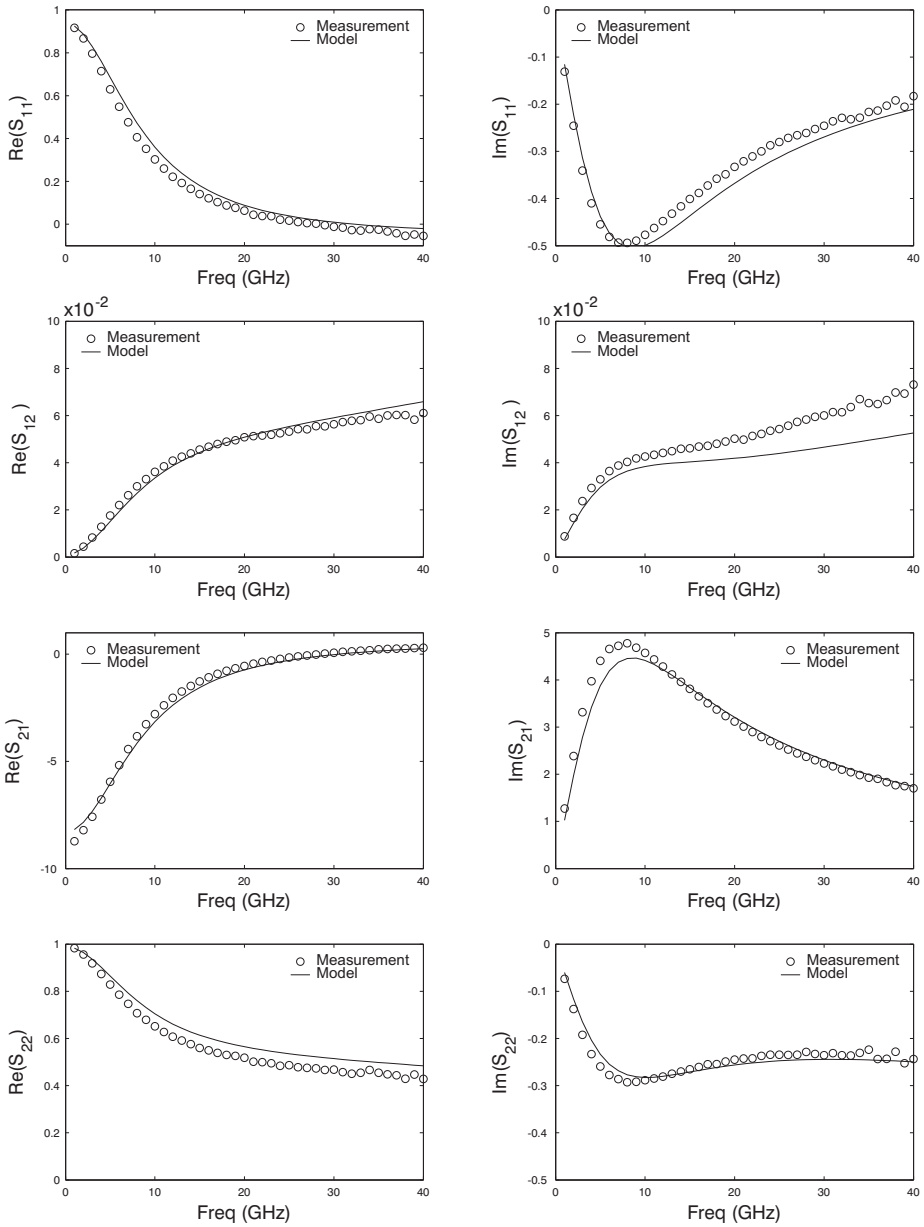


Figure 4.15: Fitting plots of S-parameters v.s. frequency (1∼40 GHz) biased at $V_{BE} = 0.95$ V and $V_{CE} = 2$ V.

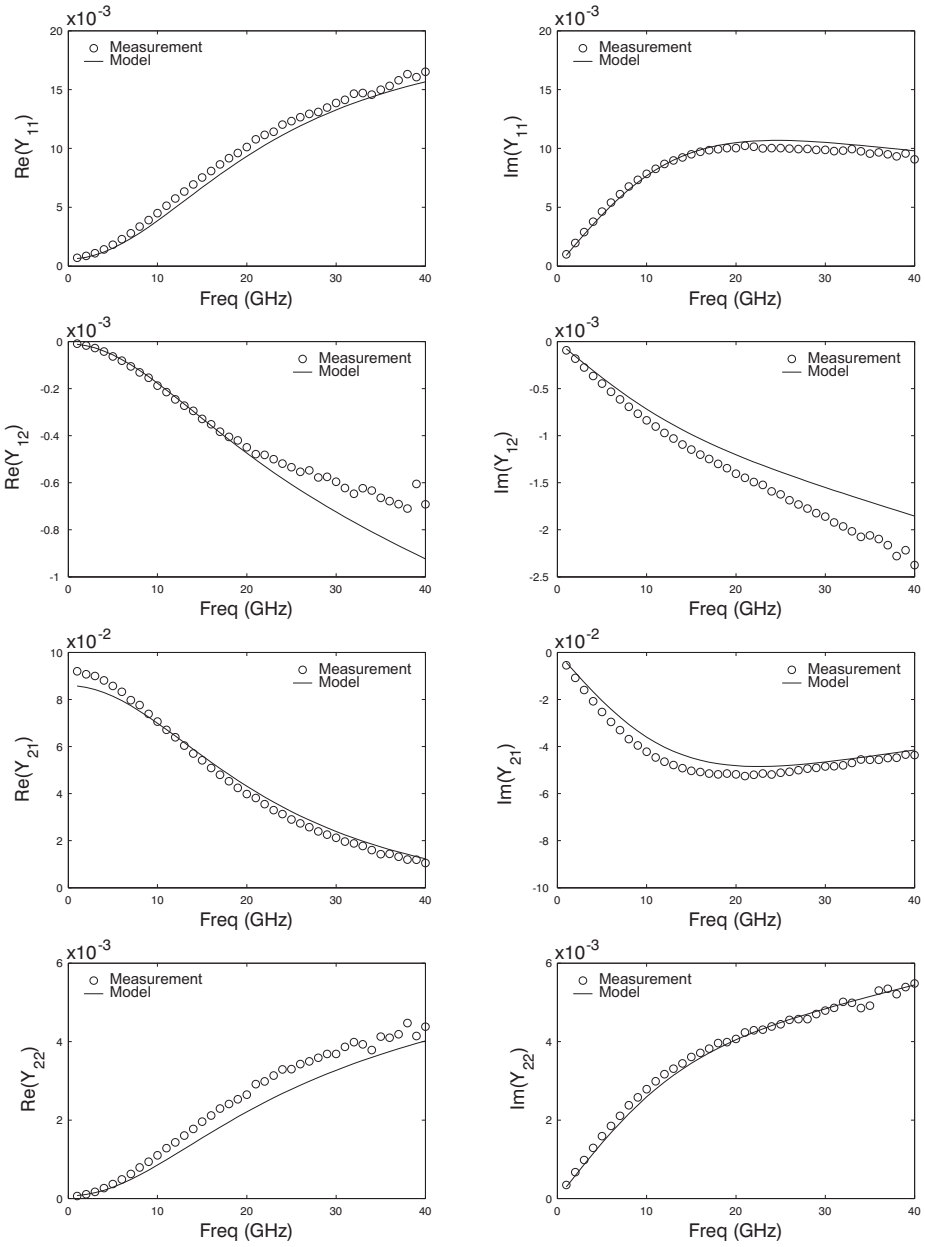


Figure 4.16: Fitting plots of Y-parameters v.s. frequency (1∼40 GHz) biased at $V_{BE} = 0.95$ V and $V_{CE} = 2$ V.

Although S_{IB} is usually measured at low I_B currents, its noise parameters AF and KF are extracted last. Note, that the fitting plot in Fig. 4.14h is derived from a W_E=0.6 μm and L_E=10.16 μm size device and the measurement data is at courtesy of TSMC. Finally, two to three iterative extractions between low-current, temperature and high-current measurement setups may be needed to get the optimum parameter set for the overall measurement setups.

4.3.3 Improved temperature scaling model for the reverse current gain

As can be seen in the fitting plot of H_{fc} at various temperatures in Fig. 4.13h, the result is not as good as for the fitting plot of H_{fe} at various temperatures in Fig. 4.13f. The distinct temperature dependence of VLR is observed in measured H_{fc} as shown in Fig. 4.13h. As a result, it is suggested to have linear temperature dependence for VLR as

$$V_{LrT} = \texttt{VLR}(1 + \texttt{VLRT}(t_N - 1)), \tag{4.29}$$

where V_{LrT} is the temperature scaled VLR and VLRT is the temperature parameter of VLR.

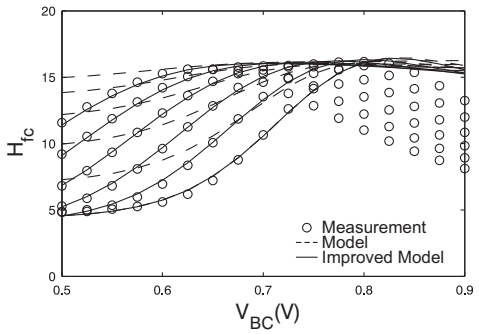


Figure 4.17: Improved temperature scaling for H_{fc}.

The temperature scaling of IBR is also suggested to de-couple from VGC, which is corresponding to many measurement setups, and changed to

$$I_{BrT} = \texttt{IBR}\; t_N^2 \exp\left(\frac{-\texttt{VGJC}}{2V_{\Delta T}}\right). \tag{4.30}$$

Where I_{BrT} is the temperature scaled `IBR` and a band-gap voltage for the BC depletion region (`VGJC`) replaces `VGC` as the new temperature parameter for `IBR`. With the new temperature parameters `VLRT` and `VGJC`, the improved model in solid line shows better fit to the measured H_{fc} at low-current compared to the original model in dashed line as shown in Fig. 4.17.

4.4 Conclusion

In this chapter, the complete modeling procedure from data acquisition to single Mextram model parameter extraction has been introduced. With a systematic parameter extraction procedure from pinched base sheet resistance to low-current, to various temperatures, to high-current and to noise parameters, the Mextram model shows good fit to the measured DC, CV, HF S-parameters/Y-parameters and $1/f$ noise characteristics from a high-speed SiGe HBT.

Chapter 5

Scalable bipolar model approach and parameter extraction

5.1 Introduction

As the complexity of the integrated circuits increases, scalable device models are becoming crucial for the optimum circuit design. MOS transistor models [119,120] typically give quite reliable scaling results since they describe only lateral (surface) current flow as shown in Fig. 5.1a. On the other hand, scaling of bipolar transistor models is significantly more involving due to current flow in both vertical and lateral directions as shown in Fig. 5.1b. Moreover, the geometry scaling rules of bipolar devices generally depend on the layout configuration, device cross-section and fabrication process details. As a result, the geometry scalable bipolar transistor models and the corresponding parameter extraction procedures are not generally available and usually exist as an in-house proprietary.

The scalable bipolar models are mainly based on the physical [121] or electrical [28] size of the transistor's emitter window. The determination of the physical device sizes typically requires Scanning Electron Microscope (SEM) inspection of special device test samples during the process steps. The same device might not be used later for model evaluation. On the other hand, it is difficult to uniquely estimate the devices' electrical sizes from electrical measurement data [25]. Besides, the electrical device size could be quite different for the model parameters representing different physical quantities, like currents or charges. Scaling based directly on the drawn device dimensions in

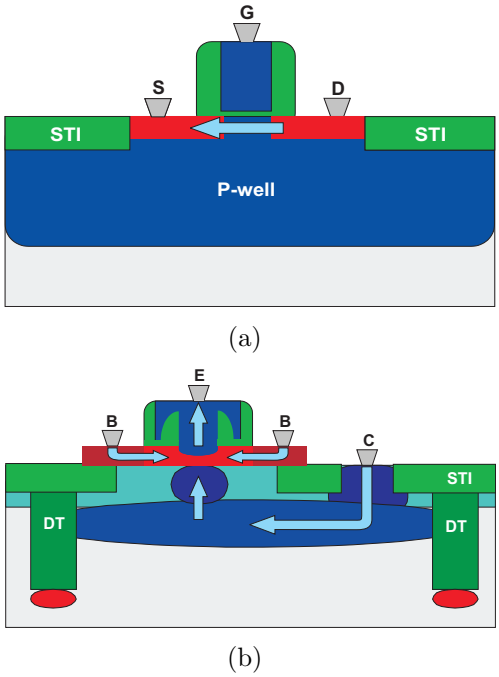


Figure 5.1: Current flow in (a) MOS and (b) bipolar devices.

the layout, if possible, would be the most convenient way from an application point of view.

For the scalable model implementation, it is possible to employ the sub-circuit netlist approach [121], which includes the geometry parameters and scaling rules in the model library. However, in that case it is difficult to have Intellectual Property (IP) protection of the scaling rules (if such a protection is required). Some model developers use a toolbox [28] that generates a model library for different geometries. It offers IP protection of the scaling rules but designers sometimes still have to come back to the model developers for a specific size, which they need for their design.

The main goal of this chapter is to propose these new concepts for a scalable bipolar transistor Mextram model and realize them efficiently. The concepts that we will propose are a scalable model, which can be scaled based on the drawn size in the layout of the device and is implemented in AHDL Verilog-A language. So, it offers IP protection of the scaling rules if the model is compiled as a shared library. In the mean time, circuits designers can still have full freedom in choosing device sizes and configurations in the valid geometry range

when linking the shared library in their design environment. Meanwhile, the geometry parameter extraction for the scalable model should be simple and better to be based on the electrical characteristics of devices with different geometries as their corresponding single model parameter extraction.

5.2 Physics of geometry scaling

The physical quantities in a bipolar transistor that scales with geometry can be separated into three categories: (I) current and charge, (II) ratio of current and charge and (III) parasitic and thermal resistance. In the following subsections, physical base geometry scaling rules for the three different categories of scalable Mextram model parameters will be derived.

5.2.1 Geometry scaling rules of current and charge

There are normally three PN junctions (BE, BC and SC) in a bipolar transistor. Fig. 5.2 shows the top view and cross-section of a PN junction, which can be separated into bulk, sidewall and corner components depending on the shape of the junction area. Therefore, a geometry scalable model parameter P that describes current or charge scaled with area of the PN junction can be described by the following expression [25]:

$$\mathrm{P} = \underbrace{P_A''(W+dW)(L+dL)}_{\text{Bulk}} + \underbrace{P_W''(W+dW) + P_L''(L+dL)}_{\text{Sidewall}} + \underbrace{P_C''}_{\text{Corner}} . \quad (5.1)$$

Where P_A'', P_W'', P_L'' and P_C'' are the constant geometry parameters that account for bulk, sidewall width, sidewall length and corner contributions, respectively. The W and L are drawn sizes of BE, BC or SC junction in the layout, while dW and dL are constant corrections between electrical and drawn sizes. Eqn. (5.1) can be expressed in terms of W and L as

$$\begin{aligned}\mathrm{P} = {} & (P_A''W + P_A''dW + P_L'')L + (P_A''dL + P_W'')W \\ & + P_A''dWdL + P_W''dW + P_L''dL + P_C''. \end{aligned} \quad (5.2)$$

If the extracted P v.s. L for different W is plotted as in Fig. 5.3, the slopes and intersections at $L = 0$ for different W can be expressed as

$$Slope = P_A^{''}W + P_A^{''}dW + P_L^{''}, \tag{5.3}$$

$$Inter. = (P_A^{''}dL + P_W^{''})W + P_A^{''}dWdL + P_W^{''}dW + P_L^{''}dL + P_C^{''}. \tag{5.4}$$

From the W dependence of (5.3) and (5.4), the slope (m_{slope}) and intersection (i_{slope}) of $Slope$ v.s. W and the slope ($m_{inter.}$) and intersection ($i_{inter.}$) of $Inter.$ v.s. W can be expressed as

$$m_{slope} = P_A^{''}, \tag{5.5}$$
$$i_{slope} = P_A^{''}dW + P_L^{''}, \tag{5.6}$$
$$m_{inter.} = P_A^{''}dL + P_W^{''}, \tag{5.7}$$
$$i_{inter.} = P_A^{''}dWdL + P_W^{''}dW + P_L^{''}dL + P_C^{''}. \tag{5.8}$$

There are, however, 6 unknown geometry parameters ($P_A^{''}$, $P_W^{''}$, $P_L^{''}$, $P_C^{''}$, dW and dL) in the 4 equations (5.5), (5.6), (5.7) and (5.8). Thus, no unique solution can be found for the unknowns. Even more measurement data can't help to find all of them. So, the remedy for it may be to reduce from 6 to 4 unknowns by determining dW and dL from SEM measurement or re-arranging (5.2) to

$$\mathrm{P} = P_A^{'}WL + P_W^{'}W + P_L^{'}L + P_C^{'}. \tag{5.9}$$

Where

$$P_A^{'} = P_A^{''}, \tag{5.10}$$
$$P_W^{'} = P_A^{''}dL + P_W^{''}, \tag{5.11}$$
$$P_L^{'} = P_A^{''}dW + P_L^{''}, \tag{5.12}$$
$$P_C^{'} = P_A^{''}dWdL + P_W^{''}dW + P_L^{''}dL + P_C^{''}. \tag{5.13}$$

are the 4 new composite geometry parameters.

If (5.9) can be further expressed in a form that is an analogy to the temperature scaling rule of the Mextram model [40] listed below:

$$\frac{\mathrm{P}}{\mathrm{P_{ref}}} = \begin{cases} f(t_N) = 1, & T_K = T_{RK}. \\ f(t_N), & T_K \neq T_{RK}. \end{cases} \tag{5.14}$$

i.e.: P is equal to its reference electrical parameter ($\mathrm{P_{ref}}$) at the reference temperature (T_{RK}) and scaled to the value at the evaluated temperature (T_K) by

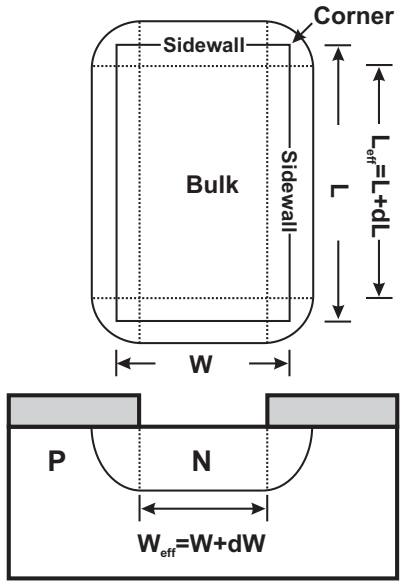


Figure 5.2: Top view and cross-section of a PN diode.

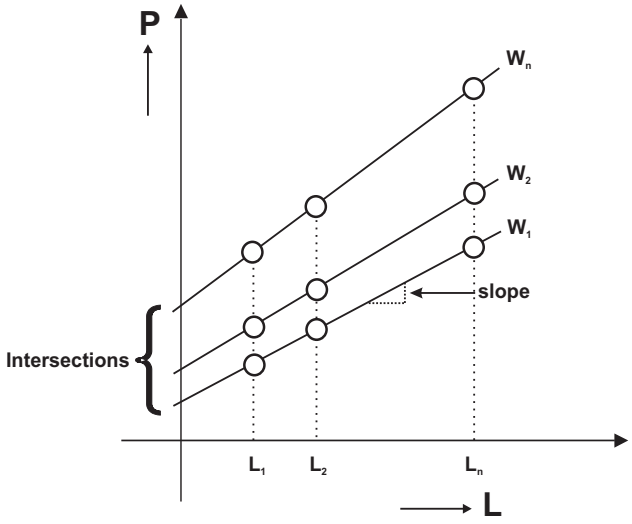


Figure 5.3: Extracted parameter P vs. L for different W.

its temperature scaling rule ($f(t_N)$) times $\mathrm{P_{ref}}$ and the normalized temperature (t_N) in $f(t_N)$ is the ratio of T_K to T_{RK}, both temperature and geometry parameters can have the same extraction methodology. Therefore, P in (5.9) is first normalized to the reference electrical parameter ($\mathrm{P_{ref}}$) at a reference geometry and temperature. Then, W and L are normalized to the reference geometry (W_R, L_R) as

$$\frac{\mathrm{P}}{\mathrm{P_{ref}}} = P_A \frac{WL}{W_R L_R} + P_W \frac{W}{W_R} + P_L \frac{L}{L_R} + P_C. \tag{5.15}$$

Now, $P_A = \frac{P_A^{'} W_R L_R}{\mathrm{P_{ref}}}$, $P_W = \frac{P_W^{'} W_R}{\mathrm{P_{ref}}}$, $P_L = \frac{P_L^{'} L_R}{\mathrm{P_{ref}}}$ and $P_C = \frac{P_C^{'}}{\mathrm{P_{ref}}}$ are the normalized geometry parameters. However, (5.15) is not the most compact form for geometry scaling since there is the constant corner term (P_C) in both reference and the evaluated geometry. When W=W_R and L=L_R, P=$\mathrm{P_{ref}}$ and (5.15) becomes

$$1 = P_A + P_W + P_L + P_C, \tag{5.16}$$

P_C can be removed by subtracting (5.16) from (5.15). Their difference will be

$$\frac{\mathrm{P}}{\mathrm{P_{ref}}} - 1 = P_A \left(\frac{WL}{W_R L_R} - 1 \right) + P_W \left(\frac{W}{W_R} - 1 \right) + P_L \left(\frac{L}{L_R} - 1 \right). \tag{5.17}$$

By re-arranging (5.17), the geometry scaling equation becomes:

$$\frac{\mathrm{P}}{\mathrm{P_{ref}}} = 1 + P_A \left(\frac{WL}{W_R L_R} - 1 \right) + P_W \left(\frac{W}{W_R} - 1 \right) + P_L \left(\frac{L}{L_R} - 1 \right), \tag{5.18}$$

which is the most compact form. When $W = W_R$ and $L = L_R$ in (5.18), $\frac{\mathrm{P}}{\mathrm{P_{ref}}} = 1$ as $T = T_{RK}$ in (5.14). A similar geometry scaling approach can be found in [122] but it is limited to emitter length scaling.

5.2.2 Geometry scaling rules for ratio of current and charge

However, some of the scalable model parameters represent the ratio of current and charge. For example, the ratio of total base charge to the BC junction capacitance (`CJC`) defines the forward early voltage (`VEF`). As a result, `VEF` will scale as the ratio of two geometry scalable model parameters:

$$\mathrm{P} = \frac{P_{A1}^{'} WL + P_{W1}^{'} W + P_{L1}^{'} L + P_{C1}^{'}}{P_{A2}^{'} WL + P_{W2}^{'} W + P_{L2}^{'} L + P_{C2}^{'}}. \tag{5.19}$$

Only the geometry parameter in the numerator (P'_{A1}, P'_{W1}, P'_{L1} and P'_{C1}) is extracted for `VEF` scaling, while the geometry parameters in the denominator (P'_{A2}, P'_{W2}, P'_{L2} and P'_{C2}) has been extracted from the scaling of `CJC`. But it is usually better to de-couple ratio parameter from the influence of the other geometry scaling parameter in case that there is a necessity to set it constant. For the modern VLSI technology, the junction depth is usually smaller than the lateral dimension of the junction according to the minimum design rules, which means that the bulk component still dominates the total junction area. As a result, by normalizing the numerator and denominator of (5.19) with their respective bulk components, the ratio parameter is then first order approximated as

$$\mathrm{P} \simeq P'_A + \frac{P'_W}{W} + \frac{P'_L}{L} + \frac{P'_C}{WL}. \tag{5.20}$$

Where P'_A, P'_W, P'_L and P'_C are composites of the geometry parameters in (5.19). The normalized equation of (5.20) will be

$$\frac{\mathrm{P}}{\mathrm{P_{ref}}} = P_A + P_W \frac{W_R}{W} + P_L \frac{L_R}{L} + P_C \frac{W_R L_R}{WL}, \tag{5.21}$$

where $P_A = \frac{P'_A}{\mathrm{P_{ref}}}$, $P_W = \frac{P'_W}{\mathrm{P_{ref}} W_R}$, $P_L = \frac{P'_L}{\mathrm{P_{ref}} L_R}$ and $P_C = \frac{P'_C}{\mathrm{P_{ref}} W_R L_R}$ are the geometry parameters for scaling of the ratio parameters. When W and L are large, P is approaching the constant P_A as there is only a bulk component. Since there is no geometry dependence of P_A in (5.21), it can be removed by substraction (5.16) from (5.21). The compact reference based geometry scaling equation for the ratio parameters will be

$$\frac{\mathrm{P}}{\mathrm{P_{ref}}} = 1 + P_W \left(\frac{W_R}{W} - 1 \right) + P_L \left(\frac{L_R}{L} - 1 \right) + P_C \left(\frac{W_R L_R}{WL} - 1 \right). \tag{5.22}$$

5.2.3 Geometry scaling rules for parasitic and thermal resistance

A parasitic resistance models a current flow path, which can be vertical, lateral or both vertical and lateral, from external to internal nodes of a bipolar device. The emitter resistance `RE` models vertical current flow path from E to $E1$ node as shown in Fig. 2.1. It scales like the reciprocal of the current as the reciprocal

of the right part of (5.18) which becomes

$$\frac{\mathrm{P}}{\mathrm{P_{ref}}} = \left(1 + P_A\left(\frac{WL}{W_R L_R} - 1\right) + P_W\left(\frac{W}{W_R} - 1\right) + P_L\left(\frac{L}{L_R} - 1\right)\right)^{-1}. \tag{5.23}$$

Thermal resistance models the heat flow path, which depends on a 3D box region of the power dissipation [31], from the heat source to the heat sink. So, the geometry scaling of thermal resistance is also modeled with (5.23) as a thermal conductance with bulk and sidewall components.

