Compact Modeling of Physical Mechanisms Present in the Weak Breakdown Regime of Advanced Si and SiGe Bipolar Transistors

Robert Setekera

Compact Modeling of Physical Mechanisms Present in the Weak Breakdown Regime of Advanced Si and SiGe Bipolar Transistors

PROEFSCHRIFT

ter verkrijging van de graad van doctor
aan de Technische Universiteit Delft,
op gezag van de Rector Magnificus Prof. ir. K. C. A. M. Luyben,
voorzitter van het College voor Promoties,
in het openbaar te verdedigen

op dinsdag 21 januari 2016 om 15:00 uur

door

Robert SETEKERA

Professional Doctorate in Engineering
van Eindhoven University of Technology, The Netherlands,
geboren te Mukono, Uganda.

Dit proefschrift is goedgekeurd door de promotor:

Prof. dr. ir. A. W. Heemink

Copromotor:

Dr. ir. R. van der Toorn

Samenstelling promotiecommissie:

Rector Magnificus,	voorzitter	Technische Universiteit Delft
Prof. dr. ir. A. W. Heemink,	promotor	Technische Universiteit Delft
Dr. ir. R. van der Toorn,	copromotor	Technische Universiteit Delft

Onafhankelijke Leden:

Prof. dr. ir. G. Jongbloed	Technische Universiteit Delft
Prof. dr. ir. J. W. Slotboom	Technische Universiteit Delft
Prof. dr. ing. L. C. N. de Vreede	Technische Universiteit Delft
Prof. dr. H. W. A. Schilders	Technische Universiteit Eindhoven
Dr. T. Vanhoucke	NXP Semiconductors, Eindhoven

Robert Setekera,
Compact Modeling of Physical Mechanisms Present in the Weak Breakdown Regime of Advanced Si and SiGe Bipolar Transistors,
PhD Thesis, Delft University of Technology,
with summary in Dutch.

Keywords: Avalanche, bipolar transistors, junction breakdown, non-local avalanche, compact modeling, parameter extraction, thermal resistance, base resistance, impact-ionization, relaxation length, self-heating, HBT, physical effects, temperature scaling.

ISBN 978-00-0000-000-0

Copyright © 2016 by Robert Setekera
All right reserved. No part of this publication may be reproduced, stored in a retrieval system, or transmitted in any form or by any means without the written permission of the copyright owner.

An electronic version of this dissertation is available at
`http://repository.tudelft.nl/`.

For what shall a it profit a man, if he shall gain the whole world, but lose his own soul?

- Mark 8:36

Contents

1 Introduction **1**
1.1 SiGe HBTs for advanced RF applications 2
1.2 Junction breakdown mechanism in SiGe HBTs 6
1.2.1 Local and non-local avalanche mechanisms 12
1.2.2 Self-heating mechanism 14
1.2.3 Forward Early effect 15
1.3 Semiconductor device modeling 16
1.3.1 Device simulation 17
1.3.2 Compact semiconductor device modeling for IC simulation 18
1.3.3 Mextram compact model for Si BJTs and SiGe HBTs 21
1.4 Thesis Motivation 26
1.5 Thesis aim/purpose 29
1.6 Thesis outline 30
References 31

2 Extraction of the base and thermal resistance of bipolar transistors **37**
2.1 Introduction 38
2.2 Derivation of the R_B and R_{TH} extraction method 39
2.3 Experimental application and evaluation of the method 44
2.3.1 Measurement setup 44
2.3.2 Determination of the extraction region 46
2.3.3 Extraction of both R_B and R_{TH} 52
2.4 Comparison to other R_B and R_{TH} extraction methods 54
2.4.1 DC extraction methods 54
2.4.2 RF extraction methods 58
2.4.3 Noise measurements method 59
2.5 Self consistency check of the extraction method 62
2.5.1 Self consistency check for R_{TH} 64
2.5.2 Self consistency check for R_B 65
2.5.3 Summary 67
2.6 Summary 68
References 71

3 Compact model for non-local avalanche effect in bipolar transistors **75**
3.1 Introduction 76
3.2 Compact model for non-local avalanche current 77
3.2.1 Derivation of the electron temperature 77
3.2.2 Non-local impact-ionization rate 78
3.2.3 Multiplication factor: $(M_n - 1)$ 79

3.2.4 Approximate analytical integration of sharply peaked functions . . . 80
3.2.5 Approximate expression for the multiplication factor $(M_n - 1)$ 82
3.2.6 Temperature scaling of model parameters 83
3.3 Implementation of the non-local avalanche model in Mextram 84
3.4 Summary 86
References 87

4 Experimental assessment of physical effects in weak avalanche regime 91
4.1 Introduction 92
4.2 Experimental measurement setup 94
4.3 Independent physical effects that interfere in weak avalanche regime . . . 97
4.4 Verification results for Si bipolar junction transistor 104
4.5 Error analysis of the non-local avalanche compact model 112
4.5.1 Dependence of the RMS error measure on the model parameters . . 112
4.5.2 Dependence of parameters on software's numerical accuracy 114
4.5.3 Effect of stochastic measurement errors on extracted parameters . . 116
4.5.4 Summary 117
4.6 Verification results for SiGe HBTs 118
4.6.1 Results corresponding to a high speed QUBiC4X SiGe-HBT 118
4.6.2 Results corresponding to a high voltage SiGe-HBT 120
4.6.3 Results corresponding to a very high voltage SiGe-HBT 122
4.6.4 Summary 124
4.7 Magnitude of non-local avalanche effect in different device types 125
4.8 Summary 126
References 127

5 Conclusions and Recommendations 133
5.1 Conclusions 134
5.2 Recommendations for future work 138
References 141

A Functions used for the non-local avalanche compact model 145
A.1 Integration of a Gaussian function 145
A.2 Smooth bounded functions 146
References 147

Summary 149

Samenvatting 152

List of Publications 155

Acknowledgements 157

About the Author 159

1

Introduction

In this introductory chapter, we give an overview of the motivation for the work in the thesis. A summary of silicon-germanium (SiGe) bipolar transistor technology is introduced with focus on present day industrial device applications. The advantages of the silicon-germanium transistor technology over the other existing transistor technologies, which makes it to a competing technology and suitable for industrial applications are discussed. Tradeoffs in relation to other transistors application requirements such as that of the peak cutoff frequency and breakdown voltages are discussed. The junction breakdown mechanism in bipolar transistors is briefly discussed with focus on transistor scaling and carrier multiplication for a biased transistor; this forms the foundation for the work in the thesis. Some of the possible physical mechanisms present in junction breakdown regime are presented together with their impact on the terminal transistor characteristics. We briefly discuss modeling approaches in semiconductor devices and their application in integrated circuit design process. The compact modeling process of semiconductor devices which is commonly used in commercial integrated circuit design tools is briefly introduced, with merits of having physic-based compact models in industrial applications. An example of a physics-based standard compact model used for simulation of present day bipolar transistors is discussed; shortcomings in this compact model in relation to the transistor characteristics in the weak collector-base breakdown regime are demonstrated. Such observable model setbacks lead to the motivation and purpose of the thesis, which are discussed here. The thesis outline is presented at the end of this chapter.

1

1.1 SiGe HBTs for advanced RF applications

The rapid growth in the demand for faster and multi-functional communication systems (e.g., low power portable cellular phones with (high-speed) internet connection) in the recent years has motivated continuous research efforts for transistor technologies which are capable of delivering: low power, high linearity, low noise, high-speed of operation, and low production cost. Silicon-Germanium (SiGe) heterojunction bipolar transistor (HBT) technology has emerged as one of the key technologies that can address most of these concerns. Its integration into bipolar complementary metal-oxide-semiconductor (BiCMOS) technologies to get SiGe BiCMOS technology [1, 2] has led to establishment of such a technology as a leading contender for a host of circuit applications such as analog, mixed signal, radio-frequency (RF), and millimeter-waves (mm-waves). Present-day SiGe BiCMOS processes integrate high performance HBTs with modern CMOS devices on a single chip, which provides cost-efficient solutions to many of the building blocks of RF and microwave transceivers [3]. In this, SiGe has shown that it compares with the high performance requirements achieved by III-V compound semiconductors e.g., GaAs HBT while using Si-based fabrication processes similar to silicon CMOS [4, 5].

SiGe technology shows enormous potential for bringing all the benefits of Si semiconductor device technology firmly into the high frequency world of analog electronics. The use of silicon provides better strength and thermal conductivity (than III-V devices), large-scale integration, large wafer size (processing can be done on 12-inch wafers which lowers the costs), and high-functionality density [6–8]. The fact that a standard Si production line and most of the standard bipolar process modules can be used for SiGe device fabrication, yields low production cost with high yield and excellent liability [7, 9, 10]. Since the first SiGe bipolar devices were demonstrated in late 1980's [11], several RF circuits and systems [1] applications (such as wireless communication, power amplifier, automotive, and fibre optics/cable communications) have driven advances in SiGe (SiGe BiCMOS) technology [12].

The key foundation of the modern SiGe BiCMOS technology goes back to theoretical work of H. Kroemer on heterojunction transistors (HBTs) published in 1957 [13]. In this work, it was postulated that with alloy grading, the energy bandgap could be altered such that the electrostatic force could be overcome by a quasi-electric field, thereby enhancing carrier transport. Though this theory (of bandgap engineering) was well known, the challenge of fabricating a very thin, high-quality SiGe layer in the base while maintaining a good control over Ge fraction, boron doping, and layer thickness, greatly hindered the realization SiGe HBTs until early 1990. This followed the studies done in later 1980s, which showed that a good heterojunction could be obtained if the SiGe layer was thin and Ge content relatively low (< 30%) [14]. Such made it possible to overcome the key barrier of a relatively large (~ 4.2%) lattice mismatch between Si and Ge, which had made it very difficult to form a heterojunction between Si and SiGe without generation of misfit dislocations at the interface. A combination of these studies and the improvements in the epitaxy techniques, made it possible to grow a thin, strained SiGe layer on top of silicon without generation of misfit dislocations [15–19]. Significant steps along the path to manufacturing of SiGe HBT included the first demonstration of high-frequency (75 GHz) operation of a SiGe HBT in a non-self aligned structure during early

1990 [10]. This result got significant worldwide attention, since the performance of SiGe HBT was roughly twice what state-of-the-art Si bipolar junction transistor (BJT) could achieve; thus it led to (ongoing) dedicated research on SiGe HBTs (SiGe BiCMOS technology), which resulted into a number of new applications (e.g., in wireless and optical fibre communication systems, cellular handsets, and radar systems) for such devices. Note that before the SiGe HBTs, heterojunction bipolar transistors were only available in III-V compound semiconductor technologies such as AlGaAs/GaAs [20, 21].

Compared to the conventional Si BJTs, SiGe HBTs have got several key advantages, such as high frequency performance, high-current gain, low noise figure, and increased Early voltage. These advantages are attributed to the extra degree of freedom provided by the introduction of a controlled amount of Ge in the neutral base. This enables bandgap engineering in a Si system to obtain impressive performance metrics. A well-engineered Ge profile in the base region of the transistor can effectively decouple key device parameters that otherwise would lead to critical performance tradeoffs in standard Si BJTs; thus resulting into improved overall performance of SiGe HBTs. This has led to SiGe HBTs replacing Si BJTs in many present day bipolar transistor applications.