The `RBV` models lateral current flow path from $B1$ to $B2$ node and its geometry scaling has been given in (4.27) based on a 2D simulation. As a result, its reference based scaling equation will be

$$\frac{\mathrm{P}}{\mathrm{P_{ref}}} = \frac{WL_R}{W_R L}\frac{(\frac{1}{12} - (\frac{1}{12} - \frac{1}{28.45})\frac{W}{L})}{(\frac{1}{12} - (\frac{1}{12} - \frac{1}{28.45})\frac{W_R}{L_R})}, \tag{5.24}$$

where $\rho_\square$ in (4.27) is replaced by the $\mathrm{P_{ref}}$.

The `RBC` models lateral current flow path from B to $B1$ node as shown in Fig. 5.4. Since the low-resistivity external base resistance behaves like a contact that surrounds the internal base, `RBC` is modeled with its conductance per length at its width (G_W), length (G_L) and corner (G_C) as

$$\mathrm{P} = (G_W W + G_L L + G_C)^{-1}. \tag{5.25}$$

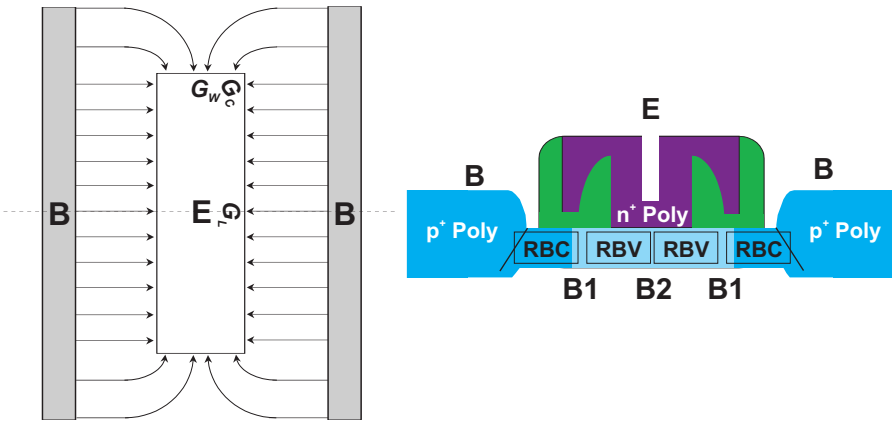
The reference based scaling equation of (5.25) will be

$$\frac{\mathrm{P}}{\mathrm{P_{ref}}} = \left(1 + P_W\left(\frac{W}{W_R} - 1\right) + P_L\left(\frac{L}{L_R} - 1\right)\right)^{-1}. \tag{5.26}$$

The collector resistance `RCC` that models both vertical and lateral current flow path from C to $C1$ node, comprises part of the un-depleted epi-collector resistance ($Rcep$) and buried-layer resistance under the emitter ($Rbli$), the extrinsic buried-layer resistance ($Rblx$) and the collector-plug resistance ($Rcpl$), as shown in Fig. 5.5. It is expressed as

$$\begin{aligned}\mathrm{P} &= Rcep + Rbli + Rblx + Rcpl \\ &= Rci + Rcx.\end{aligned} \tag{5.27}$$

where Rci is composed of vertical $Rcep$ and lateral $Rbli$ under emitter and Rcx is composed of $Rblx$ and $Rcpl$ next to emitter. As a result, Rci is modeled



(a) (b)

Figure 5.4: (a) top view and (b) cross-section of E-B regions of a SiGe HBT from external (B) to internal Base (B2) node.

with one over area $(1/WL)$ and square (W/L) of the emitter size as

$$\begin{aligned} Rci &= Rcep + Rbli \\ &= Rciv\frac{1}{WL} + Rcih\frac{W}{L}, \end{aligned} \tag{5.28}$$

where $Rciv$ and $Rcih$ are the constant geometry parameters for $Rcep$ and $Rbli$. From (5.28), the reference based scaling equation of Rci becomes:

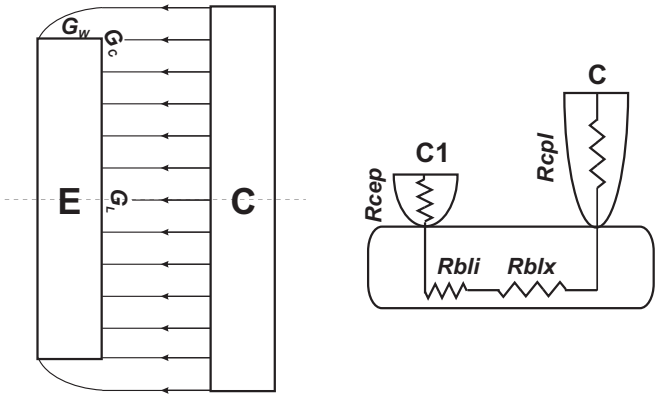
$$Rci = Rci_{ref}\left(P_I\frac{W_R L_R}{WL} + (1-P_I)\frac{WL_R}{W_R L}\right). \tag{5.29}$$

Where Rci_{ref} denotes Rci at the reference geometry and $P_I = \frac{Rciv}{Rci_{ref}W_R L_R}$ is the geometry parameter for Rci scaling. Notice also from Fig. 5.5 that Rcx is low-resistivity compared to Rci, so it is modeled with its conductance per length at its width (G_W), length (G_L) and corner (G_C) as

$$Rcx = (G_W W + G_L L + G_C)^{-1}. \tag{5.30}$$

Therefore, the reference based scaling equation for Rcx will be

$$Rcx = Rcx_{ref}\left(1 + P_W\left(\frac{W}{W_R} - 1\right) + P_L\left(\frac{L}{L_R} - 1\right)\right)^{-1}. \tag{5.31}$$



(a) (b)

Figure 5.5: (a) top view and (b) cross-section of a SiGe HBT from external (C) to internal collector (C1) node.

where Rcx_{ref} is the Rcx at the reference geometry, P_W and P_L are the width and length geometry parameters for Rcx scaling. From (5.27), (5.29) and (5.31), the referenced based scaling equation of the `RCC` for a device with one collector contact ($N_C = 1$) will be

$$\frac{\mathrm{P}}{\mathrm{P_{ref}}} = \left((1-P_X)\left(P_I\frac{W_R L_R}{WL} + (1-P_I)\frac{WL_R}{W_R L}\right) + P_X\left(1 + P_W\left(\frac{W}{W_R}-1\right) + P_L\left(\frac{L}{L_R}-1\right)\right)^{-1}\right), \tag{5.32}$$

where $P_X = \frac{Rcx_{ref}}{\mathrm{RCC_{ref}}}$ is the geometry parameter for `RCC` scaling.

5.3 Model implementation and parameter extraction

Since the temperature scaling rules and parameters are part of the standard Mextram model already, only the geometry scaling rules and parameters have to be added in the scalable model. It can be done easily by extending the standard Mextram model implementation in Verilog-A in Chapter 3 with an additional geometry scaling module and additional parameters as follows:

```
‘include ”frontdef.inc”
‘define SELFHEATING
‘define SUBSTRATE

module bjt504t_va (c, b, e, s, dt);
//External ports
inout c, b, e, s, dt;
electrical c, b, e, s, dt;
// Internal nodes
electrical c1, c2, b1, b2, e1;
electrical noi;
‘include ”parameters.inc”          //New instant and geometry
                                     parameters are added here.
‘include ”variables.inc”           //New variables are added here.
analog begin
  ‘include ”geo_scaling.inc”       //Geometry scaling rules are added
                                     here. See Appendix E for details.
  ‘include ”initialize.inc”
  ‘include ”tscaling.inc”
  ‘include ”evaluate.inc”
end
endmodule
```

In the implemented model above, new instance parameters (`WE`, `LE`) and geometry parameters (`ISA`, `ISW`, `ISL` etc.) are added in the module ”parameters.inc” to account for the drawn emitter size (W_E, L_E) and geometry scaling of the bipolar devices. The scalable electrical Mextram model parameters (`IS`, `IK`, `IBF` etc.) in ”parameters.inc” are used as the reference parameters. The scalable electrical parameters to their corresponding geometry dependent and scaling rules derived in the previous section are shown in Table 5.1 and implemented in ”geo_scaling.inc”, which is listed in Appendix E. Notice from Table 5.1 that the scaling rules are functions of the transistor's emitter, base and collector's drawn width (W_E, W_B, W_C) and length (L_E, L_B, L_C) as labeled in Fig. 5.6 depending on which physical quantities and junctions they are representing. The W_B, L_B, W_C and L_C are usually calculated from W_E and L_E based on process design rules so there is no need to put them as in-

Table 5.1: Scaling rules for the scalable Mextram model parameters.

Electrical Parameters	Geometry Dependent	Scaling Rules
IS, IK, IBF, IBR	W_E, L_E	(5.18)
CJE, CJC, IHC	W_E, L_E	(5.18)
ISS, IKS	W_B, L_B	(5.18)
CJS	W_C, L_C	(5.18)
BF, VEF, VER	W_E, L_E	(5.22)
PE, PC, TAUB	W_E, L_E	(5.22)
BRI	W_B, L_B	(5.22)
PS	W_C, L_C	(5.22)
RE, RCV, SCRCV	W_E, L_E	(5.23)
RTH	W_B, L_B	(5.23)
RBV	W_E, L_E	(5.24)
RBC	W_E, L_E	(5.26)
RCC	W_E, L_E	(5.32)

stant parameters. Following "parameter.inc", some new variables are added in the module "variables.inc" and used in the module "geo_scaling.inc" for the geometry scaling results of the scaling rules. They will replace the scalable electrical model parameters in the temperature-scaling module "tscaling.inc", where the temperature scaling results will be used later in the main module "evaluate.inc". Therefore, the geometry and temperature scaling of the bipolar transistor are taken into account in the model evaluation.

Parameter extraction for the temperature and geometry scalable bipolar transistor model Mextram is demonstrated on a high-speed SiGe HBT technology as used for single model parameter extraction in Chapter 4. Its available device geometry matrix, which includes five different emitter widths (W_1, W_2, W_{ER}, W_3, W_4= 0.2, 0.3, 0.4, 0.6, 0.9 μm) and lengths (L_{EMIN}, L_1, L_2, L_3, L_{ER}= 1.7, 2.64, 4.52, 8.28, 10.16 μm) for parameter extraction, is shown in Fig. 5.7. In Fig. 5.8, a flow chart of a unified reference, temperature and geometry parameter extraction procedure for the scalable model is shown. It has been implemented in an IC-CAP model file. In order to extract geometry parameters in IC-CAP, the emitter width and length of each DUT have to be added in the standard measurement data files as additional input parameters, which will pass to a circuit simulator e.g. ADS as instant parameters. The parameter extraction starts from a reference device's parameter extraction. The reference device should be decided so that the bulk, sidewall and corner

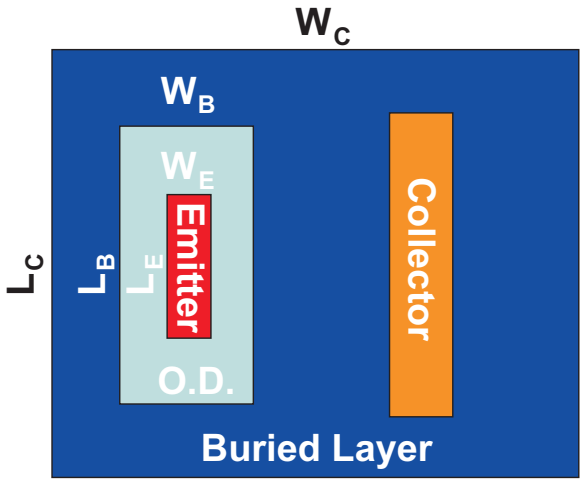


Figure 5.6: Drawn size in the device layout.

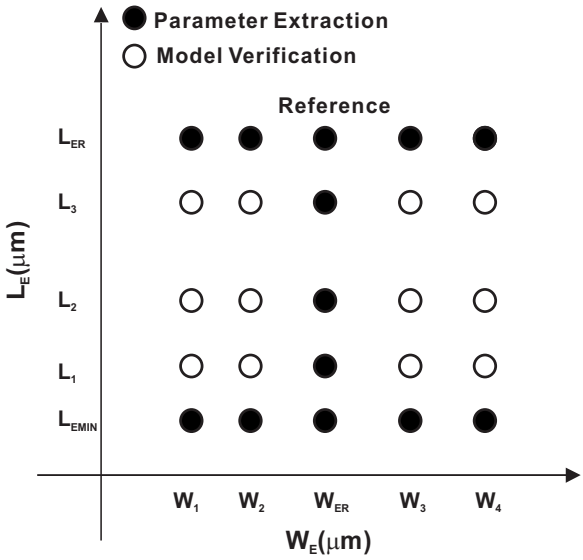


Figure 5.7: Devices' geometry matrix including five different emitter widths (W_1, W_2, W_{ER}, W_3, W_4= 0.2, 0.3, 0.4, 0.6, 0.9 μm) and lengths (L_{EMIN}, L_1, L_2, L_3, L_{ER}= 1.7, 2.64, 4.52, 8.28, 10.16 μm) for parameter extraction.

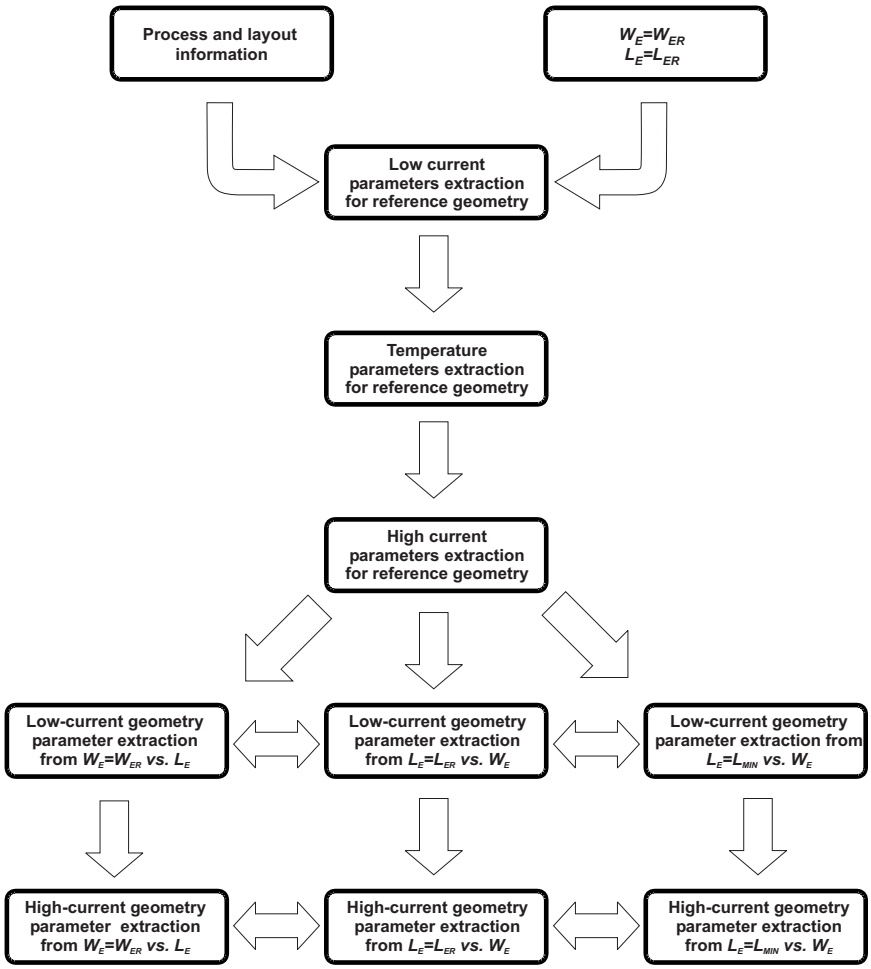


Figure 5.8: Flow chart of a unified parameter extraction procedure.

components contribute considerably to the device characteristics. As a result, it is chosen as the width is in the middle and length is the maximal of the matrix labeled in Fig. 5.7, which is the device with W_E= 0.4 μm, L_E= 10.16 μm, N_B =2 and N_C =1 used in Chapter 4 for single device model parameter extraction. The reference and temperature parameter extraction follows the single device parameter extraction in Chapter 4. The drawn emitter width, length, sheet resistance of internal and external base, number of base stripes, epi-collector doping level and epi-collector thickness are specified as technology parameters for initial parameter calculation to assist the numerical optimizer in finding the optimum parameter values during parameter extraction.

After extraction of reference and temperature parameters, geometry parameters can be extracted from the same measurement setup at various geometries in the following ways. The devices from two rows and one column in the geometry matrix used for geometry parameters extraction are marked in Fig. 5.7. They are chosen because when $L_E = L_{ER}$, P_L will be removed from the geometry scaling rule (5.18), which is then simplified to

$$\frac{\mathrm{P}}{\mathrm{P_{ref}}} = 1 + (P_A + P_W)\left(\frac{W_E}{W_{ER}} - 1\right), \tag{5.33}$$

when $W_E = W_{ER}$, (5.18) is changed to

$$\frac{\mathrm{P}}{\mathrm{P_{ref}}} = 1 + (P_A + P_L)\left(\frac{L_E}{L_{ER}} - 1\right), \tag{5.34}$$

where P_W is removed. And when $L_E = L_{MIN}$, (5.18) becomes:

$$\frac{\mathrm{P}}{\mathrm{P_{ref}}} = \left(P_A\frac{L_{MIN}}{L_{ER}} + P_W\right)\frac{W_E}{W_{ER}} + C, \tag{5.35}$$

where $C = 1 - P_A - P_W + P_L(\frac{L_{MIN}}{L_{ER}} - 1)$ is a constant. Since $L_{MIN} << L_{ER}$, P_A is scaled down by a factor $\frac{L_{MIN}}{L_{ER}}$ in (5.35). So, P_W is more significant in (5.35) than (5.33). In these three special cases, low-current to high-current geometry parameters can be extracted from their specific measurement setups, while keeping $\mathrm{P_{ref}}$ un-touched, in the following three extraction steps: (I) P_A is extracted from $L_E = L_{ER}$ v.s. W_E where the bulk component is the largest, (II) P_L is extracted from $W_E = W_{ER}$ v.s. L_E where the sidewall length component is the largest and (III) P_W is extracted from $L_E = L_{MIN}$ v.s. W_E where the sidewall width component is the largest. It usually takes only two to three repeat sequences of extraction for the optimizer to find the optimum P_A, P_W and P_L.

The essential feature of the proposed unified parameter extraction procedure is a direct extraction of the geometry parameters from the measured electrical characteristics and the Mextram model parameters as reference parameters are extracted only once for a reference geometry. The complete parameter extraction procedure is integrated in the single environment, which save time in verification of the final scalable model results with measured data. Besides, when setting $W_E = W_{ER}$ and $L_E = L_{ER}$, the scalable model is also backward compatible with the original Mextram model for using it as a single device model.

Fig. 5.9 $\sim$ Fig. 5.14 show the fitting plots of C_{BE}, C_{BC}, forward Gummel, $I_C/exp(V_{BE}/V_T)$, reverse Gummel, f_T and S-parameters of (a) $L_E = L_{ER}$ v.s. W_E (b) $W_E = W_{ER}$ v.s. L_E and (c) $L_E = L_{EMIN}$ v.s. W_E devices after geometry parameter extraction. As can be seen, simulated results in all the figures from ADS fit well to the measured results.

5.4 Conclusions

In this chapter, a reference geometry based scaling approach and its parameter extraction is proposed for the bipolar transistor model Mextram. It is based on the physical properties of Mextram parameters and scaled with the layout geometry of a bipolar transistor. The scalable electrical parameter, temperature and geometry in the scaling rules are normalized to a reference parameter, temperature and geometry. The scalable model is implemented in AHDL Verilog-A language, which can be used in many commercial simulators [74, 123–125] that offer IP protection. Along with the scalable model, a unified parameter extraction procedure for reference, temperature and geometry parameters of the scalable model has been implemented in an IC-CAP model file and demonstrated on a of high-speed SiGe HBT technology. No additional tool is needed for temperature and geometry parameter extraction and the resulting model shows good fit to CV, DC, f_T and S-parameters measured from high-speed SiGe HBTs. Therefore, the scalable model library generation is efficient and the accuracy is increased with the new scaling approach. Since the geometry scaling approach is based on the scaling of the physical properties, it is believed that it can be applied to the Gummel-Poon, HICUM and VBIC models as well.

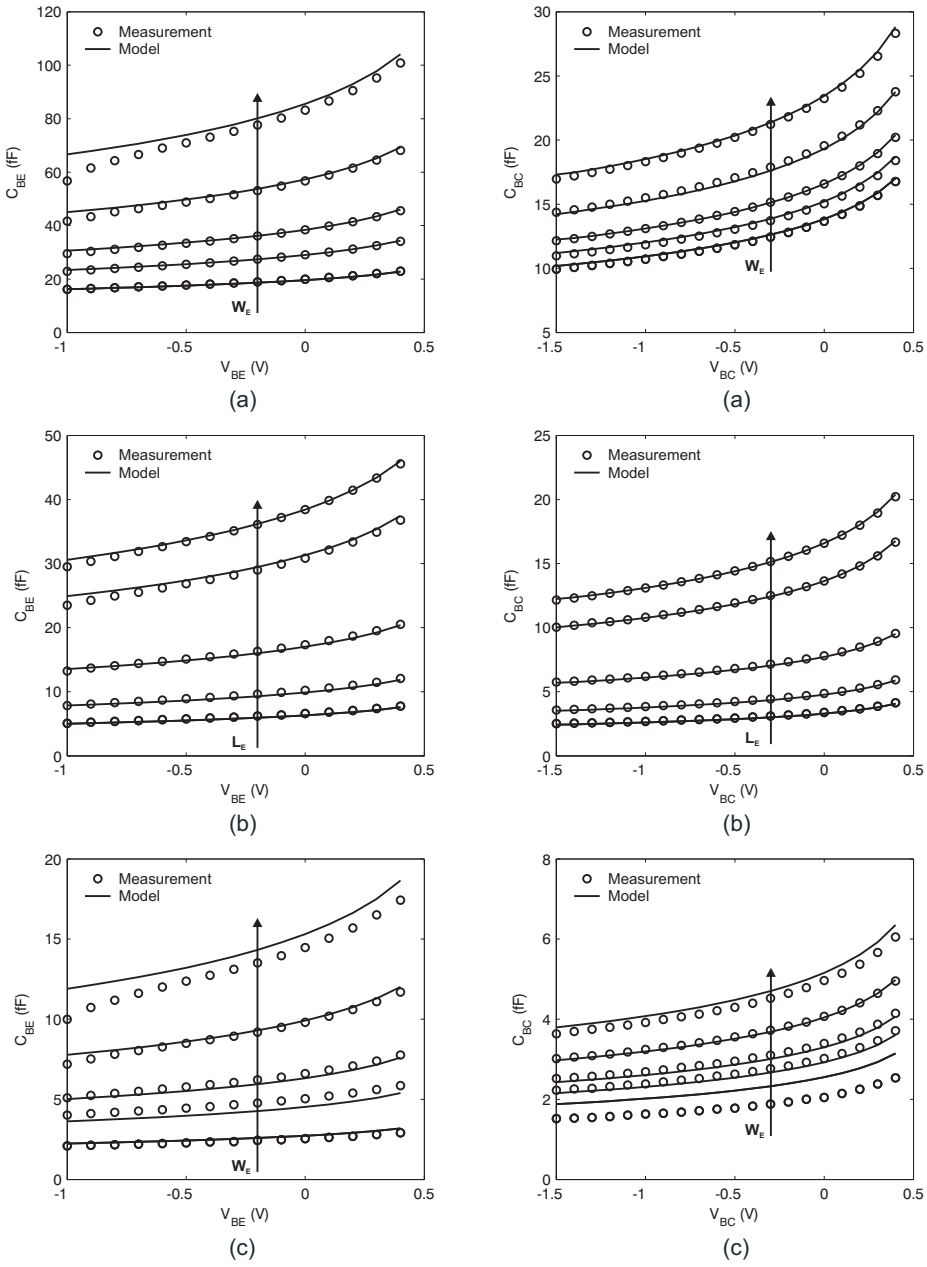


Figure 5.9: Measured and simulated C_{BE} and C_{BC} of (a) $L_E = L_{ER}$ v.s. $W_E\uparrow$ (b) $W_E = W_{ER}$ v.s. $L_E\uparrow$ and (c) $L_E = L_{EMIN}$ v.s. $W_E\uparrow$ devices at $T = T_R = 25$ °C.

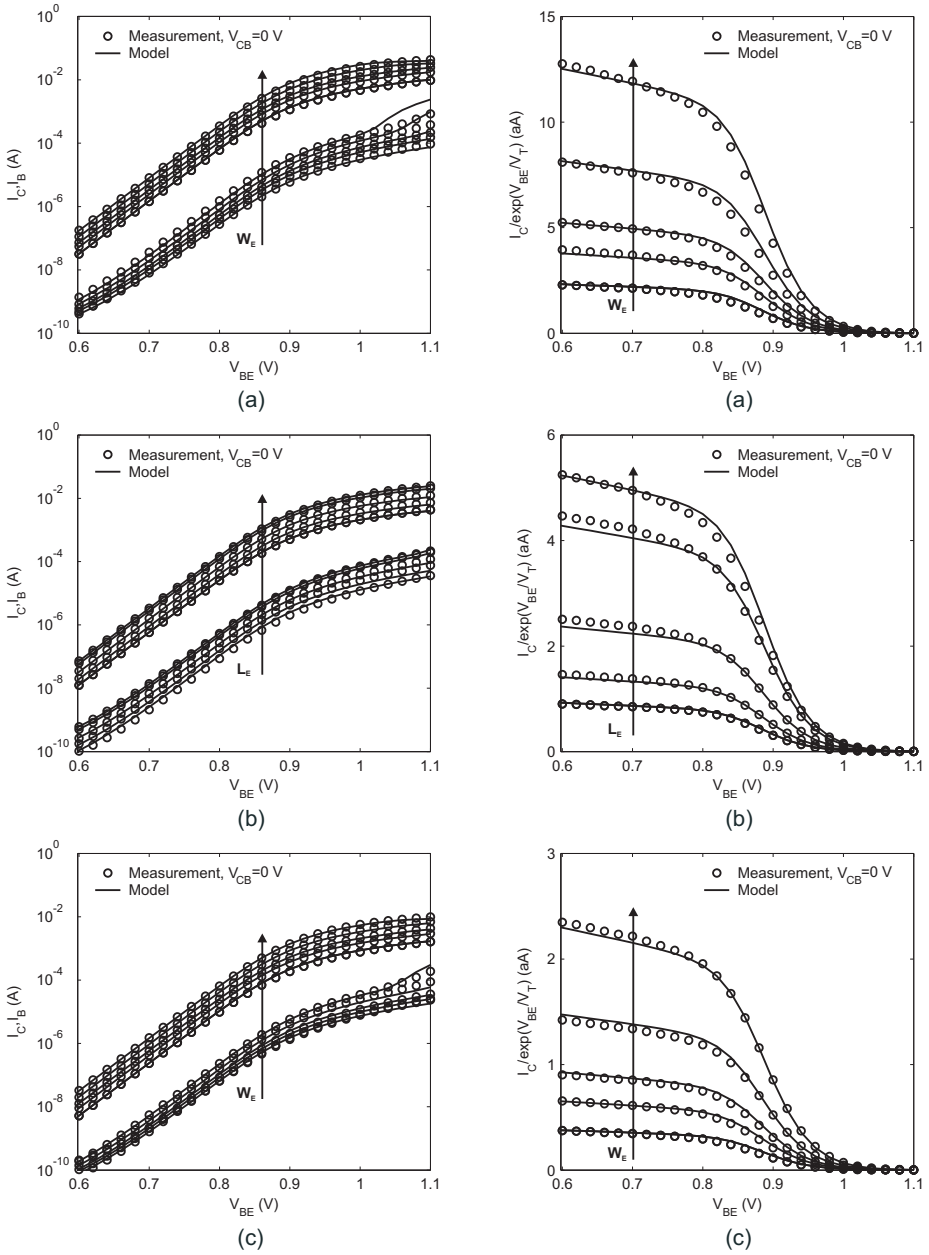


Figure 5.10: Measured and simulated forward Gummel plot and $I_C/exp(V_{BE}/V_T)$ of the forward Gummel plot of (a) $L_E = L_{ER}$ v.s. $W_E\uparrow$ (b) $W_E = W_{ER}$ v.s. $L_E\uparrow$ and (c) $L_E = L_{EMIN}$ v.s. $W_E\uparrow$ devices at $T = T_R = 25$ ^{o}C.

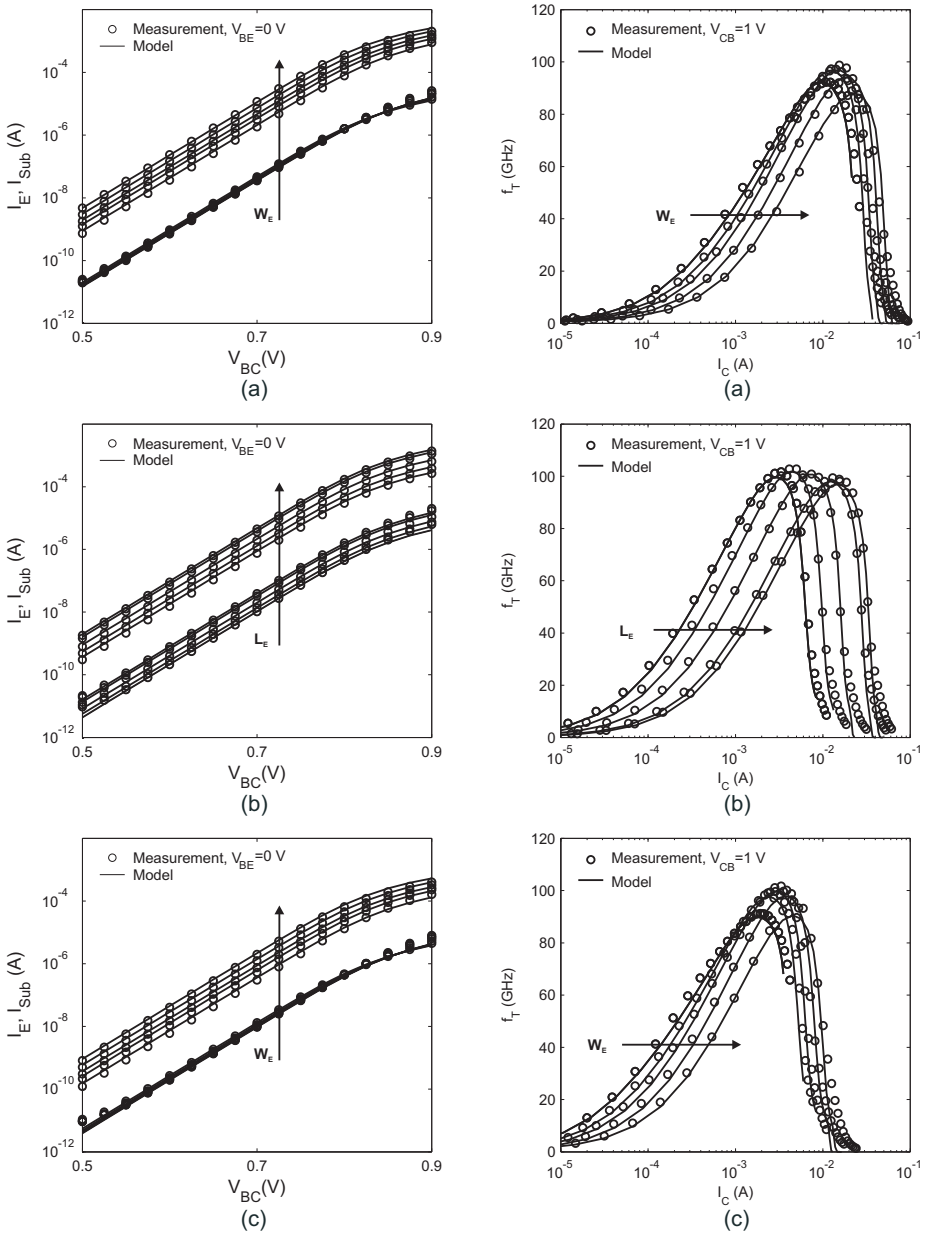


Figure 5.11: Measured and simulated reverse Gummel plot and f_T v.s. I_C of (a) $L_E = L_{ER}$ v.s. $W_E\uparrow$ (b) $W_E = W_{ER}$ v.s. $L_E\uparrow$ and (c) $L_E = L_{EMIN}$ v.s. $W_E\uparrow$ devices at $T = T_R = 25$ °C.

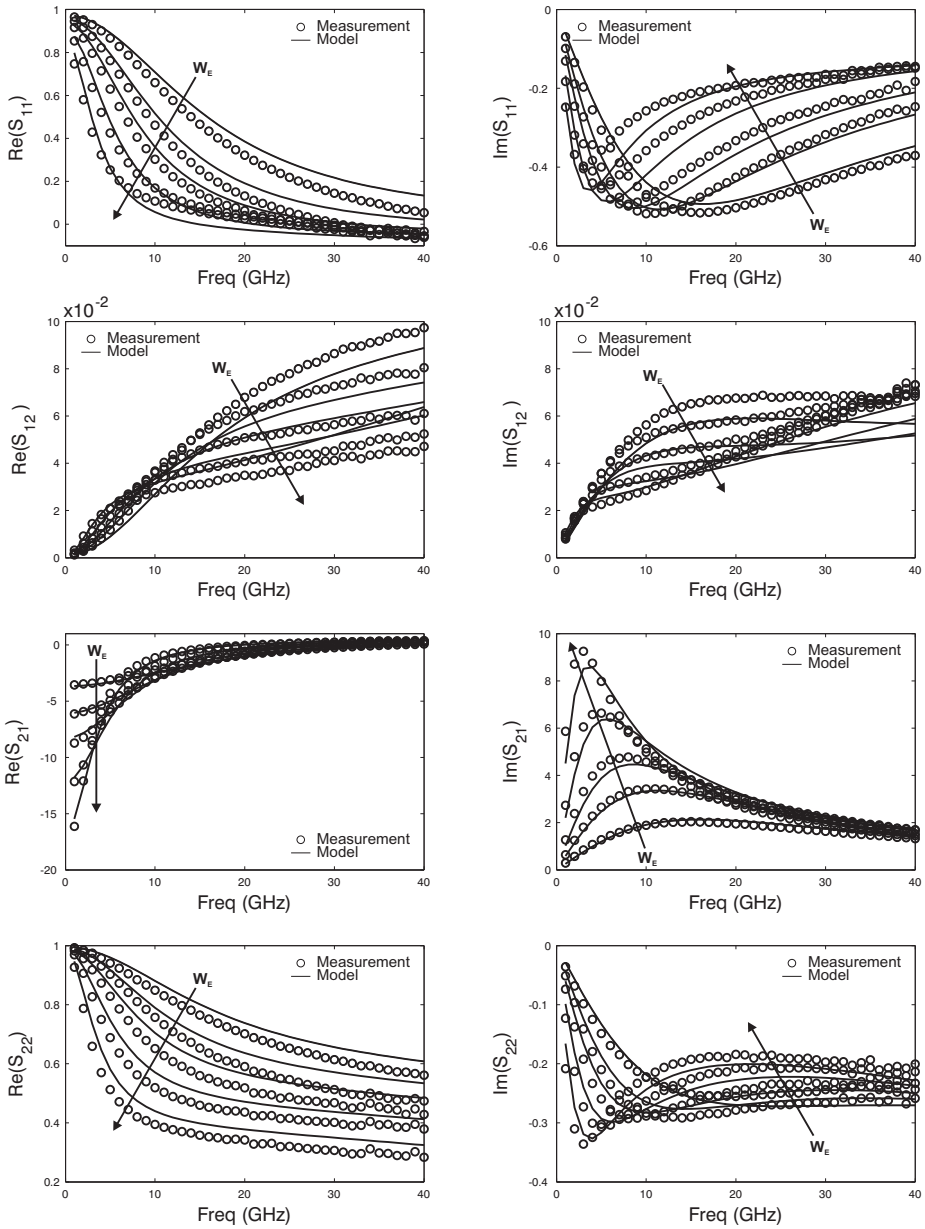


Figure 5.12: Measured and simulated S-parameters v.s. frequency (1∼40 GHz) of $L_E = L_{ER}$ v.s. $W_E\uparrow$ devices biased at $V_{BE} = 0.95$ V and $V_{CE} = 2$ V and at $T = T_R = 25$ °C.

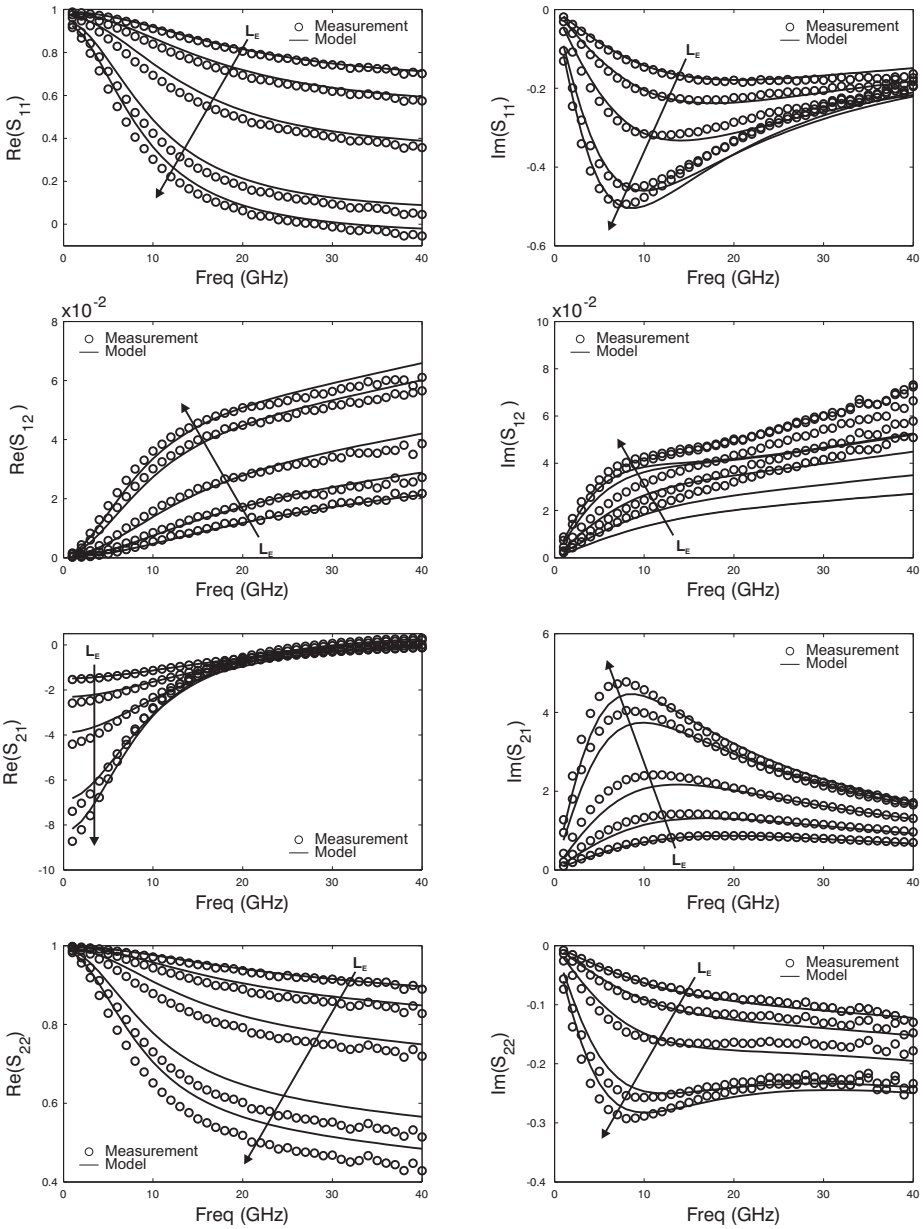


Figure 5.13: Measured and simulated S-parameters v.s. frequency (1∼40 GHz) of $W_E = W_{ER}$ v.s. $L_E\uparrow$ devices biased at $V_{BE} = 0.95$ V and $V_{CE} = 2$ V and at $T = T_R = 25$ °C.

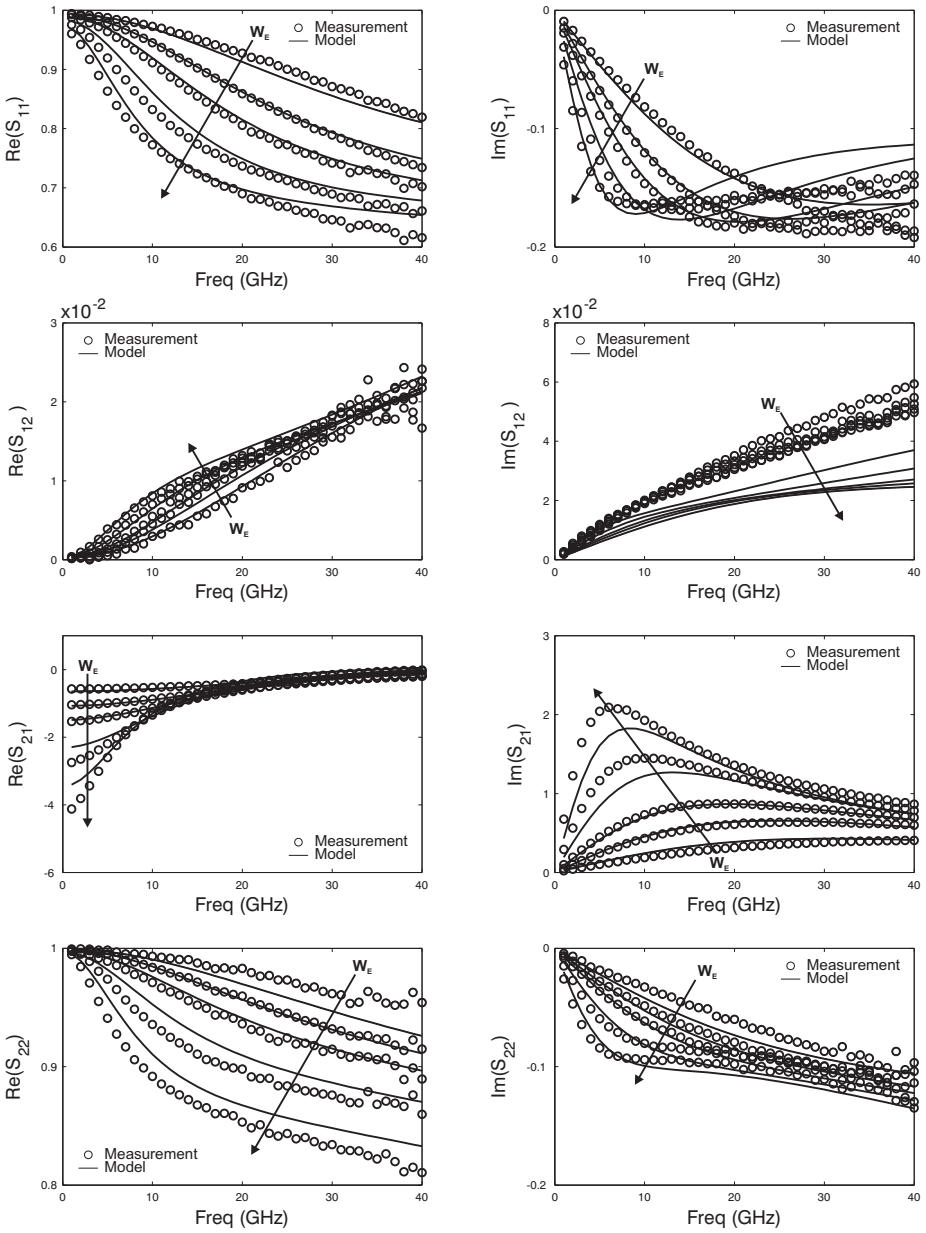


Figure 5.14: Measured and simulated S-parameters v.s. frequency (1∼40 GHz) of $L_E = L_{MIN}$ v.s. $W_E\uparrow$ devices biased at $V_{BE} = 0.95$ V and $V_{CE} = 2$ V and at $T = T_R = 25$ °C.

Chapter 6

Mextram modeling for high power application

6.1 Introduction

With the coming era of System-on-a-Chip (SoC) there is an ever increasing number of microcircuits and systems integrated on a single chip. In order to deal with the increasing complexity in circuit design, a geometry scalable model is needed for different function blocks in SoC design. However, simple dimensional device scaling is not sufficient to meet all the circuit design requirements. For example, power amplifier applications require larger emitter area (A_E) to provide sufficient power output but emitter area can't be extended without limits. When the device becomes large, the emitter current crowding effect and the distributed parasitic resistance will degrade the device performance [126, 127]. In order to provide a high output power level without sacrificing too much device performance, it is common practice to employ several transistors in parallel or create a Multi-Emitter Device (MED) in power amplifier design. Between these two options, a MED is preferred for power amplifier application because of its compactness in using minimum chip area.

The advanced bipolar transistor model Mextram, which includes various high current bipolar transistor effects [8, 55], is especially suitable for high-power applications. In this chapter, we would like to extend the previously developed reference based geometry scaling approach for the Mextram model to configuration scaling including different contact configurations and emitter numbers for optimum circuit design.

The chapter is organized as follows. In Section 6.2, scaling equations for MEDs, based on different layout configurations, are derived. The scalable model implementation and an IC-CAP model file for geometry parameter extraction from different layout configurations is introduced in Section 6.3. The conclusion of this chapter is in Section 6.4.

6.2 Derivation of scaling equations for multi-emitter devices

6.2.1 Scaling of intrinsic transistor parameters

Fig. 6.1 shows the layout of a multi-emitter SiGe HBT with its emitter (E) and collector (C) labeled. The electrical model parameters that describe currents and capacitances of the intrinsic device (`IS`, `IK`, `IBF` etc.) are simply scaled with the emitter number while the ratio parameters (`BF`, `VEF`, `VER` etc.) are kept identical to the Single-Emitter Device (SED). The resistances that define the current flow in the direction vertical to the emitter window (`RE`, `RCV` and `SCRCV`) and base resistances (`RBV` and `RBC`) scale reciprocally to the number of the emitters.

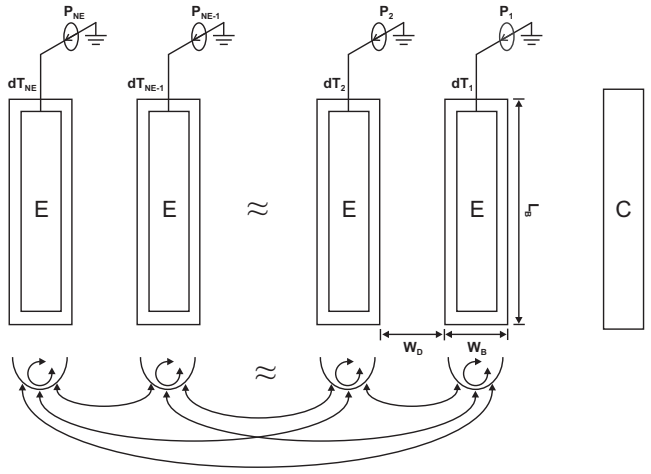


Figure 6.1: Layout of a SiGe HBT with N_E emitters and a single collector contact with self-heating and mutual-heating among emitters due to power consumption at each emitter.