Effect of SiGe HBTs scaling on transistor performance

The standard figures-of-merit for dynamic transistor performance are the cutoff frequency f_T (the frequency at which the extrapolated common-emitter current gain of the transistor becomes unity) and the maximum oscillation frequency $f_{\max}$ (the frequency at which the unilateral power gain of the transistor becomes unity) [22]. The cutoff frequency f_T has an inverse relation with the emitter-to-collector transit time, τ_{EC}, and it is usually described as [22]

$$f_T = \frac{1}{2\pi\tau_{EC}} = \frac{1}{2\pi}\left[\frac{kT}{q_e I_C}(C_{BE}+C_{CB}) + \tau_B + \tau_E + \tau_C + \frac{W_{CB}}{2v_{\text{sat}}} + (R_C+R_E)C_{CB}\right]^{-1}, \quad (1.1)$$

where I_C is the collector current, C_{BE} and C_{CB} are the base-emitter and collector-base depletion capacitances, respectively, τ_B, τ_E, τ_C are the quasi-neutral base, emitter, and collector transit times, respectively, W_{CB} is the collector-base depletion width, v_{sat} the saturation drift velocity, R_C the parasitic resistance of the collector, R_E the parasitic resistance of the emitter, q_e is the elementary charge, T the ambient temperature, and k the Boltzmann constant. The maximum oscillation frequency is related to f_T by

$$f_{\max} = \sqrt{\frac{f_T}{8\pi C_{CB} R_B}}, \quad (1.2)$$

where R_B is the parasitic resistance of the base.

Expression (1.1) shows that to achieve high f_T, the quantities in the square brackets have to be minimized, and this is achieved in RF SiGe-HBTs by employing both vertical scaling (to reduce the intrinsic delays in the device) and horizontal scaling (to reduce the device parasitics). We note that vertical scaling largely consists of reducing the neutral base, increasing the slope of germanium profile, and finally increasing the collector doping concentration [12]. On the other hand, horizontal scaling mostly consists of reducing

1

the lateral emitter width [23]. Other than improving the device speed (f_T), scaling also improves other important figures-of-merit of the device such as noise performance [10].

By grading the doping of Ge through the base [10, 12], a built-in electric field is induced across the base that decreases the base transit time (τ_B) since charge carriers are now accelerated across the base. This increases the maximum cutoff frequency f_T of the SiGe-HBT compared with a conventional Si-BJT. Addition of Ge in the base also leads to an increase in the current gain (β), which can be raised beyond tens-of-thousands if a typical doping profile of the Si-BJT is maintained. This increase in the current gain is due to the fact that in SiGe-HBT the conduction band barrier is lower than that in Si-BJT [10, 16, 17, 24], thus the concentration of the injected electrons into the base (for an NPN transistor in forward active mode) is much higher (several orders of magnitude) in SiGe-HBT than in Si-BJT. The current gain for a SiGe-HBT is expressed as:

$$\beta_{\text{SiGe}} = \beta_{\text{Si}} \exp\left(\frac{\Delta E_g(x)}{kT}\right), \tag{1.3}$$

where $\Delta E_g(x)$, is the bandgap difference between the emitter and the base of the transistor, i.e., $\Delta E_g = E_{g,\text{Si}} - E_{g,\text{SiGe}}$, with $E_{g,\text{Si}}$, the bandgap in the Si emitter and $E_{g,\text{SiGe}}$ the bandgap in the SiGe base. In SiGe-HBTs, $\Delta E_g(x)$ can be made much higher than kT; which would mean β_{SiGe} will be much higher than β_{Si} by a factor of $\exp\left(\Delta E_g(x)/kT\right)$. However, in normal device applications, such excessively large current gain is not necessarily favored, as it leads to degradation in the breakdown voltages (which limits the overall SiGe-HBT performance) without additional benefits for circuit applications [1, 10]. For most circuit applications, the current gain of about $\beta = 100$ is sufficient; so any extra gain larger than this is not much desired [1, 24]. The real advantage in the device design of SiGe-HBTs, however, lies in the fact that the extra current gain can be traded for the benefit of RF characteristics of the device. A typically practiced trade-off is to increase the base doping concentration in favor of enhanced operation speed [1, 5, 9]. The increased base doping will lower the current gain [25], but will provide benefits such as reduction in the base resistance leading to low noise figure, reduced base layer width without emitter-collector punch-through [16] leading to low τ_B, thus high peak f_T, and reduced influence of the Early effect [24]. As can be seen from (1.2) base resistance R_B is an important parameter in determining f_{max}, with a lower value R_B necessary for achieving high f_{max}. Base resistance R_B will be one of the key quantities of focus in Chapter 2 of this thesis.

Though vertical scaling of SiGe-HBTs results into high performance (in terms of maximum f_T), as we mentioned earlier it lowers the transistor breakdown voltages (BV_{CEO} and BV_{CBO}), which has raised considerations for the circuit designers, mostly those for high-power amplifiers [9, 26]. Present day advanced RF SiGe-HBTs typically show collector-emitter breakdown voltage with an open-base BV_{CEO} of around 1.5 V. This value is not necessarily the upper limit of the voltage allowed across the emitter and collector in real circuit applications [1]; rather the collector-base breakdown voltage with an open-emitter BV_{CBO} imposes an upper limit for possible transistor operation voltage. In most today's RF applications which target moderate operation frequencies relative to peak f_T, usually the high peak f_T is traded off for other benefits such as reduced power consumption, high breakdown voltage, and reduced noise [23].

An example of the tradeoff between f_T and breakdown voltage is demonstrated in Fig. 1.1 [23], which corresponds to four subsequent generation of Jazz SiGe BiCMOS devices. Here, we can see that for each generation of these devices, a tradeoff between f_T and BV_{CEO} is applied, where high values of f_T are realized for lower values of BV_{CEO} and vice verse. Reducing the collector doping concentration [23] can lower the base-collector field improving the breakdown but reducing f_T. In practice, this is accomplished by adding customized collector implant masks to create bipolar devices with higher breakdown voltages without impacting other devices built on the same wafer [23]. This technique is applied across the four generations of Jazz technology shown in Fig. 1.1. Here, each subsequent generation supports devices with higher f_T but also improves the tradeoff between f_T and breakdown voltage (BV_{CEO}) to realize high breakdown devices with higher f_T improving large-signal performance for applications such as integrated drivers and power amplifiers [23]. Therefore, in designing the SiGe-HBT (for e.g., high power

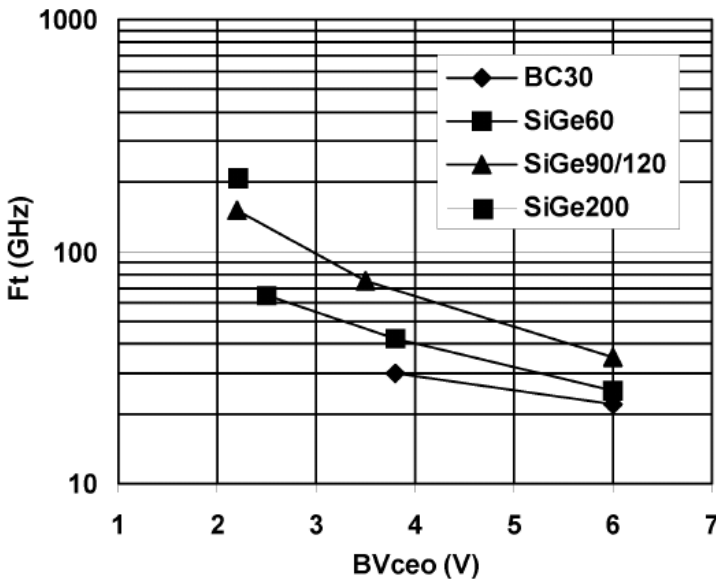


Figure 1.1: Unity cutoff frequency f_T verses BV_{CEO} plotted for four Jazz BiCMOS generations (figure get from [23]). Here, the devices are fabricated such that a tradeoff of f_T for high breakdown voltages is achieved, mostly by optimizing the collector doping enabling the integration of high speed and large-signal functions.

applications), a compromise has to be made between the requirements of breakdown voltage and speed (f_T). With regard to circuit and device prospective, these tradeoffs in modern HBT device design need to be well understood in order to achieve maximum transistor operation for a targeted application. An extended discussion of such tradeoffs in relation to junction breakdown of SiGe-HBT, is presented in the next section.

1

1.2 Junction breakdown mechanism in SiGe HBTs

For an NPN bipolar transistor biased in forward operation mode (i.e., base-emitter junction is forward biased ($V_{BE} > 0$) and base-collector junction is reverse biased ($V_{CB} < 0$)), the electron current is crossing the collector-base (CB) high field space charge region. In devices that have a high collector-epilayer doping N_{epi}, already under moderate values of V_{CB}, around its peak the electric field $E(x)$ may reach high values such that it causes the incident electrons in the depletion region to obtain high kinetic energy. When such electrons moving with high kinetic energy collide with the crystal lattice, the excess energy may be transferred to a bond electron in the valence band, promoting this carrier to the conduction band while leaving a hole behind, thus creating an electron-hole pair. This carrier generation process illustrated in Fig. 1.2 is known as impact-ionization [10, 27]. In some instances, the generated electron during impact-ionization process may gain enough kinetic energy due to the high electric field, and thus invoke an electron-hole pair upon collision with the crystal lattice; this process commonly known as avalanche multiplication [28] can go on, where the generated carriers by impact-ionization lead to generation of other charge carriers in the presence of high electric field (i.e., an avalanching process).

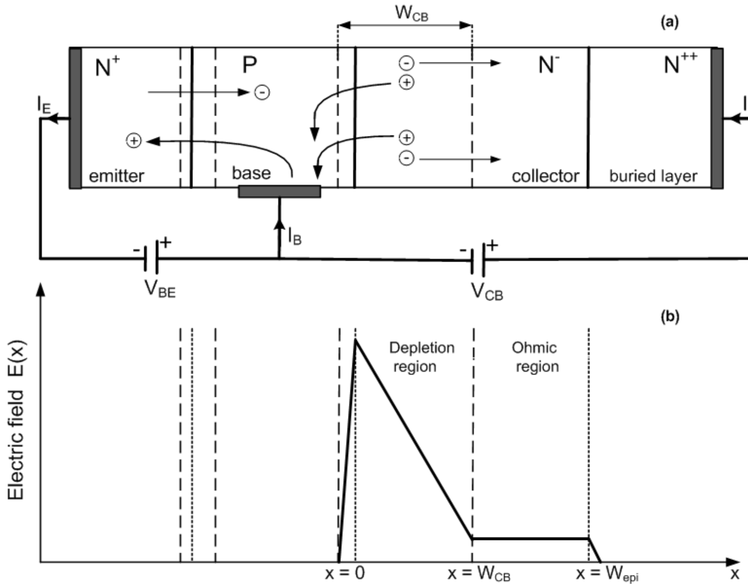


Figure 1.2: Schematic representation of an NPN bipolar transistor, in forward normal operation mode, i.e., base-emitter junction forward biased ($V_{BE} > 0$) and collector-base junction reverse biased ($V_{CB} < 0$). Here, W_{CB} is the bias dependent width of the collector-base depletion region. Due to high reverse collector-base junction bias, the (absolute value of the) electric field $E(x)$ (b) is distributed across the collector-epilayer, with maximum value at the collector-base metallurgical junction. Such high electric field leads to generation of electron-hole pairs by impact-ionization process.