6.2.2 Thermal resistance scaling

The self- and mutual-heating of emitters are also schematically presented in Fig. 6.1. It produces a non-uniform distribution of the temperature and in a real device each emitter operates at a different temperature [128]. To represent the real temperature distribution in the MED, there should be a thermal node associated with each emitter. However, it is not the case in Mextram so it is chosen to model with only one thermal node as it is in Mextram now due to simulation speed and model simplicity concern. Therefore, only the average temperature rise of the device is considered and it is assumed to be uniform across the active device region, which is the region isolated by the shallow trench isolation.

For each emitter, the temperature rise (dT_i) produced by the self- and mutual-heating effects can be modeled with an $N_E \times N_E$ thermal resistance matrix and its power dissipation P_i [129] as

$$\begin{bmatrix} dT_1 \\ dT_2 \\ \vdots \\ dT_{N_E} \end{bmatrix} = \begin{bmatrix} Rth_{1,1} & Rth_{1,2} & \cdots & Rth_{1,N_E} \\ Rth_{2,1} & Rth_{2,2} & \cdots & Rth_{2,N_E} \\ \vdots & \vdots & \vdots & \vdots \\ Rth_{N_E,1} & Rth_{N_E,2} & \cdots & Rth_{N_E,N_E} \end{bmatrix} \begin{bmatrix} P_1 \\ P_2 \\ \vdots \\ P_{N_E} \end{bmatrix} \tag{6.1}$$

The components in the thermal resistance matrix are modeled as

$$Rth_{i,j} = \begin{cases} Rth & i = j, \\ \frac{Rthc}{abs(i-j)W_D W_B L_B} & i \neq j. \end{cases} \tag{6.2}$$

where Rth and $Rthc$ are self- and mutual-heating thermal resistance parameters, W_D, W_B and L_B as labeled in Fig. 6.1 are the fixed distance of adjacent active regions, width and length of the active region respectively.

The Rth scales with the reciprocal of the active region area so its reference based geometry scaling equation based on (5.23) is expressed as

$$\begin{aligned} Rth = Rth_{ref} \Bigg(1 + \texttt{RTHA} \left(\frac{W_B L_B}{W_{BR} L_{BR}} - 1 \right) + \texttt{RTHW} \left(\frac{W_B}{W_{BR}} - 1 \right) \\ + \texttt{RTHL} \left(\frac{L_B}{L_{BR}} - 1 \right) \Bigg)^{-1}, \end{aligned} \tag{6.3}$$

where Rth_{ref} is the self-heating thermal resistance at a reference geometry and W_{BR} and L_{BR} are the reference width and the reference length of single-emitter active region. The geometry parameters `RTHA`, `RTHW` and `RTHL` account

for bulk, sidewall width and sidewall length contributions. When the device geometry is equal to the reference geometry, Rth is equal to Rth_{ref} .

If P_i, which represents the power dissipation at the finger i, is assumed to be equal to the total power dissipation P divided by N_E, the average temperature rise dT in each emitter finger will be:

$$\begin{aligned} dT &= \frac{dT_1 + dT_2 + \cdots + dT_{N_E}}{N_E} \\ &= \frac{P}{N_E^2}\Big(Rth_{1,1} + Rth_{1,2} + \cdots + Rth_{1,N_E} + Rth_{2,1} + Rth_{2,2} + \cdots + Rth_{2,N_E} \\ &\quad + \cdots + Rth_{N_E,1} + Rth_{N_E,2} + \cdots + Rth_{N_E,N_E} \Big). \end{aligned} \tag{6.4}$$

By using (6.2) in (6.4), the effective total thermal resistance (RTH) is expressed as

$$\begin{aligned} \mathrm{RTH} &= \frac{dT}{P} \\ &= \frac{1}{N_E}\left(Rth + \frac{2Rthc}{N_E W_D W_B L_B} \sum_{i=1}^{N_E-1} \frac{N_E - i}{i} \right). \end{aligned} \tag{6.5}$$

Notice, that from (6.5) when $N_E \geq 2$, RTH is increased due to the mutual heating effect.

From (6.3) and (6.5), the reference base scaling equations for the SED and MED's RTH, which includes self- and mutual-heating thermal resistance, becomes:

$$\begin{aligned} \mathrm{RTH} = \frac{Rth_{ref}}{N_E}\Bigg(\Bigg(1 + \mathtt{RTHA}\left(\frac{W_B L_B}{W_{BR} L_{BR}} - 1 \right) + \mathtt{RTHW}\left(\frac{W_B}{W_{BR}} - 1 \right) \\ + \mathtt{RTHL}\left(\frac{L_B}{L_{BR}} - 1 \right) \Bigg)^{-1} + \frac{\mathtt{2RTHC}}{N_E} \frac{W_{BR} L_{BR}}{W_B L_B} \sum_{i=1}^{N_E-1} \frac{N_E - i}{i} \Bigg), \end{aligned} \tag{6.6}$$

where $\mathtt{RTHC}=\frac{Rthc}{Rth_{ref} W_D W_{BR} L_{BR}}$ is the geometry parameter for the mutual heating thermal resistance.

6.2.3 Collector resistance scaling

Similarly to the temperature distribution, the potential distribution along the lateral collector buried-layer of a MED will be non-uniform because of the

lateral current flow. However, it is chosen to be modeled by a single internal collector node of the Mextram model due to simulation speed and model simplicity concern. Therefore, the total collector resistance (RCC) is modeled effectively. Starting from the SED with one collector contact, the reference based scaling equation will be

$$\mathrm{RCC} = \mathrm{RCC}_{\mathrm{ref}}\Bigg((1-\mathtt{RCX})\left(\mathtt{RCI}\frac{W_{ER}L_{ER}}{W_EL_E}+(1-\mathtt{RCI})\frac{W_EL_{ER}}{W_{ER}L_E}\right) + \mathtt{RCX}\left(1+\mathtt{RCW}\left(\frac{W_E}{W_{ER}}-1\right)+\mathtt{RCL}\left(\frac{L_E}{L_{ER}}-1\right)\right)^{-1}\Bigg), \tag{6.7}$$

where (6.7) comes from (5.32), $\mathrm{RCC}_{\mathrm{ref}}$ is RCC at a reference geometry, and RCI, RCX, RCW and RCL are geometry parameters.

When extending the scaling from SED to MED, the extra current flow path between the emitter fingers has to be taken into account in the RCC calculation. First, the *Rbli*, which represents the current flow path under the emitter, is modeled with a T-network as shown in the dashed box of Fig. 6.2 to connect different emitters. When the lateral current flow through *Rbli* in the buried layer turns into each emitter finger, its effective resistance will become *Rbli*/3 for a one-sided connection and *Rbli*/12 for a double-sided connection with the -*Rbli*/6 reduction resulting from the distributed resistance effect [130]. Secondly, a *Rblc* is used to model the buried-layer resistance in adjacent emitter fingers. The final impedance network of the MED from C1 to C is shown in Fig. 6.3, where the $Rcd = Rblc + Rbli$ is the combination of *Rblc* and *Rbli* between two adjacent emitters and *Rcx* is equal to *Rblx* plus *Rcpl* as defined in (5.27).

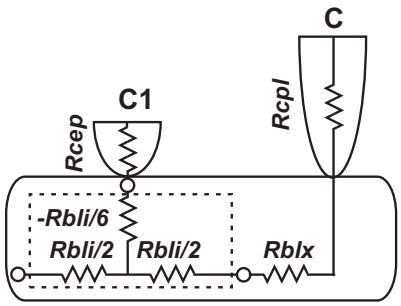


Figure 6.2: Equivalent T-network to model the 2-D current flow of *Rbli*.

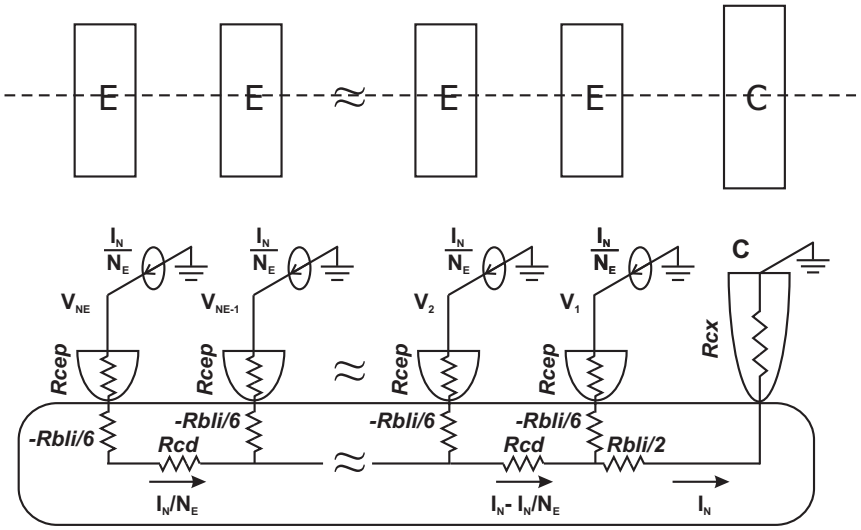


Figure 6.3: Top view and collector impedance network of a MED with single collector contact.

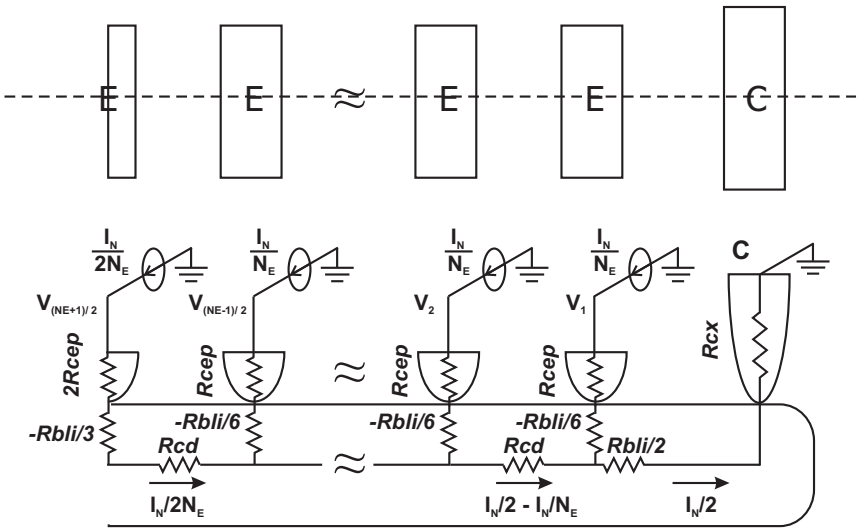


Figure 6.4: Top view and collector impedance network of an odd-emitter, two collector contacts MED.

If the total current (I_N) is equally distributed among the emitter fingers as shown in Fig. 6.3, the internal node voltage V_i can be modeled with an $N_E \times 3$ resistance matrix times three nodal currents as

$$\begin{bmatrix} V_1 \\ V_2 \\ \vdots \\ V_{N_E-1} \\ V_{N_E} \end{bmatrix} = \begin{bmatrix} Rcep - \frac{Rbli}{6} & 0 & \frac{Rbli}{2} + Rcx \\ Rcep - \frac{Rbli}{6} & (N_E - 1)Rcd & \frac{Rbli}{2} + Rcx \\ \vdots & \vdots & \vdots \\ Rcep - \frac{Rbli}{6} & \sum_{i=1}^{N_E-2} (N_E - i)Rcd & \frac{Rbli}{2} + Rcx \\ Rcep - \frac{Rbli}{6} & \sum_{i=1}^{N_E-1} (N_E - i)Rcd & \frac{Rbli}{2} + Rcx \end{bmatrix} \begin{bmatrix} \frac{I_N}{N_E} \\ \frac{I_N}{N_E} \\ I_N \end{bmatrix} \quad (6.8)$$

The effective RCC is defined as the average nodal voltage V divided by I_N as

$$\begin{aligned} \text{RCC} = \frac{V}{I_N} &= \frac{V_1 + V_2 + \cdots + V_{N_E-1} + V_{N_E}}{I_N N_E} \\ &= \frac{Rcep}{N_E} - \frac{Rbli}{6N_E} + \sum_{i=1}^{N_E-1} i^2 \frac{Rcd}{N_E^2} + \frac{Rbli}{2} + Rcx \\ &= \frac{Rcep}{N_E} + \frac{3N_E - 1}{6N_E} Rbli + \frac{(N_E - 1)(2N_E - 1)}{6N_E}(Rbli + Rblc) + Rcx \\ &= \frac{Rcep}{N_E} + \frac{N_E}{3} Rbli + \frac{(N_E - 1)(2N_E - 1)}{6N_E} Rblc + Rcx. \end{aligned} \quad (6.9)$$

The analysis for a two collector contacts MED is similar to that of a single collector contact MED due to the symmetry of the device layout. In that case, it is sufficient to consider half of the device as shown in Fig. 6.4, which is half of a MED with odd emitters. Its nodal voltages from V_1 to $V_{(N_E+1)/2}$ can be modeled with an $(N_E+1)/2 \times 3$ resistance matrix times the nodal currents as

$$\begin{bmatrix} V_1 \\ V_2 \\ \vdots \\ V_{\frac{N_E-1}{2}} \\ V_{\frac{N_E+1}{2}} \end{bmatrix} = \begin{bmatrix} Rcep - \frac{Rbli}{6} & 0 & \frac{Rbli}{2} + Rcx \\ Rcep - \frac{Rbli}{6} & (\frac{N_E}{2} - 1)Rcd & \frac{Rbli}{2} + Rcx \\ \vdots & \vdots & \vdots \\ Rcep - \frac{Rbli}{6} & \sum_{i=1}^{\frac{N_E-3}{2}} (\frac{N_E}{2} - i)Rcd & \frac{Rbli}{2} + Rcx \\ Rcep - \frac{Rbli}{6} & \sum_{i=1}^{\frac{N_E-1}{2}} (\frac{N_E}{2} - i)Rcd & \frac{Rbli}{2} + Rcx \end{bmatrix} \begin{bmatrix} \frac{I_N}{N_E} \\ \frac{I_N}{N_E} \\ \frac{I_N}{2} \end{bmatrix} \quad (6.10)$$

By averaging the N_E nodal voltages, the RCC for an odd emitter with two collector contacts MED is:

$$\begin{aligned}
\text{RCC} = \frac{V}{I_N} &= \frac{2\left(V_1 + V_2 + \cdots + V_{\frac{(N_E-1)}{2}}\right) + V_{\frac{(N_E+1)}{2}}}{I_N N_E} \\
&= \frac{Rcep}{N_E} - \frac{Rbli}{6N_E} + \sum_{i=1}^{\frac{N_E-1}{2}} (\frac{N_E}{2} - i)(N_E - 2i)\frac{Rcd}{N_E^2} + \frac{Rbli}{4} + \frac{Rcx}{2} \\
&= \frac{Rcep}{N_E} + \frac{3N_E - 2}{12N_E} Rbli + \frac{(N_E-1)(N_E-2)}{12N_E}(Rbli + Rblc) + \frac{Rcx}{2} \\
&= \frac{Rcep}{N_E} + \frac{N_E}{12} Rbli + \frac{(N_E-1)(N_E-2)}{12N_E} Rblc + \frac{Rcx}{2}. \qquad (6.11)
\end{aligned}$$

For any even number of emitters with two collector contact, its internal nodal voltages calculation is the same as for the single collector contact with half of its emitters ($\frac{N_E}{2}$) and total nodal current ($\frac{I_N}{2}$) as

$$\begin{bmatrix} V_1 \\ V_2 \\ \vdots \\ V_{\frac{N_E}{2}-1} \\ V_{\frac{N_E}{2}} \end{bmatrix} = \begin{bmatrix} Rcep - \frac{Rbli}{6} & 0 & \frac{Rbli}{2} + Rcx \\ Rcep - \frac{Rbli}{6} & (\frac{N_E}{2} - 1)Rcd & \frac{Rbli}{2} + Rcx \\ \vdots & \vdots & \vdots \\ Rcep - \frac{Rbli}{6} & \sum_{i=1}^{\frac{N_E}{2}-2} (\frac{N_E}{2} - i)Rcd & \frac{Rbli}{2} + Rcx \\ Rcep - \frac{Rbli}{6} & \sum_{i=1}^{\frac{N_E}{2}-1} (\frac{N_E}{2} - i)Rcd & \frac{Rbli}{2} + Rcx \end{bmatrix} \begin{bmatrix} \frac{I_N}{N_E} \\ \frac{I_N}{N_E} \\ \frac{I_N}{2} \end{bmatrix} \qquad (6.12)$$

Its collector resistance will then be:

$$\begin{aligned}
\text{RCC} = \frac{V}{I_N} &= \frac{2\left(V_1 + V_2 + \cdots + V_{\frac{N_E}{2}-1} + V_{\frac{N_E}{2}}\right)}{I_N N_E} \\
&= \frac{Rcep}{N_E} - \frac{Rbli}{6N_E} + 2\sum_{i=1}^{\frac{N_E}{2}-1} i^2 \frac{Rcd}{N_E^2} + \frac{Rbli}{4} + \frac{Rcx}{2} \\
&= \frac{Rcep}{N_E} + \frac{3N_E - 2}{12N_E} Rbli + \frac{(N_E-1)(N_E-2)}{12N_E}(Rbli + Rblc) + \frac{Rcx}{2} \\
&= \frac{Rcep}{N_E} + \frac{N_E}{12} Rbli + \frac{(N_E-1)(N_E-2)}{12N_E} Rblc + \frac{Rcx}{2}. \qquad (6.13)
\end{aligned}$$

It turns out that the scaling equations for devices that have odd or even number of emitters with two collector contacts are the same. Therefore, RCC from (6.9), (6.11) and (6.13) can be formed as a single analytical equation with N_E and N_C as

$$\mathrm{RCC} = \frac{Rcep}{N_E} + \frac{N_E}{3N_c^2}Rbli + \frac{(N_E-1)\left(2\frac{N_E}{N_C^2}-1\right)}{6N_E}Rblc + \frac{Rcx}{N_C}. \tag{6.14}$$

Since N_E=1 and N_C=1 can be used as the reference emitter and collector contact number, the reference based scaling equation of (6.14) for the SED and MED with different N_E and N_C will become:

$$\begin{aligned}
RC_1 &= (1-\mathrm{RCX})\left(\frac{\mathrm{RCI}}{N_E}\frac{W_{ER}L_{ER}}{W_EL_E} + \frac{N_E}{N_C^2}(1-\mathrm{RCI})\frac{W_EL_{ER}}{W_{ER}L_E}\right),\\
RC_2 &= \frac{(N_E-1)\left(2\frac{N_E}{N_C^2}-1\right)}{6N_E}\mathrm{RCXC}\frac{L_{ER}}{L_E},\\
RC_3 &= \frac{\mathrm{RCX}}{N_C}\left(1+\mathrm{RCW}\left(\frac{N_E}{N_C}\frac{W_E}{W_{ER}}-1\right)+\mathrm{RCL}\left(\frac{L_E}{L_{ER}}-1\right)\right)^{-1},\\
\mathrm{RCC} &= \mathrm{RCC_{ref}}(RC_1+RC_2+RC_3).
\end{aligned} \tag{6.15}$$

Where RCXC is the geometry parameter that models the *Rblc*. The scaling equation (6.15) also fulfills the requirement for emitter area conservation, i.e. when assigning instant parameters $N_E = ne$, $W_E = \frac{W_E}{ne}$ and RCXC=0, $\mathrm{RCC}|_{N_E=1}=\mathrm{RCC}|_{N_E=ne,\ W_E=\frac{W_E}{ne},\ \mathrm{RCXC}=0}$.

6.3 Model implementation and parameter extraction in IC-CAP

Based on the scaling equations derived in Section 6.2 and the geometry scaling results in Chapter 5, the configuration scalable model is utilized as an extension of the geometry scalable Mextram model implemented in Verilog-A as follows.

In the implemented model, new instance parameters (`NE`, `NC`) and geometry parameters (`RTHC`, `RCXC`) are added in the module "parameters.inc" to account for the layout configuration and geometry scaling of MEDs. The geometry scalable Mextram model parameters (`RTH`, `RCC`) in "parameters.inc" are still used as the reference parameters. The scaling rules for the scalable Mextram

Table 6.1: Scaling rules for multi-emitter devices.

Scalable Parameters	Geometry Dependent	Finger Dependent
IS, IK, IBF, IBR	W_E, L_E	N_E
CJE, CJC, IHC	W_E, L_E	N_E
ISS, IKS	W_B, L_B	N_E
CJS	W_C, L_C	none
BF, VEF, VER	W_E, L_E	none
PE, PC, TAUB	W_E, L_E	none
BRI	W_E, L_E	none
PS	W_C, L_C	none
RE, RCV, SCRCV	W_E, L_E	$1/N_E$
RBV, RBC	W_E, L_E	$1/N_E$
RTH	W_B, L_B	(6.6)
RCC	W_E, L_E	(6.15)

model parameters listed in Table 6.1 are implemented in "geo_scaling.inc" which is listed in Appendix E. Therefore, the configuration scaling of the bipolar transistor is taken into account in the model simulation.

An extension of the unified parameter extraction procedure for the scalable Mextram model is implemented in an Agilent IC-CAP model file. The high speed SiGe HBTs with different layout configurations as shown in Fig. 6.5 were used as a test example. The geometry parameters RTHC and RCXC are extracted from measured output characteristics of MEDs following the geometry parameter extraction.

Without re-extracting any parameter from SEDs after RTHC and RCXC extraction, Fig. 6.6 through Fig. 6.9 show the fitting plots of forward Gummel, f_T v.s. I_C and S-parameters of devices having different A_E v.s. N_E combinations. The model simulation shows a good fit to the measured data. Especially in Fig. 6.6, it shows the model simulated (in solid lines) f_T v.s. I_C derived from ADS comparing with the measured data (in circles) from N_E=1∼4, N_C=2 and different A_E devices. The model simulation results can accurately predict the degradation of the measured peak f_T with increasing emitter number. The default Mextram model parameter "MULT" scaling (in dashed line), which puts MULT SEDs in parallel, however, is not able to provide that results. Since the f_T can be simply expressed as the sum of the total forward transit time (τ_F), BE and BC junction capacitances charging time ($\frac{kT}{qI_C}(C_{BE}+C_{BC})$)

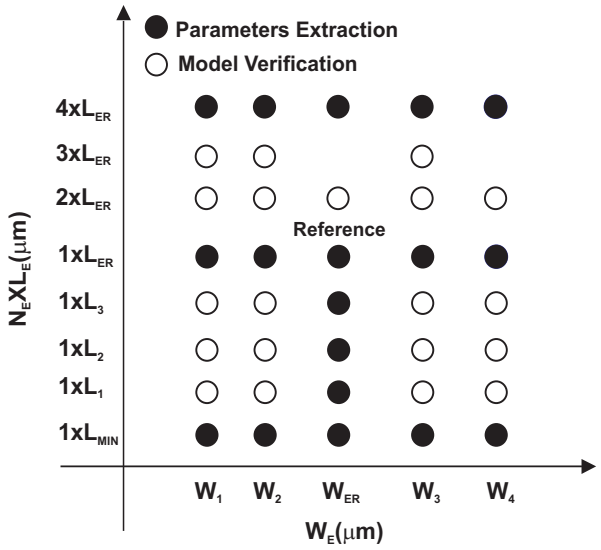


Figure 6.5: Devices' geometry matrix including multi-emitter (N_E=2, 3 and 4) devices with two collector contact (N_C=2) for geometry parameters extraction.

and collector R-C delay time ($\mathrm{RCC} \cdot C_{BC}$) [18] as

$$\frac{1}{2\pi f_T} = \tau_F + \frac{kT}{qI_C}(C_{BE} + C_{BC}) + \mathrm{RCC} \cdot C_{BC}, \tag{6.16}$$

while two capacitances charging time staying almost constant with increasing emitter fingers, the f_T-degradation arises from the increase of total forward transit time and the collector R-C delay time. Therefore, the accurate prediction of f_T-degradation with N_E from the proposed configurational scalable model validates the thermal resistance and collector resistance scaling results.

6.4 Conclusions

In this chapter, the reference geometry based scaling approach for the Mextram model has been extended to different layout configurations of bipolar devices. The proposed scaling methodology employs effective thermal and collector node assumption of multi-emitter devices.

The configuration scalable model uses almost the same set of model parameters as the geometry scalable model except `RTHC` and `RCXC`, which model the mutual heating effect and collector resistance between two adjacent emitter

fingers and two additional instance parameters `NE` and `NC` for different layout configurations. The model nicely predicts the f_T-degradation due to an increase of the mutual heating and the collector delay time with an increasing number of emitters. It is concluded that the number of emitters can't be increased without a limit as can't increase emitter area without a limit.

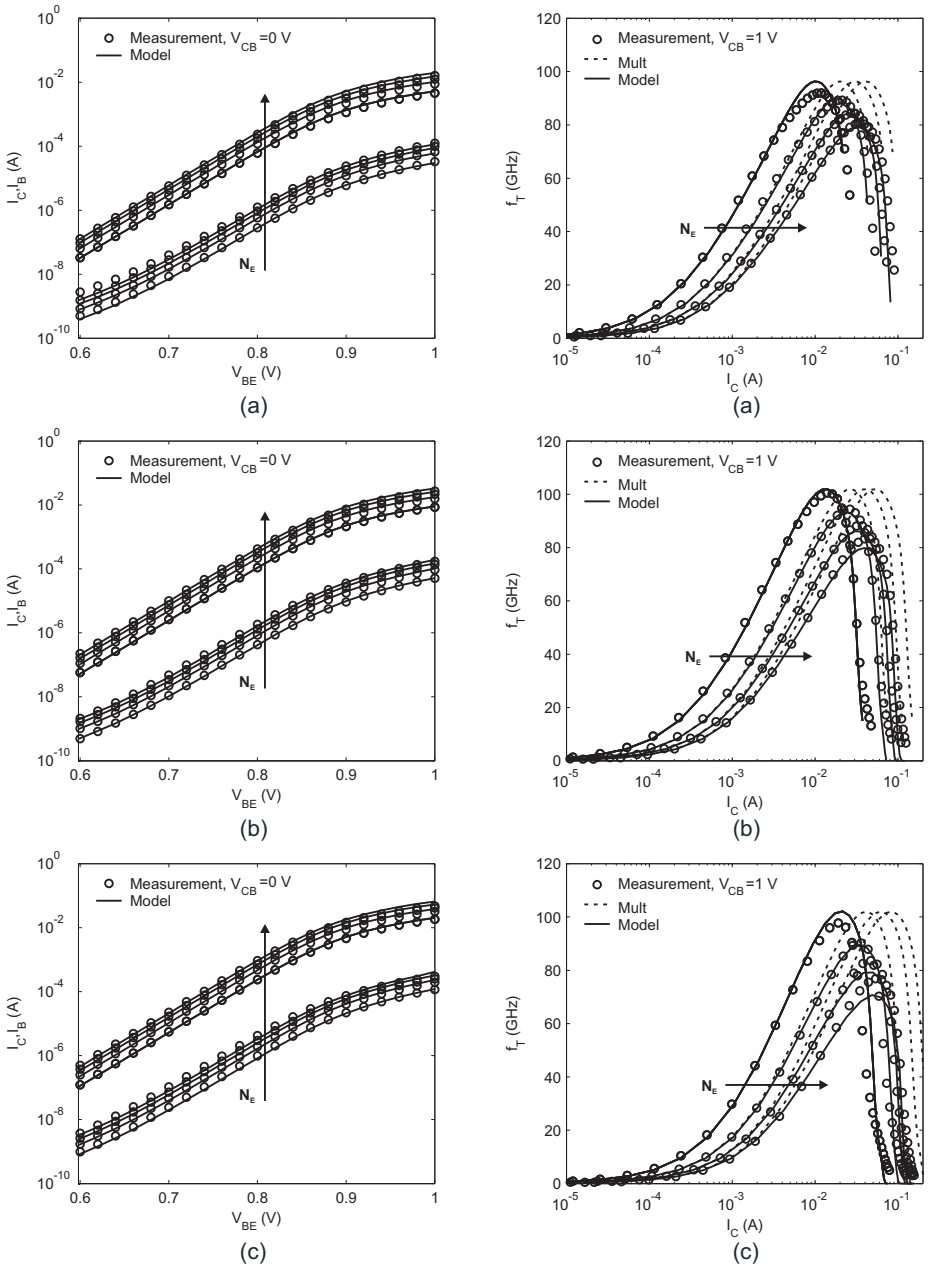


Figure 6.6: model simulated and measured forward Gummel and f_T v.s. I_C from (a) $A_E = 0.2 \times 10.16$, (b) $A_E = 0.3 \times 10.16$ and (c) $A_E = 0.6 \times 10.16$ μm^2 with $N_E = 1 \sim 4$, $N_C = 2$ at $T = T_R = 25$ °C.

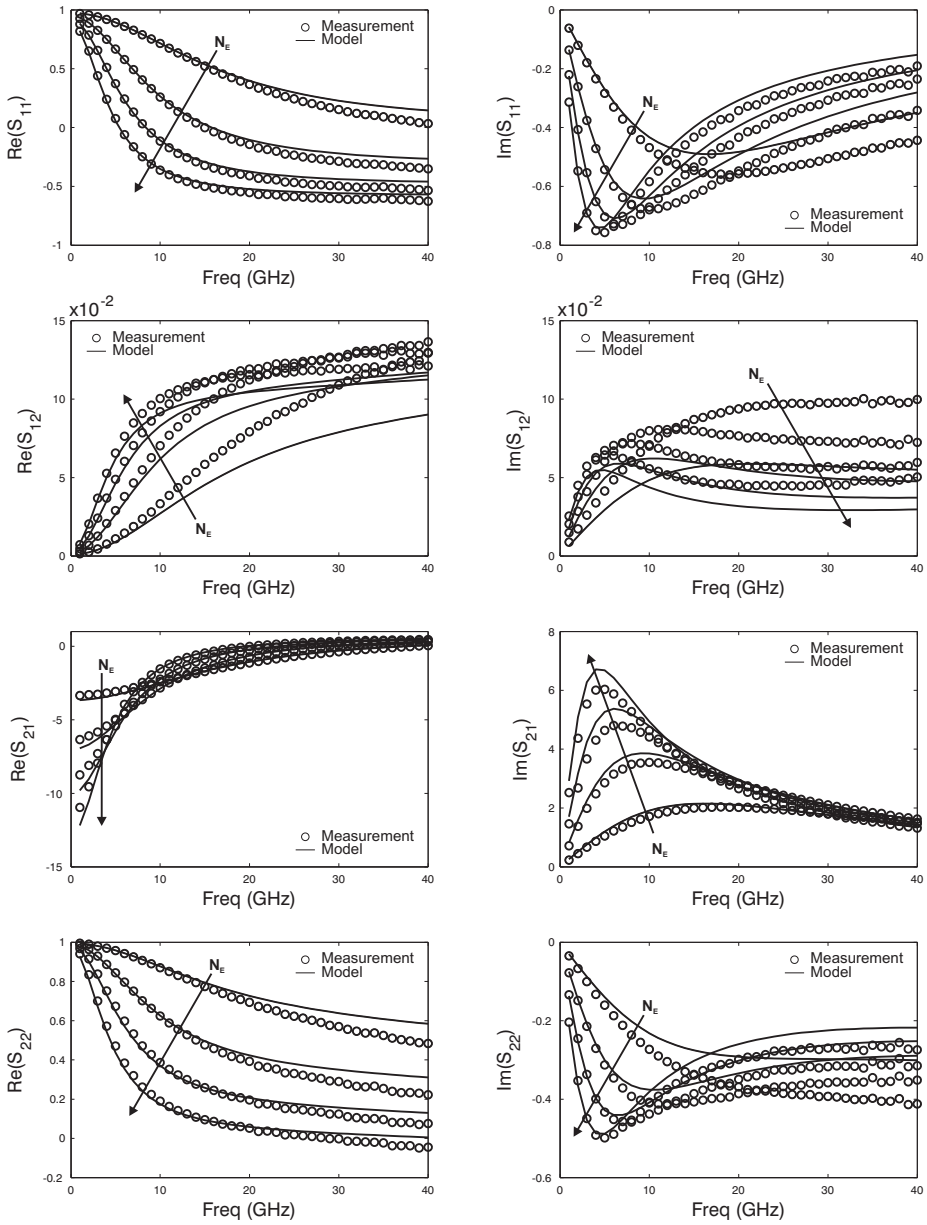


Figure 6.7: Measured and simulated S-parameters v.s. frequency (1∼40 GHz) of $A_E = 0.2{\times}10.16\ \mu m^2$, $N_C = 2$ and $N_E = 1 \sim 4$ devices biased at $V_{BE} = 0.95$ V and $V_{CE} = 2$ V and at $T = T_R = 25$ °C.

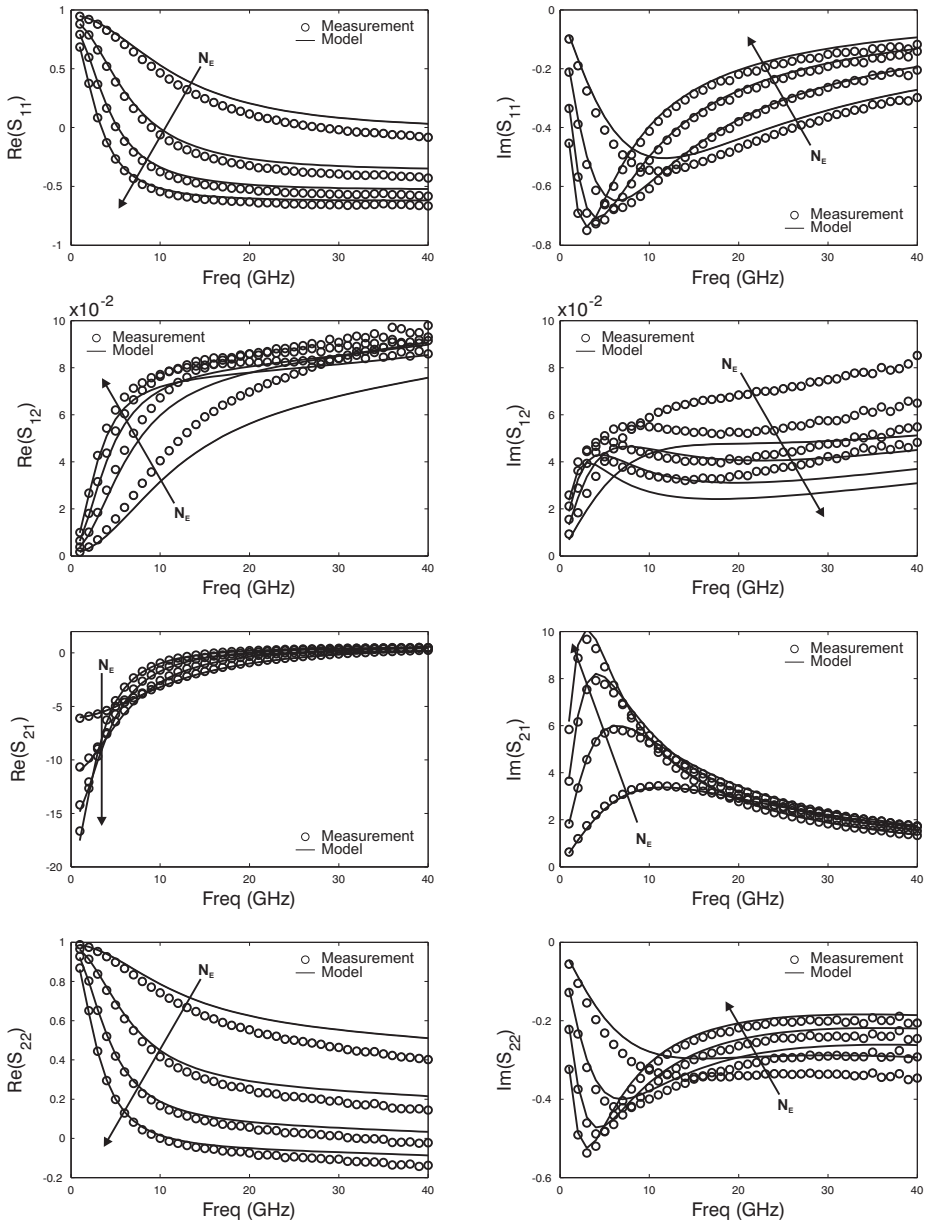


Figure 6.8: Measured and simulated S-parameters v.s. frequency (1∼40 GHz) of $A_E = 0.3{\times}10.16\ \mu m^2$, $N_C = 2$ and $N_E = 1 \sim 4$ devices biased at $V_{BE} = 0.95$ V and $V_{CE} = 2$ V and at $T = T_R = 25$ °C.

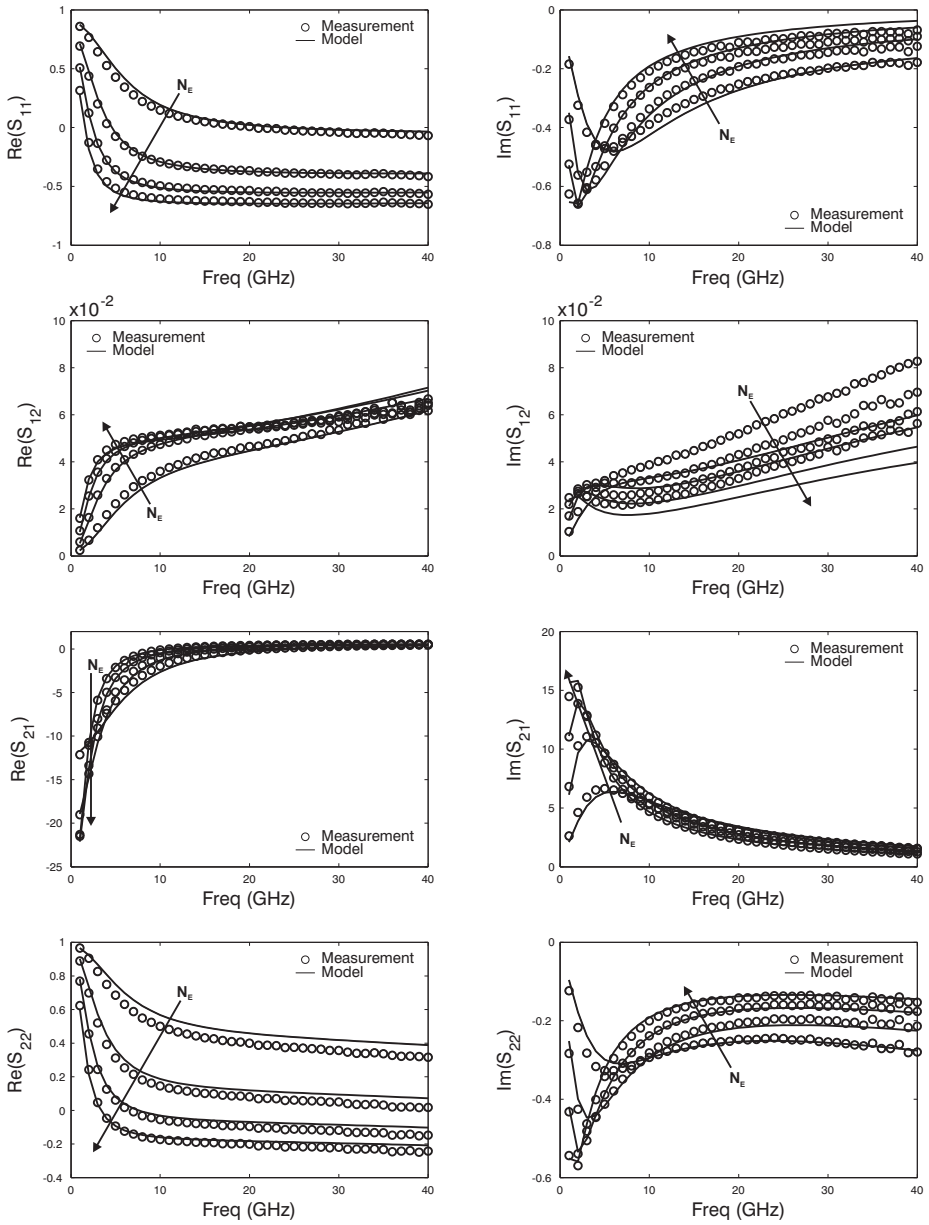


Figure 6.9: Measured and simulated S-parameters v.s. frequency (1∼40 GHz) of $A_E = 0.6{\times}10.16\ \mu m^2$, $N_C = 2$ and $N_E = 1 \sim 4$ devices biased at $V_{BE} = 0.95$ V and $V_{CE} = 2$ V and at $T = T_R = 25$ °C.

Chapter 7

Conclusions and recommendations

7.1 Conclusions

In the modern semiconductor industry, compact models play an important role in connecting semiconductor manufacturing and circuit design. Moore's law tells us that the number of transistors on a chip doubles about every two years. This means that there is a new process generation about every two years. In each process generation, there are also different kinds of devices (high-voltage, general-purpose, high-speed etc.) to meet the different design specifications. With sky-high costs for process development and masks, the semiconductor devices' model library for each process generation has to be created accurately and efficiently to reduce the circuit design cycle time. As the complexity of the integrated circuits increases, there are also more and more requirements on the functionality of compact models. One of the requirements is the scalability of the model. Not only temperature scalability but also geometry scalability is required.

In this thesis, special emphasis has been put on geometry and configuration scaling since there are deficiencies in the existing bipolar device model's scaling approach such as:

1. The scalable bipolar models are mainly based on the physical or electrical size of the transistor's junctions. The determination of the physical device sizes typically requires scanning electron microscope (SEM) inspection of special device test samples during the process steps. The same device might not be used later for the model evaluation. On the other hand, it is difficult to uniquely estimate the device electrical sizes from the measured parame-

ter v.s. drawn (layout) device dimensions. Furthermore, the electrical device size could be quite different for the model parameters representing different physical quantities, like currents or charges.

2. The scalable model implementation is not standardized. It usually employs the sub-circuit netlist approach, which includes the geometry parameters and scaling rules in the model library. However, it takes effort to have the netlist adopted by different circuit simulators and it is difficult to have IP protection of the scaling rules (if such a protection is required). Some designers use a single device model library generated from a toolbox, which is controlled by the model developer. It offers IP protection but designers sometimes still have to come back to the model developer for a specific size, which they need for their design but is not available in the model library.

3. The geometry scaling is typically not part of the bipolar model descriptions. It is performed apart from the single model parameter extraction. Therefore various auxiliary tools have been developed to extract geometry parameters based on certain geometry scaling rules, sheet resistance measurement, process data and sets of single device parameters at different geometries. However, such a heterogeneous parameter extraction procedure could be in some cases inefficient due to the switching between different tools. Moreover, the resulting scalable model might accumulate errors originating from individual parameter extraction steps in different tools.

The above mentioned deficiencies have been overcome in the new scalable bipolar transistor model Mextram. Its geometry scaling is mainly based on the physical properties of Mextram parameters but scales with the drawn size in the device layout. Therefore, there is no need for physical or electrical device size measurement. The scalable electrical parameter and the geometry in a scaling rule are normalized to the reference parameter and the reference geometry analogous to the formulation for the temperature scaling. So, the geometry scaling rules can be incorporated in the standard Mextram model implemented in AHDL Verilog-A language, which is an IEEE standard language for describing analog systems and it is adopted in many commercial simulators. The Verilog-A model can also offer IP protection if it is compiled as a shared library. In the mean time, designers still have the full freedom in choosing device geometry when the shared library is linked to the circuit simulator.

Along with the scalable model, a unified parameter extraction procedure for reference, temperature and geometry parameters of the scalable model has been implemented in an IC-CAP model file. As a result, no additional tool is

required for temperature and geometry parameter extraction of the scalable model. With a systematic parameter extraction procedure for single Mextram model parameters at a reference geometry, temperature parameters from measured data at various temperatures and geometry parameters from measured data at various selected geometries, the complete parameter extraction procedure has been demonstrated on a high-speed SiGe HBT technology. The resulting model shows good fit to CV, DC, f_T and S-parameters measured from the high-speed SiGe HBTs. As a result, the scalable model library generation is efficient and the accuracy is increased with the new scaling approach.

The scalable model is extended to different layout configurations without adding extra thermal and collector nodes based on effective thermal and collector node assumptions. Therefore, the configuration scalable model uses almost the same set of model parameters as the geometry scalable model except additional geometry parameters for the mutual heating effect and collector resistance between two adjacent fingers and additional instance parameters for different layout configurations. The model nicely predicts the f_T-degradation due to an increase of the mutual heating effect and the collector delay time with an increasing number of emitter fingers. It is concluded that the number of emitter fingers can not be increased without a limit. Moreover, since the geometry scaling approach is based on the scaling of the physical properties, it can be applied to the Gummel-Poon, HICUM and VBIC models as well.

7.2 Recommendations for the future work

Though Mextram model and its scalable variant are good enough for most applications today, there is still room to improve it for advanced circuit design. Listed below are suggestions for the future work that can be done to improve Mextram model and its parameter extraction procedure.

Predicted model from the TCAD simulation

Technology computer aided design (TCAD) tools including process [131–133] and device [134–136] simulators are usually used in assisting process and compact model development. With accurate physical models and a proper calibration procedure, TCAD simulation results can be very close to the measured data of SiGe HBTs [137]. Though mixed device/circuit simulation is available in device simulators [134–136], it is usually time consuming compared to circuit simulation with circuit simulator. So, a feasible solution may be to use

device simulation results to synthesize measured CV, DC, S-parameters and noise data. Then, parameter extraction from the synthesized measured data can be used as a predicted model for the circuit simulation. Therefore, circuit design can proceed during process development. Moreover, test structures in the test chip used for Mextram modeling purposes can be reduced in space and number if TCAD simulation results are widely used for parameter extraction.

Incorporation of process variation in the scalable Mextram model

When the bipolar device's emitter size enters the deep sub-micron era, small fluctuations during the device process steps may become significant in the device electrical characteristics. Since the process simulator can simulate the process steps during semiconductor device fabrication, it can be used to analyze relation between process variations and model parameters. Besides, the correlation between different parameters from the same process variation can also be found. By incorporating statistical variations of process parameters in the scalable Mextram model, the worst-case characteristic and the mismatch of bipolar devices can be simulated during accurate mixed-signal circuit design. Therefore, some precautions can be taken to prevent failures of mixed-signal circuits and hence to increase their yields.

Parallel models for multi-emitter bipolar devices

As discussed in Chapter 6, there are non-uniform potential and temperature distributions in the multi-emitter bipolar device due to the lateral current flow and the mutual heating effect. It is modeled with an average temperature and potential rise based on the assumption that the current flowing through each emitter finger is uniform. However, a non-uniform temperature and potential will vary the current flowing through each emitter. Therefore, a macro model, which puts Mextram models in parallel to represent the number of emitters, may give a better description of the non-uniform temperature and potential distribution for the high-power bipolar circuit design. Actually, a quasi-distributed three-dimension model, which puts bipolar models in parallel to represent part of the large emitter area, has been used to find a possible pinched point for the avalanche effect while the device is operated beyond BV_{CEO} [138].

Simplified version of Mextram model

Though there are requests from some Mextram model users to include more physical effects of bipolar devices in the Mextram model, there are also requests from others in the opposite - to simplify it. For example, the high current injection effect of parasitic pnp only exists in the reverse-active region and rarely occurs in the normal operation. The non-ohmic quasi-saturation effect only happens when there is high collector current and voltage. It does not always exist in the bipolar device with an optimized collector design [8]. Some of the exponential functions in the temperature scaling rules can be linearized as in SPICE Gummel-Poon model. From the facts that we have discussed above, it is possible to have a simplified version of Mextram to reduce simulation time and model complexity.