The amount of generated carriers by impact-ionization largely depends on the elec-

tric field distribution $E(x)$ in the collector-base depletion region (see example in Fig. 1.2 (b)). The field $E(x)$ is affected by the collector-base junction bias V_{CB}, the collector current I_C [29], and the junction doping concentration N_{epi}. This electric field can be described by Poisson's equation [22]

$$\frac{dE(x)}{dx} = \frac{\rho(x)}{\varepsilon}, \quad \text{for} \quad 0 < x < W_{\text{epi}} \tag{1.4}$$

where ε is the permittivity and $\rho = q_e(N_{\text{epi}} - n + p)$ is the charge density, with q_e the elementary charge, n the electron concentration, and p the hole concentration. Here, the base is assumed to be highly doped with respect to the collector-epilayer (which is usually the case in SiGe-HBTs), thus the depletion region exists mostly in the collector-epilayer (i.e., we assumed a mono-junction). Due to the high electric-field in the C-B depletion region (above 3×10^5 V/cm [16, 30]), the drift velocity v_d of the electrons can be assumed to be saturated (i.e., $v_d = v_{\text{sat}}$), and the hole concentration can be assumed to be negligible. In such a case, the charge density can be written as $\rho = q_e N_{\text{epi}} - (|J_{\text{C}}|/v_{\text{sat}})$, where $|J_{\text{C}}|$ is the collector current density. Thus, the field gradient (1.4) can be rewritten as [29]:

$$\frac{dE(x)}{dx} = \frac{q_e N_{\text{epi}}}{\varepsilon}\left(1 - \frac{I_{\text{C}}}{I_{\text{hc}}}\right), \tag{1.5}$$

where we define the (well known [25]) hot-carrier current $I_{\text{hc}} = q_e N_{\text{epi}} A_E v_{\text{sat}}$, with A_E the emitter area. Over the ohmic region of the epilayer (see Fig. 1.2), the electric field is low enough (in comparison to that in the depletion region) to prevent velocity saturation. In such regions, the net charge density is zero and the number of electrons equals the dope N_{epi}, with a negligible number of holes also available [31]. The integral of the spatial electric field $E(x)$ over the collector-epilayer width, gives the voltage drop across the collector-base junction (i.e., $V_{CB} + V_{d_C}$, where V_{d_C} is the junction built-in voltage), i.e., the area under the curve in Fig. 1.2 (b). This shows that a modulation of the external bias V_{CB} directly affects the junction electric-field distribution and thus the rate of avalanche multiplication.

Other than the V_{CB} and N_{epi}, the electric field distribution in the collector-epilayer can also be modulated by the collector current I_C. At high base-emitter voltage V_{BE}, more charge carriers (electrons for NPN transistor) are injected from the emitter into the C-B space-charge region; this leads to a variation of the net charge density, and thus the electric field gradient (see (1.5)) [25, 32]. For low and moderate V_{BE} the peak of the electric field is located at the C-B metallurgical junction as demonstrated in Fig. 1.2 (b), and under the local field model, the electron and hole impact-ionization coefficients are highest at the C-B metallurgical junction. When V_{BE} increases, the peak electric field at the C-B junction decreases gradually, due to the high electron injection from the emitter, this will decrease the impact-ionization rate. With further electron injection, the peak electric field decreases until a uniform distribution is attained (at $I_C = I_{\text{hc}}$), and at this point, impact-ionization at the C-B junction reaches a minimum [33]. As the charge carrier injection increases further (i.e., at $I_C > I_{\text{hc}}$), a peak electric field appears at the collector-buried layer region (see Fig. 1.2). The field starts to rise again at very high V_{BE} and collector current densities (i.e., at $I_C = I_K$, where I_K is the collector current at onset

of Kirk effect [34]), in which case impact-ionization rate would increase again due to the increased peak electric field at the collector-buried layer interface [22]. For $I_C > I_K$, the electric field in the C-B space-charge region will be approximately zero and this will lead to diffusion of holes from the base into the collector-epilayer, a mechanism known as base push-out, which degrades the transistor performance [17]. With increased electric field strength at the collector-buried layer interface (or junction) due to increased collector current (i.e., high electron injection from the emitter into the collector), avalanche breakdown at this junction may finally occur. This is commonly known as the onset of *second breakdown* [35], which is destructive as the holes generated by avalanche multiplication at this junction are accelerated by the high field into the collector-epilayer. Due to base push-out, the collector-epilayer is already flooded with holes from the base, thus the holes generated in the collector-buried layer junction by avalanche multiplication will act as an additional base current; this will lead to injection of more electrons from the emitter in order to maintain a quasi-neutral condition in the (thick) base. This will lead to a positive-feedback loop, and such mechanism is usually destructive. Second breakdown mechanism will not be addressed in the thesis, thus we will not discuss it further.

As demonstrated in Fig. 1.2, for an NPN transistor biased in normal forward operation mode, the generated electrons and holes in the collector-base space charge region by impact-ionization drift to the quasi-neutral regions under the influence of the junction electric field. Here, the impact-ionization induced electrons drifts to the collector and they lead to an increase of the collector current [33]:

$$I_C = [(M_n - 1) + 1]\, I_{C_0}\,, \tag{1.6}$$

where $(M_n - 1)$ is the avalanche multiplication factor which increases with collector-base junction voltage V_{CB} and I_{C_0} is the (temperature dependent [22]) collector current in the absence of avalanche multiplication. The impact-ionization induced holes drifts to the quasi-neutral base. Since the hole current from the base to the emitter is fixed by the base-emitter voltage V_{BE}, excess holes are forced to flow to the base terminal, as shown in Fig. 1.2. The terminal base current I_B then decreases as (assuming neutral base recombination is negligible) [27]

$$I_B = I_{B_0} - (M_n - 1)\, I_{C_0}\,, \tag{1.7}$$

where I_{B_0} is the (temperature dependent [22]) base current in the absence of avalanche multiplication [36]. Note that for this case, the transistor forward current gain $\beta = I_{C_0}/I_{B_0}$. With increasing V_{CB}, more carriers are generated by impact-ionization, leading to high avalanche current $I_{\text{avl}} = (M_n - 1)\, I_{C_0}$, which eventually become equal to I_{B_0}, at which point $I_B = 0$. This would imply that

$$I_B/I_{B_0} = 0 = 1 - (M_n - 1)(I_{C_0}/I_{B_0}), \quad \text{or} \quad 1 = (M_n - 1)\beta\,, \tag{1.8}$$

which is equivalent to the condition of open-base breakdown BV_{CEO} [16], a standard breakdown figure-of-merit for bipolar transistors [10]. For further increased V_{CB} (i.e., $V_{CB} > BV_{CEO}$), I_{avl} becomes greater than I_{B_0}, then the terminal base current I_B will become negative (base current reversal), and in such a case, the avalanche multiplication

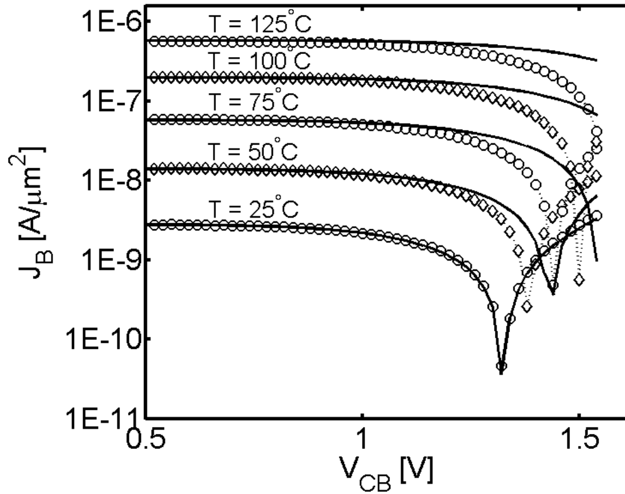
generated hole current flows out of the base terminal (see Fig. 1.2 (a)). Base current reversal is an important factor in (avalanche) breakdown related transistor instabilities at high V_{CB}, and such imposes limitations for compact modeling of the avalanche multiplication process [31, 37] in e.g., bipolar transistors. With further increase of V_{CB}, the C-B junction will eventually experience a full reverse biased junction breakdown at open-emitter breakdown voltage BV_{CBO}, a voltage which represents the absolute maximum (operational) collector voltage of the transistor [16, 32].

The behavior of the base current under the influence of avalanche multiplication, as described above, for low and moderate V_{CB} is demonstrated in Fig. 1.3, by the open symbols. The measurement results (symbols) correspond to a standard industrial RF (NPN) bipolar device of emitter area $A_E = 0.30 \times 20.7\ \mu m^2$, $BV_{CEO} \approx 1.3$ V, and $f_T / f_{max}(V_{CE} = 1.5V) \approx 120/150$ GHz. Here, the terminal base current I_B was measured as a function of V_{CB} and over a sequence of ambient temperatures i.e., $T = 25, 50, 75, 100$, and $125\,^{\circ}$C, for fixed $V_{BE} = 0.65$ V. From Fig. 1.3, we can observe that the measured base current $I_B = J_B / A_E$ is constant for low V_{CB} (since carrier multiplication is very small or negligible due to the low bias dependent junction electric field, thus I_{avl} is very small in comparison to the (external bias) fixed I_{B_0}) and then decreases significantly for high V_{CB} (since high carrier multiplication takes place due to the high bias induced junction electric field, leading to high I_{avl}, which even becomes larger than I_{B_0}), and even becomes negative. Running ahead of the introduction of the concept of simulation in Section 1.3.3 and especially Mextram in Section 1.3.3, Fig. 1.3 also shows (solid curves) computer simulated terminal characteristics. Shown are the best fitting results achievable with the industrial standard model Mextram [31, 38], release 504.10. Deviations can be observed between the symbols and the standard Mextram compact model simulations (solid curves) over the weak collector-base breakdown region, i.e., $V_{CB} > 0.8$ V; these discrepancies form the motivation for the work in the thesis, as will be discussed later in Section 1.4. In the two figures, the measurement results (symbols) are the same, but the simulations results (solid curves) correspond to an optimization of the Mextram model [39, 40]: at a lower temperature i.e., $T = 25\,^{\circ}$C in Fig. 1.3(a) and at a higher temperature i.e., $T = 125\,^{\circ}$C in Fig. 1.3(b); this will be explained further in Section 1.4.

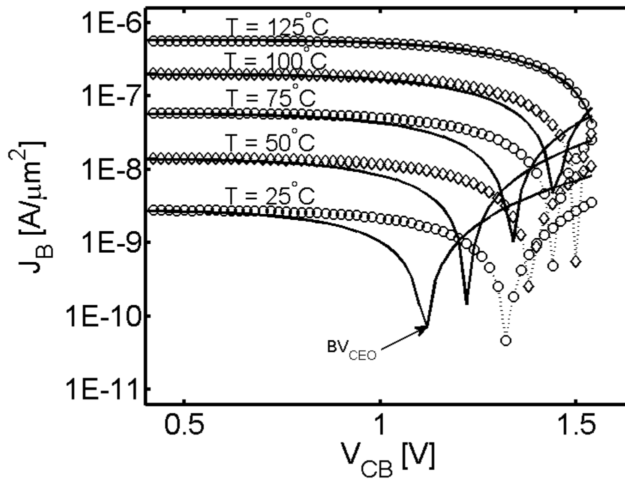
By increasing V_{BE}, the base current $I_B = I_{B_0} - I_{avl}$ can be made positive; this is because the hole diffusion current across the base-emitter junction increases at higher V_{BE} so that I_{B_0} becomes larger than I_{avl}. In Fig. 1.3, $I_B = J_B / A_E$ is observed to increase with temperature (for fixed V_{BE}), this is mainly due to the increase of the (main) transistor saturation current (I_S) with temperature, which will result into an increase of transistor main current I_N [31, 32], whose charge carriers are also involved in avalanche multiplication [38].