Appendix A

List of Mextram model parameters

level 504.6

The parameters denoted with a '$*$' are not used in the DC model

#	symbol	name	units	description
1	LEVEL	`LEVEL`	—	Model level, must be set to 504
2	T_{ref}	`TREF`	°C	Reference temperature. Default is 25°C
3	DTA	`DTA`	°C	Difference between the local ambient and global ambient temperatures: $T_{\text{local ambient}} = T_{\text{global ambient}}$ + DTA
4	EXMOD	`EXMOD`	—	Flag for extended modelling of the reverse current gain
$*5$	EXPHI	`EXPHI`	—	Flag for the distributed high-frequency effects in transient
6	EXAVL	`EXAVL`	—	Flag for extended modelling of avalanche currents
7	I_s	`IS`	A	Collector-emitter saturation current
8	I_k	`IK`	A	Collector-emitter high injection knee current
9	V_{er}	`VER`	V	Reverse Early voltage
10	V_{ef}	`VEF`	V	Forward Early voltage
11	β_f	`BF`	—	Ideal forward current gain
12	I_{Bf}	`IBF`	A	Saturation current of the non-ideal forward base current
13	m_{Lf}	`MLF`	—	Non-ideality factor of the non-ideal forward base current

#	symbol	name	units	description
14	XI_{B_1}	XIBI	—	Part of ideal base current that belongs to the sidewall
15	β_{ri}	BRI	—	Ideal reverse current gain
16	I_{Br}	IBR	A	Saturation current of the non-ideal reverse base current
17	V_{Lr}	VLR	V	Cross-over voltage of the non-ideal reverse base current
18	X_{ext}	XEXT	—	Part of I_{ex}, Q_{tex}, Q_{ex} and I_{sub} that depends on V_{BC_1} instead of $V_{B_1C_1}$
19	W_{avl}	WAVL	m	Epilayer thickness used in weak-avalanche model
20	V_{avl}	VAVL	V	Voltage determining curvature of avalanche current
21	S_{fH}	SFH	—	Current spreading factor of avalanche model (when EXAVL=1)
22	R_E	RE	Ω	Emitter resistance
23	R_{Bc}	RBC	Ω	Constant part of the base resistance
24	R_{Bv}	RBV	Ω	Zero-bias value of the variable part of the base resistance
25	R_{Cc}	RCC	Ω	Constant part of the collector resistance
26	R_{Cv}	RCV	Ω	Resistance of the un-modulated epilayer
27	SCR_{Cv}	SCRCV	Ω	Space charge resistance of the epilayer
28	I_{hc}	IHC	A	Critical current for velocity saturation in the epilayer
29	a_{x_i}	AXI	—	Smoothness parameter for the onset of quasi-saturation
∗30	C_{j_E}	CJE	F	Zero-bias emitter-base depletion capacitance
31	V_{d_E}	VDE	V	Emitter-base diffusion voltage
32	p_E	PE	—	Emitter-base grading coefficient
∗33	XC_{j_E}	XCJE	—	Fraction of the emitter-base depletion capacitance that belongs to the sidewall
∗34	C_{BEO}	CBEO	—	Emitter-base overlap capacitance
∗35	C_{j_C}	CJC	F	Zero-bias collector-base depletion capacitance

#	symbol	name	units	description
36	V_{dC}	VDC	V	Collector-base diffusion voltage
37	p_C	PC	—	Collector-base grading coefficient
38	X_p	XP	—	Constant part of C_{jC}
39	m_C	MC	—	Coefficient for the current modulation of the collector-base depletion capacitance
∗40	XC_{jC}	XCJC	—	Fraction of the collector-base depletion capacitance under the emitter
∗41	C_{BCO}	CBCO	—	Collector-base overlap capacitance
∗42	m_τ	MTAU	—	Non-ideality factor of the emitter stored charge
∗43	τ_E	TAUE	s	Minimum transit time of stored emitter charge
∗44	τ_B	TAUB	s	Transit time of stored base charge
∗45	τ_{epi}	TEPI	s	Transit time of stored epilayer charge
∗46	τ_R	TAUR	s	Transit time of reverse extrinsic stored base charge
47	dE_g	DEG	eV	Bandgap difference over the base
48	X_{rec}	XREC	—	Pre-factor of the recombination part of I_{B_1}
49	$A_{Q_{B0}}$	AQBO	—	Temperature coefficient of the zero-bias base charge
50	A_E	AE	—	Temperature coefficient of the resistivity of the emitter
51	A_B	AB	—	Temperature coefficient of the resistivity of the base
52	A_{epi}	AEPI	—	Temperature coefficient of the resistivity of the epilayer
53	A_{ex}	AEX	—	Temperature coefficient of the resistivity of the extrinsic base
54	A_C	AC	—	Temperature coefficient of the resistivity of the buried layer
55	dA_{I_s}	DAIS	—	Parameter for fine tuning of temperature dependence of collector-emitter saturation current

#	symbol	name	units	description
56	$dV_{g\beta f}$	DVGBF	V	Band-gap voltage difference of forward current gain
57	$dV_{g\beta r}$	DVGBR	V	Band-gap voltage difference of reverse current gain
58	V_{g_B}	VGB	V	Band-gap voltage of the base
59	V_{g_C}	VGC	V	Band-gap voltage of the collector
60	V_{g_j}	VGJ	V	Band-gap voltage recombination emitter-base junction
∗61	$dV_{g\tau_E}$	DVGTE	V	Band-gap voltage difference of emitter stored charge
∗62	A_f	AF	—	Exponent of the Flicker-noise
∗63	K_f	KF	—	Flicker-noise coefficient of the ideal base current
∗64	K_{fN}	KFN	—	Flicker-noise coefficient of the non-ideal base current
∗65	K_{avl}	KAVL	—	Switch for white noise contribution due to avalanche
66	I_{Ss}	ISS	A	Base-substrate saturation current
67	I_{ks}	IKS	A	Base-substrate high injection knee current
∗68	C_{jS}	CJS	F	Zero-bias collector-substrate depletion capacitance
∗69	V_{dS}	VDS	V	Collector-substrate diffusion voltage
∗70	p_S	PS	—	Collector-substrate grading coefficient
71	V_{gS}	VGS	V	Band-gap voltage of the substrate
72	A_S	AS	—	For a closed buried layer: $A_S = A_C$, and for an open buried layer: $A_S = A_{epi}$
73	R_{th}	RTH	°C/W	Thermal resistance
∗74	C_{th}	CTH	J/°C	Thermal capacitance
75	A_{th}	ATH	—	Temperature coefficient of the thermal resistance
76	MULT	MULT	—	Multiplication factor

Appendix B

Tunneling current model implementation in Verilog-A

```
//Temperature scaling for tunneling current
BTJE_T = BTJE * pow(VDE / VDE_T, PE) * (1+ABTJE*(tN-1));
VTJE_T = VTJE * pow(VDE_T / VDE, PE) * (1+AVTJE*(tN-1));
BTJE_TM = BTJE_T * MULT;

// Base-emitter tunneling current
Ibet = 0.0;
if (Vb2e1 < 0.0) begin
  Vqte        =   VDE_T / (1 - PE) - Vte;
  eVb2e1_vt2  =   pow(eVb2e1, 0.5);
  Vb2e1_bbt   =   Vb2e1 * (1.0 - eVb2e1_vt2)/ (1.0 + eVb2e1_vt2);
  Ibet        =   BTJE_TM * Vb2e1_bbt * Vqte * exp(-VTJE_T / Vqte);
end
I(b2, e1) < + TYPE * Ibet;
```

Note: The complete Mextram Verilog-A code is available in
http://hitec.ewi.tudelft.nl/mug/

Appendix C

Definition of 2-port Y, Z, H, ABCD-parameter and their transformation

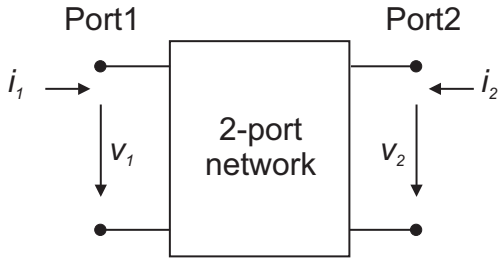


Fig. C.1: Terminal voltages and currents in a 2-port network .

$$\begin{bmatrix} i_1 \\ i_2 \end{bmatrix} = \begin{bmatrix} y_{11} y_{12} \\ y_{21} y_{22} \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} \qquad \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \begin{bmatrix} z_{11} z_{12} \\ z_{21} z_{22} \end{bmatrix} \begin{bmatrix} i_1 \\ i_2 \end{bmatrix}$$

$$\begin{bmatrix} v_1 \\ i_2 \end{bmatrix} = \begin{bmatrix} h_{11} h_{12} \\ h_{21} h_{22} \end{bmatrix} \begin{bmatrix} i_1 \\ v_2 \end{bmatrix} \qquad \begin{bmatrix} v_1 \\ i_1 \end{bmatrix} = \begin{bmatrix} AB \\ CD \end{bmatrix} \begin{bmatrix} v_2 \\ -i_2 \end{bmatrix}$$

Table C.1: S, Z, Y, H and ABCD - Parameters Conversion Table

	S	**Z**	**Y**	**H**	**ABCD**
S	$S_{11}S_{12}$ $S_{21}S_{22}$	$S_{11}=\frac{(Z'_{11}-1)(Z'_{22}+1)-Z'_{12}Z'_{21}}{(Z'_{11}+1)(Z'_{22}+1)-Z'_{12}Z'_{21}}$ $S_{12}=\frac{2Z'_{12}}{(Z'_{11}+1)(Z'_{22}+1)-Z'_{12}Z'_{21}}$ $S_{21}=\frac{2Z'_{21}}{(Z'_{11}+1)(Z'_{22}+1)-Z'_{12}Z'_{21}}$ $S_{22}=\frac{(Z'_{11}+1)(Z'_{22}-1)-Z'_{12}Z'_{21}}{(Z'_{11}+1)(Z'_{22}+1)-Z'_{12}Z'_{21}}$	$S_{11}=\frac{(1-Y'_{11})(1+Y'_{22})+Y'_{12}Y'_{21}}{(1+Y'_{11})(1+Y'_{22})-Y'_{12}Y'_{21}}$ $S_{12}=\frac{-2Y'_{12}}{(1+Y'_{11})(1+Y'_{22})-Y'_{12}Y'_{21}}$ $S_{21}=\frac{-2Y'_{21}}{(1+Y'_{11})(1+Y'_{22})-Y'_{12}Y'_{21}}$ $S_{22}=\frac{(1+Y'_{11})(1-Y'_{22})+Y'_{12}Y'_{21}}{(1+Y'_{11})(1+Y'_{22})-Y'_{12}Y'_{21}}$	$S_{11}=\frac{(H'_{11}-1)(H'_{22}+1)-H'_{12}H'_{21}}{(H'_{11}+1)(H'_{22}+1)-H'_{12}H'_{21}}$ $S_{12}=\frac{2H'_{12}}{(H'_{11}+1)(H'_{22}+1)-H'_{12}H'_{21}}$ $S_{21}=\frac{-2H'_{21}}{(H'_{11}+1)(H'_{22}+1)-H'_{12}H'_{21}}$ $S_{22}=\frac{(1+H'_{11})(1-H'_{22})+H'_{12}H'_{21}}{(H'_{11}+1)(H'_{22}+1)-H'_{12}H'_{21}}$	$S_{11}=\frac{A'+B'-C'-D'}{A'+B'+C'+D'}$ $S_{12}=\frac{2(A'D'-B'C')}{A'+B'+C'+D'}$ $S_{21}=\frac{2}{A'+B'+C'+D'}$ $S_{22}=\frac{-A'+B'-C'+D'}{A'+B'+C'+D'}$
Z	$Z'_{11}=\frac{(1+S_{11})(1-S_{22})+S_{12}S_{21}}{(1-S_{11})(1-S_{22})-S_{12}S_{21}}$ $Z'_{12}=\frac{2S_{12}}{(1-S_{11})(1-S_{22})-S_{12}S_{21}}$ $Z'_{21}=\frac{2S_{21}}{(1-S_{11})(1-S_{22})-S_{12}S_{21}}$ $Z'_{22}=\frac{(1-S_{11})(1+S_{22})+S_{12}S_{21}}{(1-S_{11})(1-S_{22})-S_{12}S_{21}}$	$Z_{11}Z_{12}$ $Z_{21}Z_{22}$	$\frac{Y_{22}}{\lvert Y\rvert}\ \frac{-Y_{12}}{\lvert Y\rvert}$ $\frac{-Y_{21}}{\lvert Y\rvert}\ \frac{Y_{11}}{\lvert Y\rvert}$	$\frac{\lvert H\rvert}{H_{22}}\ \frac{H_{12}}{H_{22}}$ $\frac{-H_{21}}{H_{22}}\ \frac{1}{H_{22}}$	$\frac{A}{C}\ \frac{AD-BC}{C}$ $\frac{1}{C}\ \frac{D}{C}$
Y	$Y'_{11}=\frac{(1-S_{11})(1+S_{22})+S_{12}S_{21}}{(1+S_{11})(1+S_{22})-S_{12}S_{21}}$ $Y'_{12}=\frac{-2S_{12}}{(1+S_{11})(1+S_{22})-S_{12}S_{21}}$ $Y'_{21}=\frac{-2S_{21}}{(1+S_{11})(1+S_{22})-S_{12}S_{21}}$ $Y'_{22}=\frac{(1+S_{11})(1-S_{22})+S_{12}S_{21}}{(1+S_{11})(1+S_{22})-S_{12}S_{21}}$	$\frac{Z_{22}}{\lvert Z\rvert}\ \frac{-Z_{12}}{\lvert Z\rvert}$ $\frac{-Z_{21}}{\lvert Z\rvert}\ \frac{Z_{11}}{\lvert Z\rvert}$	$Y_{11}Y_{12}$ $Y_{21}Y_{22}$	$\frac{1}{H_{11}}\ \frac{-H_{12}}{H_{11}}$ $\frac{H_{21}}{H_{11}}\ \frac{\lvert H\rvert}{H_{11}}$	$\frac{D}{B}\ \frac{BC-AD}{B}$ $\frac{-1}{B}\ \frac{A}{B}$
H	$H'_{11}=\frac{(1+S_{11})(1+S_{22})-S_{12}S_{21}}{(1-S_{11})(1+S_{22})+S_{12}S_{21}}$ $H'_{12}=\frac{2S_{12}}{(1-S_{11})(1+S_{22})+S_{12}S_{21}}$ $H'_{21}=\frac{-2S_{21}}{(1-S_{11})(1+S_{22})+S_{12}S_{21}}$ $H'_{22}=\frac{(1-S_{11})(1-S_{22})-S_{12}S_{21}}{(1-S_{11})(1+S_{22})+S_{12}S_{21}}$	$\frac{\lvert Z\rvert}{Z_{22}}\ \frac{Z_{12}}{Z_{22}}$ $\frac{-Z_{21}}{Z_{22}}\ \frac{1}{Z_{22}}$	$\frac{1}{Y_{11}}\ \frac{-Y_{12}}{Y_{11}}$ $\frac{Y_{21}}{Y_{11}}\ \frac{\lvert Y\rvert}{Y_{11}}$	$H_{11}H_{12}$ $H_{21}H_{22}$	$\frac{B}{D}\ \frac{AD-BC}{D}$ $\frac{-1}{B}\ \frac{A}{B}$
ABCD	$A'=\frac{(1+S_{11})(1-S_{22})+S_{12}S_{21}}{2S_{21}}$ $B'=\frac{(1+S_{11})(1+S_{22})-S_{12}S_{21}}{2S_{21}}$ $C'=\frac{(1-S_{11})(1-S_{22})-S_{12}S_{21}}{2S_{21}}$ $D'=\frac{(1-S_{11})(1+S_{22})+S_{12}S_{21}}{2S_{21}}$	$\frac{Z_{11}}{Z_{21}}\ \frac{\lvert Z\rvert}{Z_{21}}$ $\frac{1}{Z_{21}}\ \frac{Z_{22}}{Z_{21}}$	$\frac{-Y_{22}}{Y_{21}}\ \frac{-1}{Y_{21}}$ $\frac{-\lvert Y\rvert}{Y_{21}}\ \frac{-Y_{11}}{Y_{21}}$	$\frac{-\lvert H\rvert}{H_{21}}\ \frac{-H_{11}}{H_{21}}$ $\frac{-H_{22}}{H_{21}}\ \frac{-1}{H_{21}}$	AB CD

Note: 2-port parameters with ' are normalized to the characteristic impedance Z_0 (ex: Z'_{11}=Z_{11}/Z_0, Y'_{11}=$Y_{11}\cdot Z_0$ etc.)

Appendix D

Behavior substrate resistance model implementation in Verilog-A

```
module rsub_va (s, g);
inout s, g;
electrical s, g;
// Numerator Coefficients
parameter real N0 = 3E3;
parameter real N1 = -4.339E9;
parameter real N2 = -2.818E9;
parameter real N3 = -116.6E6;
parameter real N4 = -1.016E6;
parameter real N5 = 33.98E3;

// Denominator Coefficients
parameter real D0 = 1;
parameter real D1 = -1.826E3;
parameter real D2 = 466.9E3;
parameter real D3 = 506.9E3;
parameter real D4 = 63.69E3;
parameter real D5 = -70.14;
parameter real D6 = 1;

// Scaling factor parameter
parameter real SCALE_PAR = 1E9;

// Numerator and denominator Vectors
```

```
real num[0:5];
real den[0:6];

// Scaling factor and 2*pi
real pii, pii2;
real scale_factor;

analog begin
    pii = 3.1415926;
    pii2 = 2 * pii;
    scale_factor = 1 / (pii2 * SCALE_PAR);
    num[0] = N0;
    num[1] = N1 * scale_factor;
    num[2] = N2 * pow(scale_factor, 2);
    num[3] = N3 * pow(scale_factor, 3);
    num[4] = N4 * pow(scale_factor, 4);
    num[5] = N5 * pow(scale_factor, 5);
    den[0]  = D0;
    den[1]  = D1 * scale_factor;
    den[2]  = D2 * pow(scale_factor, 2);
    den[3]  = D3 * pow(scale_factor, 3);
    den[4]  = D4 * pow(scale_factor, 4);
    den[5]  = D5 * pow(scale_factor, 5);
    den[6]  = D6 * pow(scale_factor, 6);
    V(s,g) < + laplace_nd(I(s,g), num, den);
end
endmodule
```

Appendix E

Geometry scaling rules implementation in Verilog-A

```
//Geometry scaling rules for a scalable Mextram model
 ne          =  NE;
 nb          =  NE+1.0;
 nc          =  NC;
 we          =  W;
 le          =  L;
 weref       =  WREF;
 leref       =  LREF;
 wb          =  we+detob;
 lb          =  le+detob;
 wbref       =  weref+detob;
 lbref       =  leref+detob;
 wc          =  we+detoc;
 lc          =  le+detoc;
 wcref       =  weref+detoc;
 lcref       =  leref+detoc;
 IS_scale    =  ne*IS*(1.0+ISA*(we*le/(weref*leref)-1.0)
                +ISW*(we/weref-1.0)+ISL*(le/leref-1.0));
 IK_scale    =  ne*IK*(1.0+IKA*(we*le/(weref*leref)-1.0)
                +IKW*(we/weref-1.0)+IKL*(le/leref-1.0));
```

```
RE_scale    =  RE/(ne*(1.0+REA*(we*le/(weref*leref)-1.0)
               +REW*(we/weref-1.0)+REL*(le/leref-1.0)))+0.7;
RBV_scale   =  (1.0/ne)*RBV*we/weref*leref/le*(1.0/12.0-(1.0/12.0
               -1.0/28.45)*we/le)/(1.0/12.0-(1.0/12.0-1.0/28.45)
               *weref/leref);
RBC_scale   =  RBC/(ne*(1.0+RBCL*(le/leref-1.0)
               +RBCW*(we/weref-1.0)));
rc1         =  (1.0-RCX)*(RCI/ne*weref*leref/we/le+(1.0-RCI)
               *we/le*leref/weref*ne/nc/nc);
rc2         =  RCXC*leref/le*(ne-1.0)*(2.0*ne/nc/nc-1.0)/ne;
rc3         =  RCX/nc/(1.0+RCW*(ne/nc*we/weref-1.0)
               +RCL*(le/leref-1.0));
RCC_scale   =  RCC*(rc1+rc2+rc3);
CJE_scale   =  ne*CJE*(1.0+CJEA*(we*le/(weref*leref)-1.0)
               +CJEW*(we/weref-1.0)+CJEL*(le/leref-1.0));
PE_scale    =  PE*(1.0+PEW*(weref/we-1.0)+PEL*(leref/le-1.0)
               +PEC*(weref*leref/(we*le)-1.0));
CJC_scale   =  ne*CJC*(1.0+CJCA*(we*le/(weref*leref)-1.0)
               +CJCW*(we/weref-1.0)+CJCL*(le/leref-1.0));
PC_scale    =  PC*(1.0+PCW*(weref/we-1.0)+PCL*(leref/le-1.0)
               +PCC*(weref*leref/(we*le)-1.0));
CJS_scale   =  CJS*(1.0+CJSA*(wc*lc/(wcref*lcref)-1.0)
               +CJSW*(wc/wcref-1.0)+CJSL*(lc/lcref-1.0));
PS_scale    =  PS*(1.0+PSW*(wcref/wc-1.0)+PSL*(lcref/lc-1.0)
               +PSC*(wcref*lcref/(wc*lc)-1.0));
BF_scale    =  BF*(1.0+BFW*(weref/we-1.0)+BFL*(leref/le-1.0)
               +BFC*(weref*leref/(we*le)-1.0));
VEF_scale   =  VEF*(1.0+VEFW*(weref/we-1.0)+VEFL
               *(leref/le-1.0)+VEFC*(weref*leref/(we*le)-1.0));
VER_scale   =  VER*(1.0+VERW*(weref/we-1.0)+VERL
               *(leref/le-1.0)+VERC*(weref*leref/(we*le)-1.0));
IBF_scale   =  ne*IBF*(1.0+IBFA*(we*le/(weref*leref)-1.0)
               +IBFW*(we/weref-1.0)+IBFL*(le/leref-1.0));
ISS_scale   =  ne*ISS*(1.0+ISSA*(wb*lb/(wbref*lbref)-1.0)
               +ISSW*(wb/wbref-1.0)+ISSL*(lb/lbref-1.0));
```

```
BRI_scale     =  BRI*(1.0+BRIW*(weref/we-1.0)+BRIL*(leref/le-1.0)
                 +BRIC*(weref*leref/(we*le)-1.0));
IBR_scale     =  ne*IBR*(1.0+IBRA*(we*le/(weref*leref)-1.0)
                 +IBRW*(we/weref-1.0)+IBRL*(le/leref-1.0));
IKS_scale     =  ne*IKS*(1.0+IKSA*(wb*lb/(wbref*lbref)-1.0)
                 +IKSW*(wb/wbref-1.0)+IKSL*(lb/lbref-1.0));
TAUB_scale    =  TAUB*(1.0+TAUBW*(weref/we-1.0)+TAUBL
                 *(leref/le-1.0)+TAUBC*(weref*leref/(we*le)-1.0));
RCV_scale     =  RCV/(ne*(1.0+RCVA*(we*le/(weref*leref)-1.0)
                 +RCVW*(we/weref-1.0)+RCVL*(le/leref-1.0)));
IHC_scale     =  ne*IHC*(1.0+IHCA*(we*le/(weref*leref)-1.0)
                 +IHCW*(we/weref-1.0)+IHCL*(le/leref-1.0));
SCRCV_scale =  SCRCV/(ne*(1.0+SCRCVA*(we*le/(weref*leref)-1.0)
                 +SCRCVW*(we/weref-1.0)+SCRCVL*(le/leref-1.0)));
sum1          =  0;
ii            =  1;
while  (ii<   =  ne-1.0) begin
  sum_temp    =  ne/ii-1.0;
  sum1        =  sum1+sum_temp;
  ii          =  ii+1;
end
RTH_scale     =  RTH/ne*(1.0/(1.0+RTHA*(wb*lb/wbref/lbref-1.0)
                 +RTHW*(wb/wbref-1.0)+RTHL*(lb/lbref-1.0))
                 +2.0*RTHC*lbref/lb*wbref/wb*sum1/ne);
```

Bibliography

[1] T. Masuda *et al.*, "45 GHz transimpedance 32 dB limiting amplifier and 40 gb/s 1:4 high-sensitivity demultiplexer with decision circuit using SiGe HBT's for 40 gb/s optical receiver," in *Dig. Tech. Papers, 2000 IEEE International Solid-State Circuits Conference.(ISSC 2000)*, San Francisco, CA, USA, Feb. 2000, pp. 60–61.

[2] D. Szmyd, R. Brock, N. Bell, S. Harker, G. Patrizi, J. Fraser, and R. Dondero, "QUBiC4 a silicon RF-BiCMOS technology for wireless communication ICs," in *Proc. 2001 IEEE Bipolar/BiCMOS Circuits and Technology Meeting (BCTM 2001)*, Minneapolis, MN, USA, Sept. 2001, pp. 60–63.

[3] J. N. Burghartz, M. Irmschera, F. Letzkusa, J. Kretzb, and D. Resnickc, "Lithography for the 32-nm node and beyond," in *Presentation notes of 2006 IEEE Bipolar/BiCMOS Circuits and Technology Meeting (BCTM 2006)*, Maastricht, Netherlands, Oct. 2006, p. 10.3.