Dependence of breakdown voltage on transistor bias conditions

The speed of advanced SiGe-HBTs, as discussed, is normally increased at the expense of the breakdown voltage. This has forced circuit designers to operate such devices close or beyond the open-base collector-emitter breakdown voltage BV_{CEO} in order to meet the desired specifications. However, bipolar devices are fundamentally limited by the avalanche multiplication in the collector-base region [24], therefore, circuits should



(a)



(b)

Figure 1.3: Measured (symbols) and simulated (solid curves) collector-base breakdown characteristics. The model simulations correspond to the standard Mextram compact model [31]. Shown is the base current density (J_B) as a function of collector-base voltage (V_{CB}), for a sequence of ambient temperatures (i.e., $T = 25, 50, 75, 100$, and $125\,°$C), as observed for an advanced high-speed (NPN) SiGe-HBT, for fixed base-emitter voltage $V_{BE} = 0.65$ V. In (a), the Mextram model parameters i.e., $V_{\rm avl}$ and $W_{\rm avl}$ are optimized for target temperature $T = 25\,°$C, while in (b) they are optimized for $T = 125\,°$C. From (a) and (b), we can observe deviations between the measured and Mextram model simulated characteristics over the weak avalanche regime, i.e., $V_{CB} > 0.8$ V. In addition, it can be observed from (b) that the Mextram model underestimates the breakdown voltage BV_{CEO}.

be designed to always unconditionally operate below the open-emitter collector-base breakdown voltage BV_{CBO}.

For a forward biased (NPN) transistor (e.g., see Fig. 1.2 (a)), the generated holes by impact-ionization in the (high) reversed biased collector-base (C-B) junction drifts to the neutral base (under the influence of the junction electric field). These holes (in the base region) can either exit the transistor through the base terminal (as explained earlier) or they can stay within the transistor and diffuse to the emitter across the base-emitter (B-E) junction. The avalanche current due to holes which are injected into the emitter, is amplified by the forward current gain β across the B-E junction, and this in turn increases the collector current (i.e., trigger additional injection of electrons from the emitter into the collector through the base). The injected electrons will contribute to the increased avalanche multiplication in the C-B junction, hence, more avalanche generated holes will be injected into the emitter and this will cause a further increase of electron injection from the emitter; thus forming a positive-feedback loop, which can lead to a faster transistor breakdown [24].

The strength of the positive feedback, which modulates the breakdown voltage, increases with increasing external impedance seen by the base terminal, since (some of) the avalanche multiplication generated holes are increasingly forced to stay within the device with larger base terminal impedance [41] (which would resist the flow of the generated holes out of the transistor through the base terminal). Therefore, the configuration of the base connection, which affects the base terminal impedance, has a direct impact on the breakdown voltage [24, 41]. Open-base circuit configuration corresponds to infinite external impedance and maximized positive feedback, rendering the corresponding breakdown voltage BV_{CEO} to be the smallest breakdown voltage across collector and emitter; since the generated holes by junction breakdown cannot exit through the base, and end up accelerating the breakdown mechanism [1]. The opposite extreme happens with the base shorted to a grounded emitter, where the external impedance is effectively zero and the positive feedback is absent (the generated holes easily find exit to ground), leading to the corresponding collector-emitter breakdown BV_{CES} being the largest. This junction breakdown mechanism is basically the same as the open-emitter circuit collector-base junction breakdown (with corresponding breakdown voltage BV_{CBO}) [1, 41]. A general case is the one in which the base is connected to the grounded emitter through a finite resistance or impedance value, with corresponding emitter to collector breakdown voltage BV_{CER} [1]. We note that BV_{CER} should have a value between BV_{CEO} and BV_{CES}. BV_{CEO} and BV_{CBO}, imposes the lower and upper limits of transistors' breakdown voltages, and they are usually provided to circuit designers as part of the figures-of-merit of the transistor. The dependence of the bipolar transistor breakdown voltage on the source impedance is usually exploited in practical power amplifier (PA) design to significantly increase the safe operating voltage range.

A family of physical mechanisms in addition to avalanche multiplication may be present in the (collector-base) junction breakdown regime in SiGe HBTs [22, 24, 42–46]. Some of these mechanisms are already included in standard compact models for bipolar transistors, e.g., Mextram [38], and will be taken advantage of later in the thesis. We will

discuss *some* of these (significant and relevant) physical mechanisms in the next subsections.

1.2.1 Local and non-local avalanche mechanisms

The collector doping profile directly influences impact-ionization effects and breakdown voltage, which are important factors for power applications of SiGe HBTs. High-speed performance and breakdown voltage cannot be optimized independently. As we discussed earlier, the high collector doping level results in a high electric field in the collector-base (C-B) space-charge region of a (forward biased) bipolar transistor. Such high C-B junction electric field leads to generation of electron-hole pairs by impact-ionization process. Here, the incident charge carriers (electrons for an NPN transistor) in the C-B depletion region gain sufficient kinetic energy to invoke electron-hole pairs by impact-ionization; the generated electrons and holes are accelerated in opposite direction by the electric field. Due to the high electric field, such generated carriers can also gain enough kinetic energy to invoke further electron-hole pairs by impact-ionization and this can go on (i.e., avalanching). This avalanching process is commonly known as *strong* avalanche multiplication [24, 27]. If avalanching is assumed negligible, then the process is termed as *weak* avalanche multiplication [37, 47]; this is the common avalanche process addressed in standard compact models for bipolar transistors (e.g., Mextram [38] and HiCUM [37]), and we will follow the same trend in this thesis. To our present knowledge, there is no published literature addressing compact modeling of strong avalanche mechanism (in Si/SiGe bipolar transistors). This is motivated by the circumstance that a compact model that would incorporate a realistic model for strong avalanche effects would not be (generally) numerically stable: at and beyond the BV_{CBO} vertical asymptote. Such may lead to convergence issues of the compact model as a whole (when implemented in a circuit simulator) during circuit simulation.

The rate of impact ionization strongly depends on the electric field distribution in the collector-base space-charge region; with high avalanche multiplication taking place where the electric field is highest. In the so-called *local* perception of the avalanche multiplication process the electron and hole impact-ionization coefficients are assumed to actually and directly be functions of the electric field strength. This would imply locality of impact-ionization in the sense that avalanche is assumed to occur at locations of high electric fields (e.g., at the metallurgical C-B junction) [27], without any temporal or spatial delay.

In practical circumstances, for incident carriers (electrons) to gain the required sufficient energy to initiate impact-ionization, they must travel a *nonnegligible distance* within the collector under the influence of the electric field [33, 42, 48], thus ionization events do not occur in simple and instantaneous correspondence to maximum electric field. In other words, carriers gaining the threshold energy is not an instantaneous process, but a gradual process, and charge carriers travels a *certain distance* in the collector before attaining the threshold energy to initiate impact-ionization. This distance is known as the *energy relaxation length* λ_e, and such avalanche multiplication process is referred to as *non-local avalanche* [49, 50]. For an NPN bipolar transistor, under non-local avalanche process, the peak electron impact-ionization coefficient shifts from the

base-collector metallurgical junction into the collector, and the peak hole impact impact-ionization coefficient shifts from the base-collector metallurgical junction into the base [36]. In the local avalanche process, carriers multiplication is assumed to occur locally where the electric field is maximum. In spatially varying electric field, however, the carrier energy usually lags the electric field because of the finite energy relaxation time [36, 48] of the carriers. This would mean that the local avalanche process would yield large impact-ionization rate and thus low junction breakdown voltages [49], in contrast to the non-local avalanche process. In the non-local avalanche modeling process, impact-ionization coefficient is expressed in terms of (average) carrier energy or carrier temperature ($T_e(x)$), rather than the electric field (E(x)) [42], as will be elaborated in Chapter 3.

Dependence of avalanche multiplication process on ambient temperature

The applied ambient temperature directly affects the type of the avalanche multiplication process [33, 50] which occurs in a reversed biased collector-base junction of (an NPN) bipolar transistor in forward biased operation mode. This can be well understood by considering a cloud of electrons (conceived as a gas) moving across a crystal lattice of a semiconductor device under the influence of electric field $E(x)$, at a certain constant ambient temperature. These electrons will accelerate as a whole in the same direction, but they will also experience random motions or vibrations due to the thermal energy caused by the ambient temperature. Due to the work done by applied field on the electron cloud, electrons will also gain more kinetic energy, which will results into more collisions with the crystal lattice, as well as among themselves, and thus more chaotic random motion. This will result into an increase in the electron temperature ($T_e(x)$) [51], a temperature which is locally assigned to the electron cloud. Such as phenomena can be described by equation (3.1) of Chapter 3. The random motion of the electrons (moving with different velocities) within a cloud under the influence of the thermally vibrating crystal lattice and electric field takes sometime to stabilize. In this, thermal equilibrium is locally reached among the electrons within the cloud. The randomizing process has got a certain time scale and within this time, the electron cloud moves with a certain (average) velocity (approximately equal to the saturation velocity (v_{sat})) and travels a distance equal to the relaxation length (λ_e). As will be seen in Chapter 3, λ_e is a characteristic length scale over which the electron temperature increases when the electrons are forced by a macroscopic electric field.

When the ambient temperature is increased, thermal crystal lattice vibrations will increase as a result. The idea is now that a more intensely vibrating crystal lattice will have a stronger randomizing impact on the traveling electron cloud; hence the length-scale λ_e will be shorter at higher temperatures. Hence, at high temperatures λ_e will be relatively small in comparison to the available collector-base depletion width. As will be demonstrated in Chapter 4, this would imply that at lower ambient temperatures, the physical avalanche process will be dominated by non-local avalanche mechanism, while at higher temperatures the avalanche process will show a more local character. This is the one of the motivations for having two plots in Fig. 1.3, with the aim of showing the

results which can be obtained in industrial domain (Fig. 1.3(a)), and understanding the physics behind the avalanche breakdown mechanisms (Fig. 1.3(b)).

1.2.2 Self-heating mechanism

For biased transistor in operation (e.g., in normal forward operation mode, with base-emitter junction forward biased and collector-base junction reverse biased), it will dissipate power. The generated power will have an influence on the device temperature and its surroundings. Due to the power dissipation, the local device temperature will increase (or device will get warmer); this is called self-heating. To describe self-heating, we need to determine the dissipated power and how this is related to the increase in device temperature.

From device physics [24], power that flows in device can be calculated as the sum over currents times voltage drops. For the thesis, only power dissipation under time independent boundary conditions will be relevant; in this, by considering a three-terminal bipolar transistor (see Fig. 1.2), the dissipated power P_{diss} can be describe as

$$P_{\text{diss}} = I_C V_{CE} + I_B V_{BE}. \tag{1.9}$$

Since normally the collector current is larger than the base current and collector voltage is larger than the base voltage, the first term is usually dominant [31]. We note that not all power that flows in a transistor is dissipated; part of it will be stored as the energy on capacitors and is released later on. Therefore, to determine the dissipated power in a transistor, we need to add all the contributions of the dissipated elements, i.e., all DC-currents times voltage drops [31].

For reasons of computational efficiency, the increase ΔT in device temperature T due to the dissipated power P_{diss}, is in practice only approximately modeled as linear effect, i.e.,

$$\Delta T = R_{TH} P_{\text{diss}}, \tag{1.10}$$

where R_{TH} is a constant coefficient of proportionality, known as the thermal resistance. In circuit simulation then, the heat flow can be simulated fully analogous to electric ohmic conduction and is straight forward to implement by an equivalent electrical circuit. From relation (1.10), we can see that the thermal resistance can be determined as $R_{TH} = \Delta T / P_{\text{diss}}$, then the junction temperature can be described as:

$$T = T_{\text{amb}} + \Delta T, \tag{1.11}$$

where T_{amb} is ambient temperature. We can see that in order to address self-heating mechanism for a given transistor (device), the device thermal resistance R_{TH} need to be accurately determined. Self-heating affects the shape of the measured current-voltage characteristics, affects the two-port parameters (e.g., f_T and f_{max}) and precision analog circuits that depend on close transistor matching [17]; it may also lead to bias point instabilities [52, 53].