[4] R. M. Warner, "Microelectronics: Its unusual origin and personality," *IEEE Trans. Electron Devices*, vol. 48, pp. 2457–2467, Nov. 2001.

[5] J. N. Ellis and V. S. Osadchy, "The 2N3055: a case history," *IEEE Trans. Electron Devices*, vol. 48, pp. 2477–2484, Nov. 2001.

[6] D. D. Tang, P. M. Solomon, T. H. Ning, R. D. Isaac, and R. E. Burger, "1.25 μm deep-groove-isolated self-aligned bipolar circuits," *IEEE J. Solid-State Circuits*, vol. 17, pp. 925–931, Oct. 1982.

[7] G. P. Li, T. H. Ning, C. T. Chuang, M. B. Ketchen, D. D. Tang, and J. Mauer, "An advanced high-performance trench-isolated self-aligned bipolar technology," *IEEE Trans. Electron Devices*, vol. 34, pp. 2246–2254, Nov. 1987.

[8] C. T. Kirk, "A theory of transistor cut-off frequency (f_T) fall-off at high current densities," *IEEE Trans. Electron Devices*, vol. 9, pp. 164–174, Mar. 1962.

[9] D. D. Tang *et al.*, "73 ps bipolar ECL circuits," in *Dig. Tech. Papers, 2000 IEEE International Solid-State Circuits Conference.(ISSC 1986)*, Feb. 1986, pp. 104–105.

[10] S. S. lyer, G. L. Patton, S. L. Delage, S. Tiwari, and J. M. C. Stork, "Silicon-germanium base heterojunction bipolar transistor by molecular beam epitaxy," in *Proc. 1987 IEEE International Electron Devices Meeting (IEDM 1987)*, Dec. 1987, pp. 874 – 876.

[11] B. S. Meyerson, "Low-temperature silicon epitaxy by ultra-high vacuum/chemical vapor deposition," *Appl. Phys. Lett.*, vol. 48, no. 3, pp. 797–799, 1986.

[12] H. K. Gummel, "An charge control relation for bipolar transistors," *Bell Syst Tech. Journal*, vol. 49, pp. 115–120, 1970.

[13] 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.

[14] L. D. Lanzerotti, J. C. Sturm, E. Stach, R. Hull, T. Buyuklimanli, and C. Magee, "Suppression of boron outdiffusion in SiGe HBTs by carbon incorporation," in *Proc. 1996 IEEE International Electron Devices Meeting (IEDM 1996)*, Dec. 1996, pp. 249 – 252.

[15] B. Heinemann *et al.*, "A low-parasitic collector construction for high-speed SiGe:C HBTs," in *Proc. 2004 IEEE International Electron Devices Meeting (IEDM 2004)*, Dec. 2004, pp. 251 – 253.

[16] A. J. Joseph *et al.*, "Status and direction of communication technologies-SiGe BiCMOS and RFCMOS," *Proc. IEEE*, vol. 93, pp. 1539–1558, Sept. 2005.

[17] J. J. Ebers and J. L. Moll, "Large-signal behavior of junction transistor," *Proc. IRE*, vol. 42, pp. 1761–1772, 1954.

[18] T. E. Idleman, F. S. Jenkins, W. J. McCalla, and D. O. Pederson, "SLIC- a simulator for linear integrated circuits," *IEEE J. Solid-State Circuits*, vol. 6, pp. 188 – 203, Aug. 1971.

[19] L. W. Nagel and D. O. Pederson, "Simulation program with integrated circuit emphasis (SPICE)," in *Proc. 16th Midwest Symposium on circuit theory*, Waterloo, Canada, Apr. 1973.

[20] T. K. Ning, "History and future perspective of the modern silicon bipolar transistor," *IEEE Trans. Electron Devices*, vol. 48, pp. 2483–2491, Nov. 2001.

[21] I. Getreu, *Modeling the Bipolar Transistor*. Amsterdam: Elsevier, 1978.

[22] D. O. Pederson, "Recent developments in electronic circuit simulators," in *Proc. International Symposium on circuit theory*, Los Angeles, USA, Apr. 1972, pp. 105–111.

[23] H. K. Gummel and H. C. Poon, "An integral charge control model of bipolar transistors," *Bell Syst Tech. Journal*, vol. 49, pp. 827–852, 1970.

[24] J. M. Early, "Effect of space-charge layer widening in junction transistors." *Proc. IRE*, vol. 40, pp. 1401–1406, 1952.

[25] H. C. de Graaff and F. M. Klaasen, *Compact Transistor Modeling for Circuit Design*. New York: Springer-Verlag, 1990.

[26] G. M. Kull, L. W. Nagel, S. Lee, P. Lloyd, E. J. Prendergast, and H. Dirks, "A unified circuit model for bipolar transistors including quasi-saturation effects," *IEEE Trans. Electron Devices*, vol. 32, pp. 1103–1113, June 1985.

[27] H. C. Graaff and W. J. Kloosterman, "New formulation of the current and charge relations in bipolar transistor modeling for cacd purposes," *IEEE Trans. Electron Devices*, vol. 32, pp. 2415–2419, Nov. 1985.

[28] H. Stubing and H.-M. Rein, "A compact physical large-signal model for high-speed bipolar transistor at high current densities Part I: one-dimensional model," *IEEE Trans. Electron Devices*, vol. 34, pp. 1741–1751, Aug. 1987.

[29] C. McAndrew *et al.*, "VBIC95, the vertical bipolar inter-company model," *IEEE J. Solid-State Circuits*, vol. 31, pp. 1476 – 1483, Oct. 1996.

[30] D. Sheridan, M. R. Murty, W. Ansley, and D. Harame, "IBM bipolar model standardization: VBIC-MEXTRAM-HICUM," Compact Model Council (CMC)," Tech. Report, Apr. 2002.

[31] J. C. J. Paasschens, W. J. Kloosterman, and R. van der Toorn, "Model derivation of Mextram 504. The physics behind the model," Koninklijke Philips Electronics N.V. 2005, Nat. lab. Unclassified Report NL-UR 2002/806, Mar. 2005.

[32] S. Mijalkovic and J. N. Burghartz, "Compact modeling of SiGe HBTs," in *Proc. 12th European gallium arsenide & other compound semiconductor application symposium (GAAS 2004)*, Oct. 2004, pp. 247 – 250.

[33] M. Schroter, M. Friedrich, and H.-M. Rein, "A generalized integral charge-control relation and its application to compact models for silicon-based HBTs," *IEEE Trans. Electron Devices*, vol. 40, pp. 2036–2046, Nov. 1993.

[34] J. C. J. Paasschens, W. J. Kloosterman, and R. J. Havens, "Modeling two SiGe HBT specific features for circuit simulation," in *Proc. 2001 IEEE Bipolar/BiCMOS Circuits and Technology Meeting (BCTM 2001)*, Minneapolis, MN, USA, Oct. 2001, pp. 38–41.

[35] V. Cuoco, "Smoothie-a model for linearity optimization of fet devices in rf applications," Delft University of Technology, PhD dissertation ISBN:978-90-6464-017-9, 2006.

[36] Q. J. Zhang and K. C. Gupta, *Neural Networks for RF and Microwave Design.* Boston, MA: Artech House, 2000.

[37] J. D. Cressler and G. Niu, *Silicon-Germanium Heterjunction Bipolar Transistors.* Boston, MA: Artech House, 2003.

[38] H. C. de Graaff, W. J. Kloosterman, and T. N. Jansen, "Compact bipolar transistor model for CACD with accurate description of collector behaviour," in *Proc. 18th Conf. Solid St. Dev. and Material*, Tokyo, Japan, 1986, p. 287.

[39] H. C. de Graaff and W. J. Kloosterman, "The Mextram bipolar transistor model," Koninklijke Philips Electronics N.V., Nat. lab. Unclassified Report NL-UR 006/94, 1994.

[40] J. C. J. Paasschens and W. J. Kloosterman, "The Mextram bipolar transistor model - Level 504.6," Koninklijke Philips Electronics N.V. 2000/2005, Nat. lab. Unclassified Report NL-UR 2000/811, Mar. 2005.

[41] S. L. Salmon, D. Cressler, R. C. Jaeger, and D. L. Harame, "The influence of Ge grading on the bias and temperature characteristics of SiGe HBT's for precision analog circuits," *IEEE Trans. Electron Devices*, vol. 47, pp. 292–298, Feb. 2000.

[42] L. C. N. de Vreede, H. C. de Graaff, K. Mouthaan, M. de Kok, J. L. Tauritz, and R. G. F. Baets, "Advanced modeling of distorsion effects in bipolar transistors using the Mextram model," *IEEE J. Solid-State Circuits*, vol. 31, pp. 114–121, Jan. 1996.

[43] G. Niu, J. D. Cressler, and A. J. Joseph, "Quantifying neutral base recombination and the effects of collector-base junction traps in UHV/CVD HBT's," *IEEE Trans. Electron Devices*, vol. 45, pp. 2499–2504, Dec. 1998.

[44] H.-C. Wu, S. Mijalkovic, and J. Burghartz, "A unified parameter extraction procedure for scalable bipolar transistor model Mextram," in *Proc. NSTI Nanotech 2006*, Boston, MA, USA, May 2006, pp. 872–875.

[45] H. C. Wu, S. Mijalković, and J. N. Burghartz, "Parameter extraction of scalable Mextram model for high-speed SiGe HBTs," in *Proc. 2004 IEEE Bipolar/BiCMOS Circuits and Technology Meeting (BCTM 2004)*, Montreal, Canada, Sept. 2004, pp. 140–143.

[46] H. C. Wu, S. Mijalkovic, and J. N. Burghartz, "A referenced geometry based configuration scalable Mextram model for bipolar transistors," in *Proc. 2006 IEEE International Workshop on Behavioral Modeling and Simulation (BMAS 2006)*, San Jose, CA, USA, Sept. 2006, pp. 50–55.

[47] P. Deixler *et al.*, "Exploration for high perfomance SiGe-heterojunction bipolar transistor integration," in *Proc. 2001 IEEE Bipolar/BiCMOS Circuits and Technology Meeting (BCTM 2001)*, Minneapolis, MN, USA, Oct. 2001, pp. 30–33.

[48] U. Erben and E. Sonmez, "Fully differential 5 to 6 GHz low noise amplifier using SiGe HBT technnology," *Electronics Letters*, vol. 40, pp. 39–40, 2004.

[49] W. C. Hua, T. Y. Yang, and C. W. Liu, "The comparison of isolation technologies and device models on SiGe bipolar low noise amplifier," *Applied Surface Science*, vol. 224, pp. 425–428, 2004.

[50] S. W. Yoon *et al.*, "A distributed scalable SiGe power device large-signal model based on Mextram 504," in *2004 IEEE Radio Frequency Integrated Circuits (RFIC) Symposium*, Fort Worth, TX, USA, June 2004, pp. 175 – 178.

[51] F. V. Rijs, S. Bertonnaud, R. Vanoppen, and R. Dekker, "RF power large-signal modeling with Mextram," in *Proc. 1996 IEEE Bipolar/BiCMOS Circuits and Technology Meeting (BCTM 1996)*, Minneapolis, MN, USA, Sept. 1996, pp. 57–60.

[52] F. Yuan, Z. Pei, J. W. Shi, S. T. Chang, and C. W. Liu, "Mextram modeling of Si/SiGe heterojunction phototransistors," in *Proc. 2003 International Semiconductor Device Research Symposium*, Washington, DC, USA, Dec. 2003, pp. 92–93.

[53] C. C. McAndrew and L. W. Nagel, "Early effect modeling in SPICE," *IEEE J. Solid-State Circuits*, vol. 31, pp. 136–138, Jan. 1996.

[54] S. Mijalković, "Generalized early factor for compact modeling of bipolar transistors with non-uniform base," *Electronics Letters*, vol. 39, pp. 1757–1758, 2003.

[55] H. C. de Graaff and W. J. Kloosterman, "Modeling of the collector epilayer of a bipolar transistor in the Mextram model," *IEEE Trans. Electron Devices*, vol. 42, pp. 274–282, Feb. 1995.

[56] J. C. J. Paasschens, W. J. Kloosterman, R. J. Havens, and H. C. de Graaff, "Improved compact modeling of output conductance and cutoff frequency of bipolar transistors," *IEEE J. Solid-State Circuits*, vol. 36, pp. 1390–1398, Sept. 2001.

[57] M. Reisch, *High-frequency Bipolar Transistors*. Berlin: Springer-Verlag, 2003.

[58] G. Rey, "Effets de la défocalisaion sur le comportement des transistors á jonctions," *Solid-State Electronics*, vol. 12, no. 8, pp. 645–659, 1969.

[59] H. Groendijk, "Modeling base crowding in a bipolar transistor," *IEEE Trans. Electron Devices*, vol. 20, pp. 329–330, Mar. 1973.

[60] W. Shockley and W. T. Read, "Statistics of recombination of holes and electrons," *Phys. Rev*, vol. 87, no. 5, pp. 835–842, 1952.

[61] R. N. Hall, "Electron-hole recombination in Germanium," *Phys. Rev*, vol. 87, no. 2, p. 387, 1952.

[62] W. J. Kloosterman and H. C. de Graaff, "Avalanche multiplication in compact bipolar transistor model for circuit simulation," *IEEE Trans. Electron Devices*, vol. 36, pp. 1376–1380, July 1989.

[63] W. J. Kloosterman, J. C. J. Paasschens, and R. J. Havens, "A comprehensive bipolar avalanche multiplication compact model for circuit simulation," in *Proc. 2000 IEEE Bipolar/BiCMOS Circuits and Technology Meeting (BCTM 2000)*, Minneapolis, MN, USA, Oct. 2000, pp. 172–175.

[64] A. G. Chynoweth, "Ionization rates for electron and holes in silicon," *Phys. Rev.*, vol. 109, pp. 1537–1540, 1958.

[65] H.-C. Wu, S. Mijalkovic, and J. N. Burghartz, "Mixed compact and behavior modeling using AHDL Verilog-A," in *Proc. 2003 IEEE International Workshop on Behavioral Modeling and Simulation (BMAS 2003)*, San Jose, CA, USA, Oct. 2003, pp. 139–143.

[66] M. Kleeme and E. Barke, "An extended transistor model for substrate crosstalk analysis," in *Proc IEEE 1999 Custom Integrated Circuit Conference*, San Diego, CA, USA, May 1999, pp. 579–582.

[67] L. C. N. de Vreede, H. C. de Graaff, J. L. Tauritz, and R. G. F. Baets, "Extension of the collector charge description for compact bipolar epilayer models," *IEEE Trans. Electron Devices*, vol. 45, pp. 277–285, Jan. 1998.

[68] S. M. Sze, *Semiconductor Devices, Physics and Technology*. New York: John Wiley, 1985.

[69] J. C. J. Paasschens, "Compact modelling of the noise of a bipolar transistor under DC and AC current crowding conditions," *IEEE Trans. Electron Devices*, vol. 51, pp. 1483–1495, Sept. 2004.

[70] M. C. A. M. Koolen and J. C. J. Aerts, "The influence of non-ideal base current on 1/f noise behavior of bipolar transistors," in *Proc. 1990 IEEE Bipolar/BiCMOS Circuits and Technology Meeting (BCTM 1990)*, Minneapolis, MN, USA, Sept. 1991, pp. 232–235.

[71] J. C. J. Paasschens and R. de Kort, "Modelling the excess noise due to avalanche multiplication in (hetero-junction) bipolar transistors," in *Proc. 2004 IEEE Bipolar/BiCMOS Circuits and Technology Meeting (BCTM 2004)*, Montreal, Canada, Sept. 2004, pp. 108–111.

[72] R. J. McIntyre, "Multiplication noise in uniform avalanche diodes," *IEEE Trans. Electron Devices*, vol. 13, pp. 164–168, Jan. 1966.

[73] K. Kundert and O. Zinke, *The Designer's Guide to Verilog-AMS*. Boston: Kluwer Academic Publishers, 2004.

[74] *Advanced Design System 2005A*, Agilent Technologies, 2005.

[75] B. Troyanovsky, P. O'Halloran, and M. Mierzwinski, "Analog RF model development with Verilog-A," in *2005 IEEE MTT-S International Microwave Symposium*, Long Beach, CA, USA, June 2005, pp. 287 – 290.

[76] J. C. J. Paasschens, W. J. Kloosterman, and R. J. Havens, "Parameter extraction for the bipolar transistor model Mextram - Level 504," Koninklijke Philips Electronics N.V. 2001, Nat. lab. Unclassified Report NL-UR 2001/801, May 2001.

[77] G. A. M. Hurkx, "On the modelling of tunneling currents in reverse-biased P-N junctions," *Solid-State Electronics*, vol. 32, no. 8, pp. 665–668, 1989.

[78] M. Schroter, Z. Yan, T. Y. Lee, and W. Shi, "A compact tunneling current and collector breakdown model," in *Proc. 1998 IEEE Bipolar/BiCMOS Circuits and Technology Meeting (BCTM 1998)*, Minneapolis, MN, USA, Sept. 1998, pp. 203–206.

[79] J. C. J. Paasschens, "A compact model for the emitter-base tunneling current of bipolar transistors," Philips Research Technical Report, 2004.

[80] M. Pfost and H.-M. Rsin, "Modeling and measurement of substrate coupling in Si-Bipolar IC's up to 40 GHz," *IEEE J. Solid-State Circuits*, vol. 33, pp. 582–591, Apr. 1998.

[81] S. Fregonese, D. Celi, T. Zimmer, C. Maneux, and P. Y. Sulima, "A scalable substrate network for compact modeling of deep trench insulated HBT," *Solid-State Electronics*, vol. 49, no. 10, pp. 1623–1631, 2005.

[82] S. D. Harker, R. J. Havens, J. C. J. Paasschens, D. Szmyd, L. F. Tiemeijer, and E. F. Weagel, "An S-parameter technique for substrate resistance characterization of rf bipolar transistors," in *Proc. 2000 IEEE Bipolar/BiCMOS Circuits and Technology Meeting (BCTM 2000)*, Minneapolis, MN, USA, Sept. 2000, pp. 176–179.

[83] J. C. H. Lin *et al.*, "State of art RF/Analog foundry technology," in *Proc. 2002 IEEE Bipolar/BiCMOS Circuits and Technology Meeting (BCTM 2002)*, Minneapolis, MN, USA, Sept. 2002, pp. 73–79.

[84] G. J. Coram, "How to (and how not to) write a compact model in Verilog-A," in *Proc. 2004 IEEE International Workshop on Behavioral Modeling and Simulation (BMAS 2004)*, San Jose, CA, USA, Oct. 2004, pp. 97–106.

[85] *Agilent 4155C/4156C user's guide: Measurement and Analysis*, Agilent Technologies, 2001.

[86] *Precision DC probe holder for advanced on-wafer process and device characterization*, Cascade Microtech, 2003.

[87] *Agilent 4284A precision LCR meter 20 Hz $\sim$ 1 MHz*, Agilent Technologies, 1988.

[88] D. MacSweeney, K. G. MaCarthy, L. Floyd, A. Mathewson, P. Hurley, J. A. Power, and S. C. Kelly, "Influence of probing configuration and data set size for bipolar junction capacitance determination," in *Proc. 2002 IEEE Int. Conference on Microelectronics Test Structures (ICMTS 2002)*, Cork, Ireland, Apr. 2002, pp. 127–132.

[89] *Network Analyzer Basics*, Agilent Technologies, 1998.

[90] *8510C Network Analyzer Operating and Programming Manual*, Agilent Technologies, 1998.

[91] *On-Wafer Vector Network Analyzer Calibration and Measurements*, Cascade Microtech, 1998.

[92] *11612V Option K21 Bias Network User's Guide*, Agilent Technologies, 2004.

[93] *SussCal 5.0 calibration software for Vector Network Analyzers*, SUSS MicroTec, 2002.

[94] *WinCal 3.0 calibration software for Vector Network Analyzers*, Cascade Microtech, 2001.

[95] P. J. van Wijnen, H. R. Claessen, and E. A. Wolsheimer, "A new straightforward calibration and correction procedure for "On wafer" high frequency S-Parameter measurements (45 MHz - 18 GHz)," in *Proc. 1987 IEEE Bipolar/BiCMOS Circuits and Technology Meeting (BCTM 1987)*, Minneapolis, MN, USA, Sept. 1987, pp. 70–73.

[96] M. C. A. M. Koolen, J. A. M. Geelen, and M. P. J. G. Versleijen, "An improved de-embedding technique for on-wafer high-frequency characterization," in *Proc. 1991 IEEE Bipolar/BiCMOS Circuits and Technology Meeting (BCTM 1991)*, Minneapolis, MN, USA, Sept. 1991, pp. 188–191.

[97] H. Cho and D. E. Burk, "A three-step method for the de-embedding of high-frequency S-parameter measurements," *IEEE Trans. Electron Devices*, vol. 38, pp. 1371–1375, June 1991.

[98] T. E. Kolding, "A four-step method for de-embedding g igahertz on-wafer CMOS measurement," *IEEE Trans. Electron Devices*, vol. 47, pp. 734–740, Apr. 2000.

[99] E. P. Vandamme, M. M. P. Schreurs, and C. van Dinther, "Improved three-step de-embedding method to accurately account for the influence of pad parasitics in silicon on-wafer RF test-structures," *IEEE Trans. Electron Devices*, vol. 48, pp. 737–742, Apr. 2001.

[100] L. F. Tiemeijer and R. J. Havens, "A calibrated lumped-element de-embedding technique for on-wafer rf characterization of high-quality inductors and high-speed transistors," *IEEE Trans. Electron Devices*, vol. 50, pp. 822–829, Mar. 2003.

[101] V. Rizzoli, F. Mastri, and C. Cecchetti, "Computer-aided noise analysis of MESFET and HEMT mixers," *IEEE Trans. Microwave Theory Tech.*, vol. 37, pp. 1401–1410, Sept. 1989.

[102] S. Bousnina, C. F. P. Mandeville, A. B. Kouki, and F. M. Ghannouchi, "An accurate on-wafer de-embedding technique with application to HBT devices characterization," *IEEE Trans. Microwave Theory Tech.*, vol. 50, pp. 420–424, Feb. 2002.

[103] Q. Liang, J. D. Cressler, G. Niu, Y. Lu, G. Freeman, D. C. Ahlgren, R. M. Malladi, K. Newton, and D. L. Harame, "A simple four-port parasitic de-embedding methodology for high-frequency scattering parameter and noise characterization of SiGe HBTs," *IEEE Trans. Microwave Theory Tech.*, vol. 51, pp. 2165–2174, Nov. 2003.

[104] L. F. Tiemeijer, R. J. Havens, A. B. M. Jansman, and Y. Bouttement, "Comparison of the "pad-open-short" and "open-short-load" deembedding techniques for accurate on-wafer RF characterization of high-quality passives," *IEEE Trans. Microwave Theory Tech.*, vol. 53, pp. 723–729, Feb. 2005.

[105] A. Blaum, O. Pilloud, G. Scalea, J. Victory, and F. Sischka, "A new robust on-wafer 1/f noise measurement and characterization system," in *Proc. 2001 IEEE Int. Conference on Microelectronic test structures (ICMTS 2001)*, Kobe, Japan, Mar. 2001, pp. 125–130.

[106] S. M. Sze, *Semiconductor Devices Physics and Technology 2nd Edition.* New York: John Wiley & Sons, Inc., 2001.

[107] W. M. C. Sansen and R. G. Mayer, "Characterization and measurement of the base and emitter resistance of bipolar transistors," *IEEE J. Solid-State Circuits*, vol. 7, pp. 492 – 498, Dec. 1972.

[108] S. Lee, B. R. Ryum, and S. W. Kang, "A new parameter extraction technique for small-signal equivalent circuit of ploysilicon emitter bipolar transistors," *IEEE Trans. Electron Devices*, vol. 41, pp. 233–138, Feb. 1994.

[109] T. Fuse and Y. Sasaki, "An analysis of small-signal and large-signal base resistance for submicrometer BJT's," *IEEE Trans. Electron Devices*, vol. 42, pp. 534–539, Mar. 1995.

[110] W. J. Kloosterman, J. C. J. Paasschens, and D. B. M. Klaassen, "Improved extraction of base and emitter resistance from small-signal high frequency admittance measurement," in *Proc. 1999 IEEE Bipolar/BiCMOS Circuits and Technology Meeting (BCTM 1999)*, Minneapolis, MN, USA, Sept. 1999, pp. 93–96.

[111] J. R. Hauser, "The effect of distributed base potential on emitter-current injection density and effective base resistance for stripe transistor geometries," *IEEE Trans. Electron Devices*, vol. 11, pp. 238–242, May 1964.

[112] H. Groendijk, "Modeling base crowding in a bipolar transistor," *IEEE Trans. Electron Devices*, vol. 20, pp. 329–330, Mar. 1973.

[113] H. M. Rein and M. Schroter, "Experimental determination of the internal base resistance of bipolar transistors under forward-bias conditions," *Solid-State Electronics*, vol. 34, no. 3, pp. 301–308, 1991.

[114] M. Schroter, "Simulation and modeling of the low-frequency base resistance of bipolar transistors and its dependence on current and geometry," *IEEE Trans. Electron Devices*, vol. 38, pp. 538–544, Mar. 1991.

[115] *IC-CAP 2006*, Agilent Technologies, 2006.

[116] T. K. Ning and D. D. Tang, "Method for determining the emitter and base series resistances of bipolar transistors," *IEEE Trans. Electron Devices*, vol. 31, pp. 409–412, Apr. 1984.

[117] H. C. Wu, S. Mijalkovic, and J. N. Burghartz, "Extraction of collector resistances for device characterization and compact models," *Solid-State Electronics*, vol. 50, no. 9, pp. 1475–1478, 2006.

[118] D. B. M. Klaassen, "A unified mobility model for device simulation-II. temperature dependence of carrier mobility and lifetime," *Solid-State Electronics*, vol. 35, no. 7, pp. 961–967, 1992.

[119] G. D. J. Smit, R. van Langevelde, A. J. Scholten, D. B. M. Klaassen, G. Gildnblat, X. Li, H. Wang, and W. Wu. (2006) PSP 102.0 technical note. [Online]. Available: http://www.nxp.com/models/mos_models/psp/

[120] X. Xi, M. Dunga, J. He, W. Liu, K. M. Cao, X. Jin, J. J. Ou, M. Chan, A. M. Niknejad, and C. Hu. (2004) Bsim4.5.0 user's manual. [Online]. Available: http://www-device.eecs.berkeley.edu/bsim3/bsim4.html

[121] K. M. Walter *et al.*, "A scaleable, statistical SPICE Gummel-Poon model for SiGe HBT's," *IEEE J. Solid-State Circuits*, vol. 33, pp. 1439–1444, Sept. 1998.

[122] A. Chakravorty, R. F. Scholz, D. Knoll, A. Fox, B. Senapati, and C. K. Maiti, "Implementation of a scalable VBIC model for SiGe:C HBTs," *Solid-State Electronics*, vol. 50, no. 3, pp. 399–407, 2006.

[123] *Virtuoso Spectre Circuit Simulator User Guide*, Cadence, 2005.

[124] *Hspice Simulation and Analysis User Guide*, Synopsys, 2006.

[125] *Eldo Verilog-A User's Manual*, Mentor Graphics, 2006.

[126] T. Hashimoto *et al.*, "Direction to improve SiGe BiCMOS technology featuring 200-GHz SiGe HBT and 80-nm gate CMOS," in *Proc. 2003 IEEE International Electron Devices Meeting (IEDM 2003)*, Dec. 2003, pp. 129 – 132.

[127] M. Hsieh *et al.*, "Frequency response improvement of 120 GHz f_T SiGe HBT by optimizing the contact configurations," in *2004 IEEE MTT-S International Microwave Symposium*, Fort Worth, Texas, USA, June 2004, pp. 967 – 970.

[128] M. Latif and P. R. Bryant, "Network analysis approach to multidimensional modeling of transistors including thermal effects," *IEEE Trans. Computer-Aided Design*, vol. 1, pp. 94 – 101, Apr. 1982.

[129] N. Nenadovic *et al.*, "Extraction and modeling of self-heating and mutual thermal coupling impedance of bipolar transistors," *IEEE J. Solid-State Circuits*, vol. 39, pp. 1764 – 1772, Oct. 2004.

[130] J. C. J. Paasschens, "Modeling distributed resistance effect," Private Communication.

[131] *Taurus-Process User Guide, Version W2004.9*, Synopsys, 2004.

[132] *DIOS TCAD Release 8.0*, Integrated Systems Engineering AG, Zurich, 2002.

[133] *ATHENA 2004 Release Notes*, Silvaco, 2004.

[134] *Taurus-Device User Guide, Version W2004.9*, Synopsys, 2004.

[135] *MINIMOS-NT 2.0 User's Guide*, Technische Universitat Wien, Austria, 2002.

[136] *ATLAS 2004 Release Notes*, Silvaco, 2004.

[137] V. Palankovski and S. Selberherr, "The state-of-the-art in simulation for optimization of SiGe-HBTs," *Applied Surface Science*, vol. 224, no. 1-4, pp. 312–319, 2004.

[138] M. Rickelt and H.-M. Rein, "A novel transistor model for simulating avalanche-breakdown effects in Si bipolar circuits," *IEEE J. Solid-State Circuits*, vol. 37, pp. 1184–1197, Sept. 2002.

Summary

Title: A Scalable Mextram Model for Advanced Bipolar Circuit Design
By: Hsien-Chang Wu

With the blossom of the wireless mobile technology starting from the 90's, Si/SiGe bipolar devices are widely used in radio frequency integrated circuits (RFICs) due to their superior high-frequency performance compared to CMOS devices for a given process generation. In this thesis, a general geometry scaling approach and an efficient parameter extraction procedure to support advanced bipolar circuit design is presented. The thesis is summarized as follows.

In Chapter 1, the evolution of bipolar transistor technology from the early point contact transistor to the most advanced SiGe:C HBT, as well as the main applications of bipolar transistors in each period of time, are described. Then, it is discussed how compact model development has been driven by technology progress. It is further mentioned how the different generations of bipolar models are classified and what major differences between the subsequent generations of models are.

In Chapter 2, the detailed branch currents, charges and noises sources in the equivalent circuit of the Mextram model associated with their physical meaning are described.

In Chapter 3, a "short-cut" for compact model implementation on basis of the high-level language AHDL Verilog-A is introduced. Compared to the handcrafted C-code approach, the Verilog-A approach drastically reduces the time needed to implement a compact model as well as the effort and time needed to test the implementation. The Verilog-A Mextram model shows promising accuracy and simulation speed compared to a C-code Mextram model. Moreover, two additional features (tunneling current and behavioral substrate resistance) for the Mextram model are implemented in Verilog-A demonstrating that it is very suitable for practicing new modeling ideas.

In Chapter 4, the bipolar transistor measurements (including DC, CV, S-parameters and 1/f noise), test structure design for the device measurement and pad parasitic de-embedding are discussed. As a result, a transistor array is used as a way to increase the measured capacitance value of a small geometry transistor for CV measurement. Different de-embedding methods including open, open-short and 4-port in combination with dummy open, short and loads structures are introduced for de-embedding at different frequency ranges. Following the device measurements and test structure design, different measurement setups for the relevant Mextram model parameters are used to demonstrate single device parameter extraction from a high-speed SiGe HBT.

In Chapter 5, a reference based geometry scalable model and its parameter extraction is proposed. The scaling rules for the scalable Mextram model are based on device physics and re-formulated in terms of reference parameters, geometry parameters and drawn sizes in the device layout for geometry scaling. Then, the geometry scaling rules are incorporated in the standard Mextram model implemented in Verilog-A language. Therefore, the scalable model can be employed in the ADS simulator and the parameters can be systematically extracted from the measured data of various geometries within an IC-CAP model file, which is extended from the single device parameter extraction procedure. The new scalable model is tested with the same high-speed SiGe HBT technology used in Chap. 4. The simulation data from the new scalable model show good fit to CV, DC, f_T and S-parameters measurements from the SiGe HBTs of various geometries.

In Chapter 6, an extension of the geometry scalable Mextram model to the configuration scalable model including different emitter numbers and collector contacts for high-power application is described. The distributed temperature and internal collector voltage are modeled by their average values. Therefore, the configuration scalable model uses almost the same set of model parameters as the geometry scalable model except additional geometry parameters representing the mutual heating effect and the collector resistance between two adjacent emitter fingers and additional instance parameters representing different layout configurations. The configuration model nicely predicts the f_T-degradation measured from the SiGe HBTs due to an increase of the mutual heating and the collector delay time with increasing emitter numbers.

Finally, Chapter 7 presents the conclusion of the thesis and some suggestions for the future work on the Mextram model.

Samenvatting

Een Schaalbaar Mextram Model voor Geavanceerd Bipolair Circuit Ontwerp

By: Hsien-Chang Wu

In de draadloze mobiele communcatietechnologie, die bloeit sinds de negentiger jaren van de twintigste eeuw, worden Si en SiGe bipolaire transistoren breed toegepast in *geïntegreerde circuits voor radio-frequenties* (RFIC's), vanwege hun superieure hoog-frequent prestaties vergeleken met CMOS componenten voor een gegeven procesgeneratie. In dit proefschrift worden een algemene aanpak van geometrie-schaling en een efficiënte parameterextractieprocedure gepresenteerd ter ondersteuning van geavanceerd bipolair circuit ontwerp. Het proefschrift wordt hier als volgt samengevat.

In hoofdstuk 1 wordt de evolutie van de bipolaire transistor-technologie beschreven, vanaf de vroege puntcontacttransistor tot de meest geavanceerde SiGe:C heterojunctie-bipolaire transistor (HBT), samen met de voornaamste toepassing van bipolaire transistoren in elk tijdvak. Vervolgens wordt besproken hoe de ontwikkeling van compacte modellen telkens gedreven werd door de voortgang van de technologie. Voorts wordt aangegeven hoe de verschillende generaties van bipolaire modellen worden geclassificeerd en wat de belangrijke verschillen tussen de opeenvolgende generaties van modellen zijn.

In hoofdstuk 2 worden de details beschreven van de stromen, ladingen en ruisbronnen in het equivalente circuit volgens het Mextram model, in samenhang met een beschrijving van hun fysische betekenis.

In hoofdstuk 3, wordt een laagdrempelige technologie geïntroduceerd voor implementatie van compacte modellen. Het betreft toepassing van de hogere programmeertaal *AHDL Verilog-A*. Vergeleken met implementatie in handgeschreven C-code vergt een implementatie in Verilog-A aanzienlijk minder inspanning en tijd en ook de inspanning en tijd die nodig zijn voor het testen van de implementatie worden bij implementatie in Verilog-A drastisch gereduceerd. De implementatie in Verilog-A van het Mextram model laat een

veelbelovende nauwkeurigheid en simulatiesnelheid zien, vergeleken met een implementatie in C-code van het model. Voorts worden twee toegevoegde aspecten (tunnelstroom en effectieve substraatweerstand) van het Mextram model geïmplementeerd in Verilog-A, als demonstratie van de grote geschiktheid ervan voor het uitwerken van nieuwe model ideeën.

In hoofdstuk 4 worden metingen aan bipolaire transistoren bediscussieerd – dit omvat DC, CV, S-parameters en $1/f$-ruis – alsmede het ontwerp van teststructuren voor metingen aan *devices* en de correctie van gemeten waarden voor parasitaire invloeden van contactstructuren. Een matrix van transistoren wordt gebruikt om de te meten capaciteit te vergroten van kleine transistoren, ten bate van CV-metingen. Verschillende *de-embedding* methoden – *open*, *open-short* en *4-poort* in combinatie met *dummy-open, short* en *load* structuren – worden geïntroduceerd ten bate van *de-embedding* in verschillende frequentie-regimes. Ten bate van extractie van Mextram model-parameters voor individuele transistoren worden verschillende meet-configuraties gebruikt. Dit wordt gedemonstreerd aan de hand van een SiGe HBT voor hoog-frequent toepassingen.

In hoofdstuk 5 wordt een model voorgesteld, inclusief een parameterextractiemethodiek, voor geometrie-schaling, een zogenaamd *referentie-gebaseerd, geometrie-schaalbaar model*. De schaalregels voor het schaalbare Mextram model zijn gebaseerd op *device*-fysica en geformuleerd in termen van referentieparameters, geometrieparameters en getekende maten van de *device layout*. De geometrie-schaalregels zijn opgenomen in het standaard Mextram model door middel van een implemenatie in de Verilog-A taal. Het schaalbare model kan zo worden gebruikt binnen een *IC-CAP model file* met een *ADS*-circuit simulator en de parameters kunnen systematisch worden geëxtraheerd uit de gemeten data van transistoren van verschillende geometrische afmetingen. Het betreft een uitbreiding van de parameterextractieprocedure voor individuele *devices*. Het nieuwe schaalbare model is getest met de zelfde SiGe HBT hoogfrequent technologie die ook in hoofdstuk 4 is gebruikt. De simulaties met het nieuwe schaalbare model laten een goede *fit* zien aan CV, DC, f_T en S-parameter metingen aan SiGe HBT's van verschillende afmetingen.

Hoofdstuk 6 bevat een beschrijving van een uitbreiding van het *geometrie-schaalbaar Mextram model* naar een *configuratie-schaalbaar model*, dat bedoeld is voor transistoren met meerdere emittervingers en meerdere collectorcontacten. De ruimtelijk verdeelde temperatuur en interne collectorspanning worden erin gerepresenteerd door hun respectievelijke gemiddelde waarden. Bijgevolg heeft het *configuratie-schaalbare model* vrijwel dezelfde model-para-

meterset als het *geometrie-schaalbare Mextram model*. Toegevoegd zijn alleen parameters voor de wederzijdse opwarming van emitter-vingers, voor de collector-weerstand tussen emitter-vingers, alsmede parameters die verschillende configuraties representeren. Het configuratie-model geeft een bevredigende voorspelling van de f_T*-degradatie* die wordt waargenomen bij SiGe HBT's bij toenemend aantal emitter-vingers ten gevolge van de toenemende onderlinge opwarming van de vingers en de toenemende collector *delay-time*.

Hoofdstuk 7, ten slotte, presenteert de conclusies van de dissertatie en enkele suggesties voor toekomstig werk aan het Mextram model.

List of publication

Journal papers

1. H. C. Wu, S. Mijalkovic and J. N. Burghartz, "Extraction of Collector Resistances for Device Characterization and Compact Models," Solid-State Electronics, Vol. 50, Issues 9-10, pp. 1475-1478, 2006.
2. H. C. Wu, J. B. Kuo, "A compact velocity-overshoot model for deep-submicron bipolar devices considering energy transport," IEEE Transactions on Electron Devices, Vol. 45, No. 2, February 1998, pp. 417-422.
3. B. Y. Chen, H. C. Wu, J. B. Kuo, "A 2D simulation study on the d.c. and transient behaviour of a SiGe-base heterojunction bipolar device with an extended germanium profile into the n-epi region with application to an ECL buffer," Solid-State Electronics, Vol. 39, No. 6, June 1996, pp. 935-939.
4. H. C. Wu, S. Mijalkovic and J. N. Burghartz, "A Reference Geometry Based Scaling Approach for Bipolar Transistor Model Mextram," To be submitted to IEEE Transactions on Electron Devices.

Conference Proceedings

1. H. C. Wu, S. Mijalkovic and J. Burghartz, "A Reference Geometry Based Scaling Approach for Bipolar Transistor Model Mextram," in Proc. IEEE Eurocon 2005, Nov. 2005, pp. 1239-1242.
2. A. Kautaniuk, H. C. Wu, S. Mijalkovic, J. N. Burghartz, "Verification of Mextram 504.6 in Verilog-A and C-code (SiMKit 2.1.1) implementation," in Proc. SAFE 2005, November 2005, Veldhoven, the Netherlands, pp. 19-22.
3. H. C. Wu, S. Mijalkovic, J. N. Burghartz, "Extraction of Collector Resistance for Compact Models," in Proc. SAFE 2004, November 2004, Veld-

hoven, the Netherlands, pp. 808-811.
4. H. C. Wu, S. Mijalkovic, J. N. Burghartz, "Parameters Extraction of a Scalable Mextram Model for High-Speed SiGe HBTs," in Proc. 2004 IEEE Bipolar/BiCMOS Circuits and Technology Meeting (BCTM 2004), September 12-14, 2004, Montreal, Canada, pp. 140-143.
5. V. Cuoco, W. C. E. Neo, L. C. N. de Vreede, H. C. de Graaff, L. K. Nanver, H.-C. Wu, H. F. F. Jos, J. N. Burghartz, "A New Extraction Technique for the Series Resistances of Semiconductor Devices Based on the Intrinsic Properties of Bias-Dependent Y-Parameters," in Proc. 2004 IEEE Bipolar/BiCMOS Circuits and Technology Meeting (BCTM 2004), September 12-14, 2004, Montreal, Canada, pp. 148-151.
6. H. C. Wu, S. Mijalkovic, J. N. Burghartz, "Extended Mextram Model To Wide Frequency Range," in Proc. SAFE 2003, November 25-26, 2003, Veldhoven, the Netherlands, ISBN 90-73461-39-1, pp. 668-671.

Workshops

1. H. C. Wu, S. Mijalkovic and J. N. Burghartz, "A Referenced Geometry Based Configuration Scalable Mextram Model for Bipolar Transistors," in Proc. 2006 IEEE International Behavioral Modeling and Simulation Workshop (BMAS 2006), Sep. 14-15, 2006, San Jose, CA, USA, pp 50-55.
2. H. C. Wu, S. Mijalkovic and J. N. Burghartz, "A Unified Parameter Extraction Procedure for Scalable Bipolar Transistor Model Mextram," in Proc. Workshop on Compact Modeling in NSTI Nanotech 2006, Boston, MA, U.S.A, Vol. 3, pp. 872-875.
3. S. Mijalkovic and H. C. Wu, "Efficient Generation of Reduced-Order Circuit and Device Models for Wide Frequency Applications," in Proc. Workshop on Compact Modeling in NSTI Nanotech 2006, Boston, MA, U.S.A, Vol. 1, pp. 713-716.
4. H. C. Wu, S. Mijalkovic, J. G. Macias, J. N. Burghartz, "Mixed Compact and Behavior Modeling Using AHDL Verilog-A," in Proc. 2003 IEEE International Workshop on Behavioral Modeling and Simulation (BMAS 2003), Oct. 7-8, 2003, San Jose, CA, USA, pp. 139-143.

Contributed talks

1. "A Reference Geometry and Configuration Scalable Mextram Model for Bipolar Transistors," in Presentation Notes of the Workshop on Compact Modeling for RF/Microwave Applications, Maastricht, the Netherlands, October, 2006.
2. "Extraction of Collector Resistance for Mextram Compact Transistor Model" in Presentation Notes of the Workshop on Compact Modeling for RF/ Microwave Applications, Montral, Canada, September, 2004.

Other publication

1. H. C. Wu, "A Referenced Based Scalable Mextram Model for advanced circuit design, " in DIMES White Papers 2006.

Acknowledgments

I still remember the date when I arrived in the Netherlands for a Ph.D. interview. It was a cold and rainy morning at 9 Jan. 2002. When I got off the train at Delft central station, the first idea came to my mind is to go back to the Schipol airport and took the same flight back to Taiwan for the sunny weather. But somehow I still decided to continue my journey. When the taxi dropped me off near the hotel, it was still dark. I couldn't find my hotel since there is no big fascia or road sign like in Taiwan. While standing in the rain with my luggage and searching for my hotel, there was a lady coming to me and asked if I needed help. After telling her my problem, she showed me where the hotel is. My first impression on Dutch people is their friendliness. When I entered the hotel, I was totally wet. So, the reception smiled at me and asked: are you swimming here? Of course not, I replied helplessly. I know it, I just try to make you feel easy he replied. Unfortunately, the hotel rooms are totally full now so you have to wait until 9:00 when somebody checks out. Then, we can prepare a room for you as soon as possible. Before that, you can enjoy the delicious breakfast that we prepared for you. I sit down and hope everything can finish soon.

After a good sleep at the date of arrival, I went to DIMES the next day. I got a warm welcome from Jeanne. She briefly introduced DIMES to me while I was waiting for the first interview. Though my interview was quite successful, I still couldn't make up my mind to move to the Netherlands immediately. After all, it is much more challenging to study in the Netherlands than to study in Taiwan. After discussing with Prof. Joachim Burghartz, he allowed me to stay in the Netherlands for the first two years and then went back to Taiwan to finish the rest of my Ph.D. But it turned out that I liked the environment so much so I spent almost five years in the Netherlands. I really appreciate Prof. Joachim Burghartz and Prof. Kees Beenakker for giving me the chance to pursue my Ph.D. degree. I would like to thank Prof. Joachim

Burghartz and Dr. Slobodan Mijalkovic for their guidance in the past five years.

Besides, I also would like to thank following people for helping my Ph.D. thesis come true:
Dr. Reinout Woltjer and Dr. Leo de Vreede for giving me "more and more" time to finish my Ph.D. thesis.
Dr. Ramses van der Toorn and Dr. Jeroen Paasschens for helpful discussions and thorough review on my Ph.D. thesis.
Prof. Lis Nanver and Dr. Henk van Zeijl for kindly offering their bipolar process knowledge.
Dr. Dirk Klaassen for supporting Mextram Model standardization.
My former manager Mr. Yu-Tai Chia in TSMC for his support of my study in the Netherlands.

It was not always smoothly during my Ph.D. study. I would like to thank following people for assisting my work:
Atef Akhnoukh, Marco Pelk and Peter Swart for assisting me in the measurement lab.
Jeanne-Marie Sirks-Wilking, Marysia Lagendijk, Birgit Rademakers, Pien Rijnink, Marian Roozenburg, Nicolette Dingeldein, Nathalie van Benthem and Willem van der Sluys for helping the administration work.
Jan Chris for setting up my unix environment.

Though a Ph.D. degree is a best wish of my life, I find the experience of working with different people is what I enjoy most. However, Chinese people used to say "There is no endless feast in the world". Therefore, I will cherish the friendship with my colleagues: Dr. Yan Zhuang, Dr. Behzad Rejaei, Dr. Marina Vroubel, Dr. Marian Bartek, Marco, Mark, Huseyin, Saoer, Jason, Koen, Sebastian, Edmund, Gabriel, Pedram, Ming, Nobu, Trinh, Jia, Mohamed, Tuncay, Victor, Alessandro, Francesco and many others in HiTeC, ECTM and DIMES. I wish them all the best in the future.

I also would like to thank my Taiwanese and Dutch friends who make me feel not so lonely and helpless in the Netherlands:
Mr. and Mrs. Lai in the Chinese Church Rotterdam who invite me to join their monthly family gathering. So, I can not only enjoy lots of delicious food but also know that God loves me.
Mr. and Mrs. van den Enden-Verhahen for offering me a very comfortable accommodation and helping me in my daily life.
Yu-Jhe Yang, Chung-Kai Yang and Huai-Wen Wu, who are Ph.D. candidates in EWI faulty. I wish them good luck during their Ph.D. study.

Jun-Yi Huang for designing the cover page of my Ph.D. thesis.

Finally, I would like to have a very special thank to:
My parents, who devote their love to me and give their unconditional support to my work.
My sister, who helps me to manage miscellaneous work in Taiwan and send me plenty of food supplies from Taiwan.
My brother, who gives me many ideas about the future and tells me his working experience in China.

About the Author

Hsien-Chang Wu, was born in Taipei, Taiwan in November 1970. He got his bachelor degree from Tatung University in 1994 and Master degree from National Taiwan University in 1996 both in Electrical Engineering. In 1996, he served in the communication unit of Taiwan, R.O.C. Army as a second lieutenant for two years. After finishing his military service, he joined SPICE model department of Taiwan Semiconductor Manufacturing Company (TSMC) as a semiconductor device modeling engineer in charge of test structure design and building model library for both SiGe HBTs and CMOS devices for the foundry customers.

In September 2002, he joined Laboratory of High frequency Technology and Components (HiTeC) Group of TU Delft in the Netherlands as a research assistant and working toward his Ph.D. degree. His Research focused on advanced methodologies for silicon bipolar transistors (BJTs) and silicon germanium hetero junction bipolar transistors (SiGe HBT) characterization, efficient parameter extraction procedures, geometry scaling of the model parameters and statistical modeling; Development and utilization of the compact model for the next generation of Si/SiGe bipolar transistors in advanced RF/Microwaves circuit design.

During his academic years, he also joins many international bipolar modeling activities. One of them is organized by Compact Modeling Council (CMC) in bipolar model standardization and evaluation of new model features reported from different semiconductor companies. The other one is in year 2003

to 2006, he has helped to organize four very successful workshops on compact modeling, which were held in France, Canada, USA and the Netherlands respectively.