In integrated circuits where a number of devices are closely connected to one-another, the various devices do dissipate power and as a result, a device can cause the temperature of the neighboring devices to rise, i.e., devices do exchange heat and influence

each other's temperature; a mechanism commonly known as *mutual-heating* [31, 39]. These two mechanisms (self- and mutual-heating) are addressed in standard compact models for bipolar transistors e.g., Mextram [31], using a (connected) external thermal network [17, 39].

Though self-heating is a simple concept, addressing it requires an experimental determination of the device thermal resistance (R_{TH}), which is generally a recognized challenge in advanced industrial bipolar devices [54]. We will demonstrate in Chapter 2 that by using the (DC) measured characteristics (taken directly on present day Si and SiGe bipolar transistors) in the weak collector-base junction breakdown regime, a good estimate for R_{TH} can be determined.

When a high reverse bias is applied across the collector-base junction of a highly scaled bipolar transistor in forward operation mode, electron-hole pairs are generated by impact-ionization. This carrier multiplication process can lead to generation a large number of charge carriers and thus a relatively high avalanche current (I_{avl}), even for lower applied base-emitter bias ($V_{BE} > 0$). Such generated I_{avl} together with the high junction bias V_{CB} can lead to significant dissipated power in the transistor, and thus a significant contribution to the device self-heating [24, 31], even in the weak avalanche regime, where I_{avl} is assumed to be smaller than the total collector current (I_C). This will become more clear in Chapter 4, where we will demonstrate (and carefully address) the significance of self-heating on the measured characteristics in the weak avalanche regime, under low V_{BE} conditions.

1.2.3 Forward Early effect

The increase in the reverse collector-base junction bias (V_{CB}), leads to an increase in the collector-base depletion width (W_{CB}), which results into a modulation of the neutral base width of the bipolar transistor. As illustrated in Fig. 1.4, the modulation of the neutral base width with increased V_{CB} leads to a modulation of the minority carrier distribution in the base. With increased V_{CB}, the gradient of the injected minority carrier electron (for an NPN transistor) distribution in the base becomes steeper; since the electron diffusion is proportional to this gradient, hence an increase in V_{CB} leads to an increase in the transistor collector current I_C (see Fig. 1.2) [55, 56]. This mechanism is commonly referred to as (forward) Early effect [16, 17, 32].

Output characteristic of the transistor is also influenced by the Early effect; here, the increase in the transistor collector current with V_{CB} is seen as a finite slope of the output characteristic [32]. This is equivalent to a finite conductance at the output of the transistor; it is usually desirable in many circuit applications to have a low output conductance [16]. Though our focus has here is on the Early effect due to the reverse collector-base bias (i.e., *forward* Early effect), but the neutral base width (thus the collector current) can also be modulated by the variation of the base-emitter junction bias; this is termed as *reverse* Early effect [25].

The magnitude of forward Early effect on the transistor characteristics, can be reduced by careful transistor design, which is usually done by increasing the transistor base doping concentration and reducing the collector-epilayer doping concentration [55]. This would imply that when a high reverse bias (V_{CB}) is applied across the collector-base

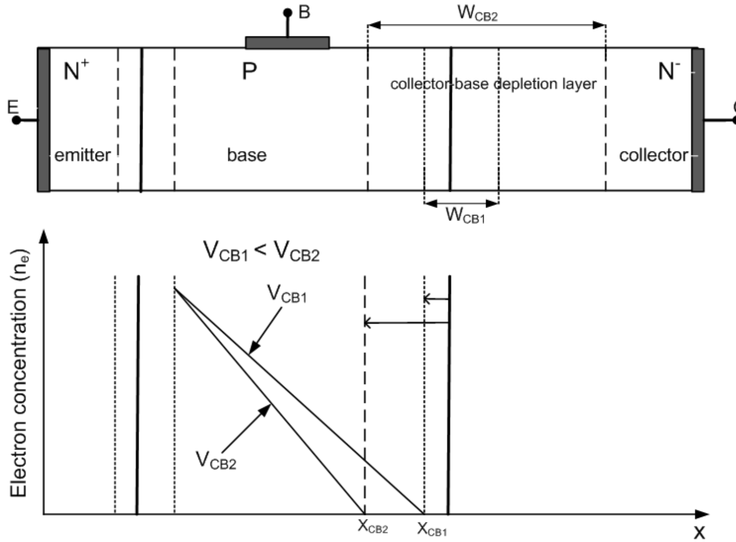


Figure 1.4: An illustration of the collector-base (forward) Early effect. We show the distribution of the minority carrier in the neutral base region for two different values of the collector-base reverse bias. When the collector-base reverse bias is increased from V_{CB1} to V_{CB2}, the neutral base width is decreased from X_{CB1} to X_{CB2}, which results into an increase of the electron gradient in the base; this leads to increase in the collector current.

junction, the lightly doped epilayer will be more depleted than the highly doped based. In such a case, the neutral base thickness and thus the collector current are less sensitive to V_{CB} variations, which implies that the Early effect is minimized. For such devices, the Early voltage V_A, which is defined as the inverse of the fractional change of collector current over V_{CB} variation (an indicator of the output conductance of the device) is usually large. The large V_A is translated into large output impedance, leading to higher voltage gain of amplifiers for a given load [1]. As discussed earlier, in present day high-speed (RF) SiGe devices, the doping concentration of the collector-epilayer is relatively high and somehow comparable to the base doping concentration. This is aimed at postponing Kirk effect [34] and reducing the collector-base junction transit time, which results into high f_T [29, 31, 33]; but this lowers V_A because of the increased base-width modulation. Therefore, in such RF-devices, there is always a tradeoff between minimizing the Early effect (i.e., increasing V_A) and maximizing other relevant characteristics e.g., f_T and breakdown voltage (BV_{CE0}), a process which generally depends on the targeted transistor application.

1.3 Semiconductor device modeling

The rapid developments in the semiconductor device technology over the past (few) decades have led to an increased interest in device modeling; which is a representation of the physical behavior of a device or network of devices (electrical circuit) by an abstract mathematical model which reliably approximates this behavior. Such a model for-

mulation may either be a closed-form expression (analytical model) or in general a system of coupled (differential) equations (to be solved numerically). The increased need to understand the detailed operation and to optimize the design of silicon very-large-scale integrated (VLSI) devices (i.e., combining many devices into a single chip) [10], very-high-speed integrated circuits (VHSIC), and compound semiconductor devices has meant that device modeling now plays a vital role in modern technology. As the scale (size) of the individual semiconductor devices decreases and the complexity of the physical structure increases, the nature of the device characteristics is shifting from those obtained from many of the classically held modeling concepts [57]. Furthermore, the challenges encountered in carrying out measurements on the devices means that greater emphasis must be put on results obtained from theoretical characteristics [58]. Semiconductor device modeling allows new device structures to be rigorously analyzed and investigated prior to fabrication, which is cost efficient. Depending on the targeted application, device modeling can either be used for *device simulation* or for *circuit simulation*; these are discussed next.

1.3.1 Device simulation

The internal operation of an electronic device e.g., a transistor, can be studied in great detail by solving the basic device (partial differential) equations subjected to relevant boundary conditions. These equations are derived from physical principles or laws, and they are used to gain detailed understanding of the structure and operation of the electronic device. The mathematical description of the device then consists of a representation of the device geometry, the material composition, doping profiles, the basic device physics equations and boundary conditions. Furthermore, a specification of the material properties, such as dependence of band-structure and carrier mobility on material composition can be part of the mathematical model. Carrier mobility also depends on the strength of the electric field, which can also be included in the model formulation. Temperature dependencies can also be taken into account. The resulting model equations are solved numerically (device simulation) for specified boundary conditions, using numerical methods such as finite element method [58], which enables the study of the device structure and operation.

Device simulation is part of the *technology computer-aided-design* (TCAD), a branch of electronic design automation (EDA) [1] that models semiconductor fabrication and device operation. Generally TCAD is aimed at designing of semiconductor processes and devices to fulfill some given specifications; the modeling of the *fabrication* is termed as *process* TCAD, while the modeling of the *device operation* is termed as *device* TCAD. Collectively process TCAD and device TCAD form the core tools for TCAD. In process TCAD, modeling of semiconductor-chip process-manufacturing steps such as lithography, deposition, etching, diffusion, and ion implantation are included [16, 24], while in device TCAD the modeling of electrical, thermal and optical behavior of semiconductor devices (e.g., transistors) based on fundamental physics such as doping profile of the device are

[1] EDA is a category of software tools for designing electronic systems such as printed-circuit boards and integrated circuits. The tools are used jointly in a design flow which is employed to design and analyze semiconductor chip or integrated circuit.

taken into account. Process TCAD requires modeling of the physical principles of fabricating, and at times also the modeling of the specific equipments used. Device TCAD mostly focuses on the physical principles at the basis of carrier transport and of optical generation in semiconductor devices.

The key goal of process TCAD is to obtain accurate prediction of the active dopant distribution in the various regions of the semiconductor device, the stress and the device geometry. The output of the process simulation (e.g., doping concentration) is used as inputs for the device TCAD, to simulate the electrical characteristics (e.g., the electric field distribution and the currents) of the device under different boundary conditions (e.g., biases) and temperatures. TCAD mostly involving solving a set of differential equations, which are derived from basic semiconductor device physics, and normally numerical approaches are used to solve such equations; here dedicated commercial software packages such as TSUPREM [59] for process simulation and MEDICI [36, 60] for device simulation are widely used. This approach of modeling the device behavior is generally complicated and demands tremendous computational time. Achieving accurate simulation results would demand more discrete nodes, which will result into large amount of computational time; thus, TCAD modeling is only used for studying the internal operation of devices, as well as for studying the relation between their structure and composition on the one hand, and their electrical characteristics on the other hand [17, 38]. It is not suitable however as a routinely applied way to model the electrical characteristics of devices, mostly those connected in a circuit network (integrated circuit (IC)). In such *circuit simulations*, *electronic computer-aided-design (ECAD)* (another branch of electronic design automation (EDA)), which is oriented towards electrical circuit design is widely used. In ECAD, simple and accurate *compact models* [25, 36, 46, 58] incorporated in circuit simulators such as SPICE are used to study the behavior and properties of electrical circuits. Such compact models used in circuit simulation will be discussed in the next section.

1.3.2 Compact semiconductor device modeling for IC simulation

For simulation of integrated circuits (ICs), which may contain thousands or even hundreds-of-thousands of electronic elements/devices (e.g., transistors, capacitors, resistors, inductors, and diodes), circuit designers prefer easy-to-simulate models (i.e., faster models in terms of CPU time) for the various employed circuit elements. Such models (*compact models*) of circuit elements minimizes the computational time, development costs and time-to-market, since they enable faster analysis/study of the behavior and properties of the electrical circuits under different relevant operation regimes (e.g., wide bias and temperature range), and thus enable circuit design adjustments in order to achieve the targeted performance. Generally the compact modeling process involve developments of models for integrated semiconductor devices for use in circuit simulations of the electrical characteristics. Compact models of the individual circuit elements are mathematical models that are sufficiently simple to be incorporated in circuit simulation and are sufficiently accurate to make the outcome of the circuit simulation software (i.e., circuit simulators) e.g., SPICE, SPECTRE and ADS, useful to circuit designers. The industry's dependence on accurate and time-efficient compact models continues

to grow as circuit operating frequencies increase and device tolerances scale down with associated increases in chip device count, and analog content in mixed-signal circuits. Compact modeling is a critical step in the design cycle of modern IC products. It is certainly the most important vehicle for information transfer from technology fabrication to circuit and product design.

Though compact models demand simplicity and accuracy, there is always a tradeoff between these two requirements; for this reason a hierarchy of models of different levels of complexity (or number of nodes in case of a equivalent circuit based model) could be offered to a circuit designer. Good convergence behavior of a compact model (under different operation regimes) is a key requirement for model implementation in circuit simulators, this demands for continuous (mathematical) equations describing the compact model, with continuous derivatives (of currents and charges) up to highest order (usually third order for bipolar transistors compact models is sufficient) [17]. For convergence it is required that compact models are globally well behaved; that is, they also need to be numerically well behaved outside their range of physical validity. This is important, because, beyond the junction breakdown the transistor my really be destroyed, but the compact model may not crash, if, e.g., during iterative search for the numerical solution of its equations, it is evaluated for bias conditions beyond BV_{CEO}.

We note that the accuracy of a compact model, does not depend only on the employed set of model equations, but also on the accuracy of the input device dependent model parameters [17, 25]; hence, great attention has to be paid on the employed methods for extracting the various model parameter values. Therefore, in order to make a compact model successful it would demand establishment of a well defined parameter extraction methodology and infrastructure which is greatly aided by properly formulated model equations and well-defined parameters. An example of such measurement based model parameter acquisition methods (for the base resistance and thermal resistance of bipolar transistors) will be presented in Chapter 2 of the thesis. We point out that in some instances the model parameters can be well known universal constants (i.e., device and technology independent), such parameters are taken as *model constants*, and they are kept fixed. An example of model constants used in e.g., Mextram compact bipolar transistor [31, 38] are A_n and B_n (see equation 1.13), the coefficients of the Chynoweths empirical law [61], whose values were experimentally determined [22, 27, 28].

In present day industrial (integrated circuit) applications, *physics-based compact models* are often preferred [36, 46, 58], particularly when concerned with geometrical or statistical or (virtual) predictive simulation. Since modern IC fabrication techniques provides transistors in a continuous range of sizes, compact models must therefore be geometrically scalable, i.e., describe the characteristics of the components for any given device size in a defined range. Electronic gadgets have to be sufficiently operational under a wide range of temperatures. Therefore, the dependence of electrical characteristics on temperature need to be modeled (*temperature scaling*). This also helps to take into account the increase of the device temperature due to self-heating (and mutual-heating) of the device. Devices actually delivered by IC fabrication techniques will show stochastic, or may be even systematic, deviations from the targeted reference devices [38]. Doping concentrations and geometrical layout may show slight variations, for example. This will result in stochastic variations in electrical characteristics, and a compact model should

be able to predict these variations. Such application demands for a compact model, can easily be realized when the model is formulated with a sound physical basis [25, 58]. For example, with a physical model, the geometrical and temperature scaling can be sufficiently achieved by applying scaling rules to the relevant physical model parameters. Physical models can also be used to predict the circuit behavior of future thought devices in the same category or study the performance of a fully virtual circuit design. These reasons and beyond, makes physics based compact models to play a central role in semiconductor industry. With new device technologies being developed for future targeted applications, compact models are undergoing continuous development in order to meet the increasing designers' demands (e.g., possibility to predict circuit performance for a new process as early as possible). In fact, the desired industrial trend is to co-develop process and circuits with the goal to have working instances of circuit designs already at the time of process qualification. This so-called concurrent engineering obviously contributes to a significant reduction in time-to-market. The corresponding cost reduction is achieved if compact models are capable of accurately predicting circuit performance early in the process development phase [37], and this calls for continuous cooperation between device physicists, modeling engineers, and circuit designers.

Other than the physical models, there are *empirical models* which are based on device behavior as observed from the measurements rather than from theoretical predictions. Analytical expressions aimed at achieving good fits with the measurements are formulated, and these are generally not based on device physics [25]. Such empirical models require a short development time with respect to physical device models, though they are limited in their applicability because of the nonlinear behavior of most devices with respect to signal level [58]. There are also *table based models*, which uses the measured device behavior directly in the circuit simulation. So instead of deriving the analytical formulas describing the nonlinear electrical performance, lookup tables are derived from the measurements, and simulations are performed by interpolating the stored values [58]. Such models can as well be developed in a shorter time, though they are only limited to the measured range, so no possibility for application outside the measured range [25]. These two later type of compact models are rarely used in circuit simulation due to their lack of scalability, thus, they will not be discussed further in the thesis.

The complexity of the compact model is dependent largely on the type of the circuit element, e.g., a linear resistor can be described simply by Ohm's law, i.e., $V = IR$ [32], but for a bipolar transistor, a more sophisticated compact model would be required. In most cases, active devices like bipolar transistors are usually represented by a small electrical circuit itself (i.e., an equivalent circuit of lumped-elements [38]), which then again is mathematically described by a set of algebraic equations, or, in case of time dependent analyses, ordinary differential equations [38]. For example, the present day standard industrial compact models for bipolar transistors i.e., Mextram [31] and HiCUM [37], are equivalent circuits themselves that consist of around 50 simpler elements, such as diodes, resistors, and capacitors [2]. We note that in this thesis, the Mextram compact model (summarized in the next section) will be used. When a device is modeled by an

[2] To establish conservation of charge in circuit simulations, it turns out that it is better to formulate the models in terms of charges, not capacitors, and hence good compact models contain functions that describe charges as function of bias conditions, not capacitors.

equivalent circuit, one will usually need a circuit simulator (equation solver) to study the behavior, properties, and electrical characteristics of that circuit. But that will be less demanding in terms computational costs than performing full device simulation (full physical modeling in space and time by partial differential equations) [38].

The different elements of the equivalent circuit, represents different physical components of the device, and physical mechanisms that influences the relevant terminal characteristics (e.g., voltages, currents, noise, etc) under all relevant practical circumstances (e.g., boundary conditions (such as biases and signals) and ambient temperature). Such physical mechanisms include: avalanche breakdown, parasitic resistances, self-heating, Early effect, charge distribution, quasi-saturation, and current crowding [24] among others; these must be sufficiently addressed in order to achieve reliable model simulations. These effects are represented componentwise in an equivalent circuit, and the various lumped elements representing the different transistor components are connected together in a way that represents (mimics) the internal structure of the device. The circuit elements are not necessarily derived from closed form analytical functions describing physics, but also may be by an empirical approach; the elements' value (as a function of e.g., bias and temperature) is calculated by computationally as inexpensive as possible equations formulated as closed-form analytical solutions, preferably in explicit form. In other words, numerical solutions of differential equations are usually prohibited, thus excluding numerical device models [37, 38, 41].

With (physics-based) element models, the time, bias, and frequency dependence of the electrical characteristics of the transistor can be addressed, thus such models can be considered as large-signal models [62]. Usually, the time-dependence of all relevant characteristics is done by representing all relevant bias-dependent charges in the device, these charges are indicated by capacitor symbols in an equivalent circuit e.g., in Mextram [31]. The time-dependence processes are then basically the processes of charging and discharging these capacitors in the network. These capacitors in advanced compact bipolar models' equivalent circuits are nonlinear network elements. This simply means that the physical charges that they represent depend on the biases in the network in a non-linear way, so that the associated capacitances would have a bias dependent value [38]. This would imply that, most of the elements of the equivalent circuit are nonlinear in this sense. Thus, by taking into account the full non-linear bias dependence of the various lumped elements, the compact model provides an appropriate basis to simulate the full, non-linear response of bipolar transistors to large signal. Then, by linearization, the small-signal AC analysis can be performed by the circuit simulator. Also from the large-signal compact model, at zero frequency, DC bias solutions can be evaluated.

1.3.3 Mextram compact model for Si BJTs and SiGe HBTs

The Mextram model is a physics-based world standard industrial compact model for simulation of both Si and SiGe bipolar transistors. This model has been developed for almost four decades [39] and over this period, it has been undergoing a number of continuous upgrades in order to capture all relevant terminal characteristics of any bipolar transistor, under different boundary conditions; the current level is Mextram 504.10 [39].

Compared to the previous levels, Mextram 504 provides better results in the description of first and higher order derivatives, which results into e.g., more accurate descriptions of the output conductances, junction charges, cutoff frequency, and low-frequency third order distortions. Mextram model generally supports accurate computer simulations of all relevant observable terminal characteristics (e.g., current and noise), as a function of applied terminal boundary conditions (e.g., junction biases) and ambient temperature. It is itself a circuit which comprises of lumped elements forming an equivalent circuit representing the physical structure of a bipolar transistor. Mextrams' equivalent circuit [31, 38] not only outlines the general mathematical structure of the compact model, but it also represents the essential physical topology of the transistor. The branches representing model currents and charges are schematically associated with different physical regions of a bipolar transistor separated by the base-emitter, base-collector, and substrate-collector 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 mathematical model formulation involves a system of algebraic equations for the time independent (DC) simulations, and a system of differential equations (nonlinear) for the time-dependent simulations. These equations are solved by the available circuit simulators e.g., SPICE and ADS for the specified stationary and time-dependent boundary conditions.

The Mextram model takes into account a number of physical mechanisms that affects the terminal characteristics (e.g, currents and depletion capacitances) of the bipolar transistor, under different bias and temperature conditions. Examples of such mechanisms include [39]: Early effects, current crowding and conductivity modulation of the base resistance, Zenner-tunneling in the base-emitter junction, high-injection effects, ohmic resistance of the collector-epilayer, quasi-saturation, Kirk effect, weak-avalanche in the collector-base junction, low-level non-ideal base currents, ohmic epilayer resistance, charge storage effects, substrate effects and parasitic PNP, high-frequency current crowing, excess phase-shift, bandgap grading effects in SiGe HBTs, temperature scaling, and self-heating [17, 24] among others. Mextram compact model does not contain extensive geometry of process scaling rules. This is because bipolar transistors come in different architectures, so that no general geometric scaling rules can be given. Therefore, geometric scaling rules of bipolar transistors have not been standardized. In practice it is added by the model users as a shell around Mextram. To enable development of such a shell, having a physical basis of the model helps: one can then find out how parameter values will scale with geometry. Some parts of the model (e.g., the increase of the avalanche current when the current density in the epilayer exceeds the doping level) are optional and can be switched on or off by setting flags; such is aimed at improving simulation speed and convergence. Though Mextram is by default formulated for NPN transistors, the model can be equally well used for PNP transistors by simple change of the current and charge polarity. Since Mextram is a physics-based model, it means that once the correct parameter set (for a selected device) has been found, the description of the device behavior is sufficiently realistic to allow accurate prediction of the model behavior under different bias and temperature conditions.

The physics-based Mextram model will be used later in the thesis, for the self-consistency check of the improved extraction method of thermal resistance R_{TH} and base

resistance R_B of bipolar transistors (in Chapter 2), which is based on the measured characteristics in the weak collector-base breakdown regime, as well as for verification (in Chapter 4) of the derived extended version of Mextram (in Chapter 3) which includes the non-local avalanche compact model. The Mextram model addresses the earlier stated physical mechanisms using a system of physics-based (coupled) mathematical formulations, which generally depend on junction voltages (external and internal), currents, and temperature. As an example, the *weak-avalanche mechanism* is addressed in Mextram as summarized next (as extracted from [31]). Note that, the weak avalanche mechanism will be presented not only as an example, but also because this aspect of Mextram, and compact modeling more in general, will be one of the key topics of the thesis.

Compact modeling of the weak collector-base avalanche mechanism in Mextram

According to the Mextram model, the total avalanche current is modeled as the integral of the product of impact-ionization coefficient (α) and the collector-epilayer current (I_{epi}), over all positions where the ionization takes place. This holds only in the *weak avalanche regime, where the generated carriers through impact-ionization do not generate extra charge carriers* [3]. Considering a weak avalanche field dependent ionization coefficient $\alpha[E(x)]$, the generated total (weak) avalanche current I_{avl} is proportional to the epilayer current I_{epi} passing through the epilayer [31], i.e.,

$$I_{\text{avl}} = I_{\text{epi}} \int_0^{W_{\text{avl}}} \alpha[E(x)]\, dx = I_{\text{epi}} G_{\text{EM}} \,, \tag{1.12}$$

where W_{avl} is the model parameter representing the effective width of the collector-epilayer, and G_{EM} is the generation factor computed from the integral of α. The electric field distribution is vital for accurate modeling of impact-ionization, and in the Mextram model [31], the electric field distribution ($E(x)$) schematically presented in Fig. 1.5 is used to define the spatial dependence of the ionization coefficient (which is based on Chynoweth's empirical law [61])

$$\alpha[E(x)] = A_n \exp\left(\frac{-B_n}{|E(x)|}\right) , \tag{1.13}$$

where A_n and B_n are material constants [27, 28]. We note that, in the local avalanche model, the ionization coefficient α is expressed as a function of the electric field $E(x)$, as in equation (1.13), but in reality (non-local avalanche model) α should be expressed as a function of the kinetic energy of the carriers or carrier temperature $T_e(x)$, as will be demonstrated in Chapter 3.

The magnitude of the electric field distribution is approximated by [27]

$$|E(x)| \approx \frac{E_{\text{max}}}{\left(1 + \frac{x}{\lambda_D}\right)} , \tag{1.14}$$

[3] We note that since avalanching is ignored, it is not actually accurate to call the whole process an avalanche, but the terminology is commonly used in compact models, thus we will keep it through out the thesis

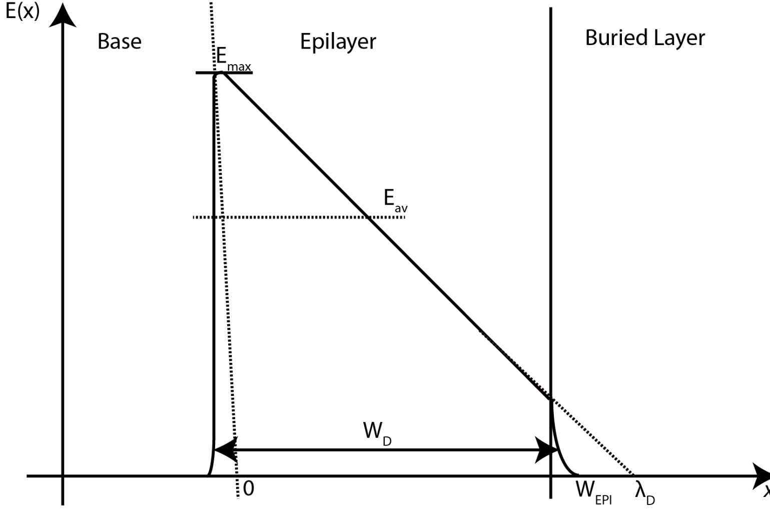


Figure 1.5: Schematic representation of the absolute value of the electric field used in the derivation of the local avalanche model in Mextram model [31]. In the figure, $x = 0$ represents the metallurgical base-collector junction.

where λ_D is the extrapolated width of the epilayer (see Fig. 1.5). On performing the integral in (1.12), using (1.13) and (1.14), we obtain the value of the generation factor G_{EM} as

$$G_{\mathrm{EM}} = \frac{A_n}{B_n}\lambda_D E_{\max}\left\{\exp\left[-\frac{B_n}{E_{\max}}\right] - \exp\left[-\frac{B_n}{E_{\max}}\left(1+\frac{W_{\mathrm{avl}}}{\lambda_D}\right)\right]\right\}. \tag{1.15}$$

In this expression, λ_D and $E_{\max}$ (which depend on collector bias and current) still need to be determined. From Fig. 1.5, the average value of the electric field E_{av} over the space-charge region is given by

$$E_{\mathrm{av}} = \frac{V_{d_C} - V_{BC}}{W_D}, \tag{1.16}$$

where W_D is the collector-base depletion width, V_{d_C} is the base-collector junction built-in voltage, and V_{BC} is the external base-collector bias without the voltage drop across the external collector resistance [63]. From this expression, we can see that at $V_{BC} = V_{d_C}$, the average electric field becomes zero; in such a case the base-collector junction is already far in forward bias and again the epilayer will be flooded with electrons and holes resulting in a low electric field. Therefore, avalanche current is taken to be zero when $V_{BC} > V_{d_C}$. Note that the expression below are such that also $E_{\max}$ will go to zero when E_{av} goes to zero. The expression (1.15) for the generation factor is such that at that point also G_{EM} will go to zero (including all its derivatives).

For charge carriers moving in the high-field collector-base depletion region (of a bipolar transistor in forward operation mode [17]), their drift velocity saturates (i.e., $v_d = v_{\mathrm{sat}}$) and the carrier electron density is then given by $n = |I_{\mathrm{epi}}|/q_e A_{\mathrm{em}} v_{\mathrm{sat}}$. The

net charge density across the depletion region is given by

$$\rho = q_e N_D = q_e N_{\rm epi} - \frac{|I_{\rm epi}|}{q_e A_{\rm em} v_{\rm sat}}.$$

Then from Poisson's equation described earlier in (1.4), the gradient of the electric field would become

$$\frac{dE}{dx} = \frac{qN_{epi}}{\varepsilon}\left(1 - \frac{I_{\rm epi}}{I_{hc}}\right), \tag{1.17}$$

where $N_{\rm epi}$ is the epilayer doping concentration and $I_{\rm hc} = q_e N_{\rm epi} A_{\rm em} v_{\rm sat}$ is known as the hot-carrier epilayer current. By considering zero epilayer current, (1.17) can be rewritten as

$$\left.\frac{dE}{dx}\right|_{I_{\rm epi}=0} = \frac{qN_{\rm epi}}{\varepsilon} = \frac{2V_{\rm avl}}{W_{\rm avl}^2}, \tag{1.18}$$

where $V_{\rm avl}$ is a dedicated Mextram model parameter which represents the measure for the derivative of the electric field, especially around the maximum electric field. The electric field E_0 at the collector-base junction can now be calculated from Fig. 1.5 as

$$E_0 = E_{\rm av} + \frac{V_{\rm avl}}{W_{\rm avl}^2}\left(1 - \frac{I_{\rm epi}}{I_{\rm hc}}\right). \tag{1.19}$$

In normal transistor operating regimes, the maximum of the electric field will be at the base-collector junction and therefore take $E_{\rm max} = E_0$. In the high current limit, the electric field will be modulated by the associated mobile charge density (Kirk effect [34]) and may possibly be subject to two or three dimensional current spreading [29, 39] effects. In this context, avalanche is likely to be less distinguishable [37]; detailed assessment of this regime is outside the scope of the thesis.

To calculate the extrapolated width of the epilayer, λ_D, relation (1.14) is used, from which it follows

$$\left|\frac{dE(x)}{dx}\right| = \frac{E_{\rm max}}{\lambda_D} = \frac{2V_{\rm avl}}{W_{\rm avl}^2}\left(1 - \frac{I_{\rm epi}}{I_{\rm hc}}\right). \tag{1.20}$$

The electric field can also be written as

$$|E(x)| = E_0 - \frac{2x}{W_D}(E_0 - E_{\rm av}), \tag{1.21}$$

which is given in such a way that the electric field at $x = W_D/2$ equals the average electric field $|E(x = W_D/2)| = E_{\rm av}$ [31]. In the case we presented here, it is valid to consider $E_0 = E_{\rm max}$. From the expression for the electric field, and from $|dE/dx| = E_{\rm max}/\lambda_D$, the width λ_D is found as

$$\lambda_D = \frac{E_{\rm max} W_D}{2(E_{\rm max} - E_{\rm av})}. \tag{1.22}$$

The collector-base depletion width W_D also needs to be calculated. Using a very simple abrupt junction depletion model, we obtain

$$W_D = \sqrt{\frac{W_{\rm avl}^2}{V_{\rm avl}}}\sqrt{\frac{V_{d_C} - V_{BC}}{1 - I_{\rm epi}/I_{\rm hc}}}. \tag{1.23}$$

We note that this formula can lead to the depletion layers larger than the (effective) epilayer width W_{epi} (here taken to be equal to model parameter W_{avl}) for $I_{epi} \geq I_{hc}$. To prevent this to occur, an empirical expression is used for the thickness over which the electric field is important. The limiting cases for this avalanche model can be found in e.g., [29, 31, 39], together with the extension of the model to cases where the effective epilayer width decreases due to base-widening and change of the sign of the slope for the electric field (see equation (1.17)) due to Kirk-effect.

Temperature dependence of I_{avl}

The temperature dependence of the avalanche current I_{avl} is based on relation 1.15. Following the earlier work [28], it was concluded that the temperature dependence of the ratio of the coefficients A_n and B_n in 1.15 roughly cancels each other and can be neglected in comparison to the exponential variation. Then the temperature dependence of B_n in the exponential term was more apparent. Following their experimental analysis on both high voltage and high frequency IC processes, it became clear that the avalanche parameter B_n increased with temperature; for all the investigated processes B_n can be expanded in a Taylor series around $T = 300$ K as

$$B_n(T) = B_n\left(1 + \gamma_1 \Delta T + \gamma_2 \Delta T^2\right), \tag{1.24}$$

where $\Delta T = (T - 300)$ K, with T the device temperature, and γ_1 and γ_2 are coefficients of expansion.

This temperature scaling rule was extensively experimentally tested and was found to be adequate [28] against various semiconductor bipolar platforms. At the time when the scaling rule (1.24) for Chynoweth's empirical law, i.e., (1.13) was established, the results were considered to be universally valid for silicon. Based on that notion, the coefficients of (1.24) were introduced in Mextram as model constants, not model parameters. The value for these coefficients (i.e., $\gamma_1 = 7.2 \times 10^{-4}$ and $\gamma_2 = -1.6 \times 10^{-6}$) were obtained by fitting from experimental data (see Fig. 1.6) taken on various bipolar IC processes [28], but once and for all. This temperature scaling rule (i.e., equation (1.24)) has been used in the standard Mextram model for decades and it looked sufficient to address the temperature dependence of the weak avalanche characteristics. But during our investigation using measured data (over temperature) taken on present day standard industrial Si and SiGe bipolar transistors, we found out that this is no longer the case, as will be discussed further in the next section.

1.4 Motivation for weak collector-base breakdown modeling

The local avalanche model of the Mextram compact model discussed in the previous section uses the temperature scaling rule presented in equation (1.24) to describe the temperature dependence of the weak avalanche current I_{avl}. This temperature scaling rule was developed in [28] and by the time of its development, was found adequate to model the temperature dependence of I_{avl} for wide range of bipolar transistors. This same

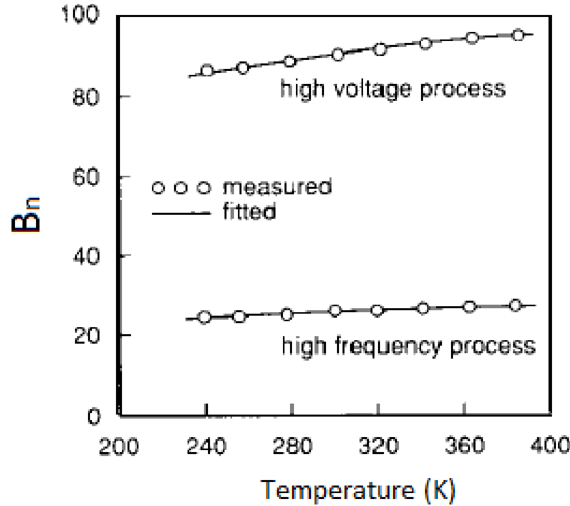


Figure 1.6: The avalanche parameter B_n verses temperature [28], with sysmbols representing the measured data and solid curves the fits using equation (1.24). The fitted curves give the values for the coefficients as $\gamma_1 = 7.2\times10^{-4}$ and $\gamma_2 = -1.6\times10^{-6}$.

model exists in the present version 504.10 of the industrial standard Mextram model and was used to simulate the characteristics in the weak avalanche regime of a selected standard industrial (high-speed (NPN)) SiGe-HBT device; the corresponding best fitting results achieved with the Mextram model were presented earlier in Fig. 1.3, by the solid curves. From these results, deviations between the measured family of characteristics (symbols) and Mextram simulated characteristics (solid curves) can be observed, as a function of both bias and ambient temperature (i.e., $T = 25, 50, 75, 100$, and $125\,^\circ$C). Such deviations are in spite of the fact that the appropriate temperature scaling rules of the relevant Mextram model parameters (e.g., equation (1.24)) were carefully and fully applied. During the time the Mextram model (discussed in the previous section) was developed, the model simulations were supposed to yield good fits to the measured data, without any need for refitting the coefficients of equation (1.24). But this is not the case in Fig. 1.3, which suggests that the physics behind the characteristics in the weak avalanche regime is now significantly different to what it was at the time the previous discussed version of Mextram model was developed. This calls for further studies to elucidate the physics behind these observed discrepancies in Fig. 1.3, and this will be one of the subjects to be addressed in the thesis.

Since a proper fit over all bias and temperature conditions is, as demonstrated in Fig. 1.3, actually not achievable, two different fit strategies have been applied, each with its own purpose. Here, the standard industrial Mextram model [31] was optimized at *two* chosen target temperatures, i.e., at the lowest ambient temperature $T_{\text{ref}} = 25\,^\circ$C, Fig. 1.3(a) and at the highest ambient temperature $T_{\text{ref}} = 125\,^\circ$C, Fig. 1.3(b). Figure 1.3(a) is intended to show what typically a person extracting parameters for the industry would

end up with, i.e., Mextram model simulations that overestimate the base current (with respect to the measured data) for $T > 25\,^{\circ}\mathrm{C}$. On the other hand, Fig. 1.3(b) is intended to more clearly arrive at a qualitative hint for the possible underlying device physics, i.e., the time dependence of the energy relaxation length (λ_e).

From these results, the shortcomings of the standard industrial compact model MEXTRAM in reproducing the measured characteristics in the weak collector-base breakdown regime of present day Si and SiGe industrial bipolar transistors as a function of bias and temperature are observable. These setbacks strongly limit application of Mextram model in simulating the behavior of transistors in the industrial domain; for example, the model underestimates device junction breakdown voltage BV_{CEO}, which misrepresents the true transistor operation capabilities. Such is a great disadvantage mostly in present day high speed (RF) devices (applied in e.g., power amplifiers), with very low BV_{CEO} (generally below 1.0 V), which has forced integrated circuit (IC) designers to exploit transistor operations above BV_{CEO} limits. The employed compact model should be able to produced reliable values for the breakdown voltage, so that the circuit designer can be able to make a correct and reliable judgement based on the targeted application. Due the increasing tightening application determined requirements (such as tolerances with respect to extreme environmental conditions e.g., in space and automotive applications), as well as self-heating [54, 64] of highly scaled RF-devices, accurate modeling of the temperature dependence with respect to device characteristics has gained undiminished attention. This means that the available standard compact transistor models e.g., Mextram [28, 31] have to accurately reproduce the measured characteristics in the various device application regimes such as weak collector-base breakdown regime (i.e., weak avalanche regime), subjected to different operation conditions.

The current Mextram compact model, addresses a number of physical effects (see Section 1.3.3) that may be significant in the collector-base breakdown regime of present day bipolar transistors [17, 31, 38]. Examples of such effects include: Early effect, local weak avalanche [27], self-heating [54], quasi-saturation, Kirk effect [34], bandgap engineering in SiGe HBTs [29], and neutral base (Auger) recombination [24]. Also from a literature review on the subject e.g., [16, 33, 48, 49], the non-local avalanche effect has been reported to be significant in this breakdown regime of mostly high scaled RF-devices. Device self-heating was also reported [52, 54] to be relevant in such RF-devices. Though all these physical mechanisms have been studied in literature, and some already included in the available standard compact models for bipolar transistors, still deviations are observed between the measured and simulated family of characteristics in the weak collector-base breakdown regime. This imposes various questions as to why this is the case: are these physical mechanisms not sufficient to tackle such a problem or are there still other physical mechanisms that need to be taken into account? Such yet unanswered hypothetical questions and may be more, will be a key subject of this thesis. Something to note is that we would like to keep the temperature scaling rule for Chynoweth's law as it was established in [28] i.e., equation (1.24). In this we would like to avoid trying to achieve the good model fits by simply adding more fitting freedom (e.g., by making γ_1 and γ_2 model parameters), but we aim at implementing the device physics as it is known from literature. We need to first understand the physics behind the observed deviations of the characteristics in the weak avalanche regime as applied to

present day bipolar devices, and then develop a physic based model that can sufficiently resolve such fitting issues subjected to different bias and temperature conditions.

1.5 Thesis aim/purpose

As illustrated in Fig. 1.3 the standard Mextram compact model as a whole is not capable of addressing the deviations between the measured family of weak collector-base breakdown characteristics and their simulated counterparts, as a function of both bias and temperature. This motivates the study of possible relevant and significant physical mechanisms in present day Si-BJTs and SiGe-HBTs that are responsible for such misfits. Therefore, the *overall/ key goal* of the thesis is to understand the (relevant) device physics behind the discrepancies demonstrated in Fig. 1.3. This naturally requires to develop a physics based compact model that serves to extend the Mextram model and resolves the problem of fitting the measured characteristics over bias and temperature.

With the above stated key goal, one may trust that self-heating will play a significant role, most especially that we want to study the characteristics as a function of temperature, and that the characteristics under study involves relatively high collector-base voltages, so significant power dissipation can be anticipated and thus self-heating effect. This suggests that thermal resistance will need to be determined, preferably in the context of weak breakdown characteristics. A series of publications have appeared recently, about a method for extraction of both thermal resistance R_{TH} and base resistance R_B, that actually uses the weak breakdown characteristics instrumentally. A review of this branch of literature will be presented in Chapter 2, and as well as an improvement of the method with respect to the accuracy of the extraction of R_{TH}. Hence, Chapter 2 serves the overall goal of the thesis: to develop understanding of the relevant device physics behind the breakdown characteristics. Indeed, it will become more clear later in the thesis that the physics behind the curves is *not* merely (avalanche) breakdown mechanisms. At the same time, it is instrumental for further experimental analysis of the measured data as presented in Chapter 4; indeed, a reliable and accurate value for R_{TH} is essential to that. We note that for the thesis, verification of the extraction of R_B can be considered as a supplementary goal. In general aspect however, extraction of R_B is important in itself, and the problem of reliably extracting R_B is notorious indeed as exemplified in e.g., [65].

In summary, we have three goals of the thesis, which include one overall goal and two derived (sub) goals:

1. Overall goal: understanding the physics behind the observed discrepancies as illustrated in Fig. 1.3.
2. Derived (sub) goals:
   (a) extending the Mextram model so as to resolve the demonstrated fitting issues of the characteristics in the weak collector-base breakdown regime, over bias and temperature.
   (b) improving relevant parameter extraction methods for R_{TH} (R_B extraction verified as a bonus).

As an additional bonus, in Chapter 4 we present the measured values for the energy relaxation length of electrons and its temperature coefficient; the values found are in accordance with earlier published values, but the values here have been measured in situ, in actual modern industrial bipolar transistors.

1.6 Thesis outline

The content of this thesis aimed at an extensive study of all possible physical mechanisms that significantly interfere in the weak collector-base breakdown regime of present day industrial Si BJTs and SiGe HBTs, and a development of an extended version of the Mextram model to achieve good model fits it is organized as follows:

Chapter 2 is dedicated to the derivation of an improved extraction method for thermal resistance (R_{TH}) and base resistance (R_B) of Si/SiGe bipolar transistors from the measurements of the transistor characteristics in the *weak collector-base breakdown regime*. This method consistently accounts for the influence of self-heating and Early effect on the internal base-emitter voltages, so as to arrive at an accurate extraction of R_{TH} and R_B. The method is demonstrated on measurement data taken on present day RF SiGe industrial bipolar transistors; which also works as an example and assessment of how self-heating significantly interferes with transistor characteristics in the weak breakdown regime. A comparison of the method with state-of-the-art is also carried out for both R_{TH} and R_B. The accuracy of the extraction method is assessed by carrying out a self-consistency check using the Mextram model simulation results as the input data for the method.

Chapter 3 contains an extended study of the relevant and significant physical mechanisms that are responsible for the observed deviations between the measured family of weak breakdown characteristics and their simulated counterparts, as a function of both bias and temperature. Such analysis will involve: (a) an extensive review of the physical mechanisms addressed by Mextram model and (b) a development a physics-based compact model formulation of non-local avalanche effects in bipolar transistors and their temperature dependencies; this model will be added to the standard Mextram model, thus leading to an extended version of Mextram. The aim of this extended version of Mextram is to simulate accurately the weak collector-base breakdown characteristics, thus resolve the fitting issue.

Chapter 4 is devoted to an experimental assessment of the family of weak breakdown characteristics by employing the physical extended version of the Mextram model developed in Chapter 3. We will demonstrate that these observed breakdown characteristics are actually not just classical local/non-local avalanche characteristics, but other physical mechanisms are significant as well in this breakdown regime. We will show that these effects can actually be distinguished, and thus taken into account independently of avalanche effects. With these interfering physical mechanisms well addressed, we will show that avalanche mechanism responsible for the deviations between the measured and simulated breakdown characteristics is non-local avalanche indeed, and with non-local effects taken into account, good model fits to the measured characteristics taken on the present day industrial Si/SiGe bipolar transistors are sufficiently achieved, over bias and temperature. We will also demonstrate that the developed model actually captures

the presupposed underlying semiconductor device physics, thus yielding values for the relaxation length of electrons and its temperature coefficient that are in accordance with earlier published literature values.

Chapter 5 summarizes the main conclusions from the work in the various chapters of the thesis, and presents several recommendations for potential future work.

References

[1] M. F. Chang, ed., *High-speed Devices and Circuits with THz Applications* (Taylor and Francis Group, 2015).

[2] Z. Ma, N. Jiang, G. Wang, and S. A. Alterovitz, *An* 18-*GHz* 300-*mW SiGe Power HBT,* IEEE Trans. Electron Devices .

[3] P. M. Mans *et al., SiGe-HBTs optimization for wireless power amplifier applications,* Hindawi Publishing Corporations .

[4] J.-S. Rieh *et al., Performance and design considerations for high speed SiGe HBTs of* $f_T/f_{max} = 375GHz/210GHz,$ (IEEE, 2003) pp. 374–377.

[5] F. Dowla, ed., *Handbook of RF and Wirelss Technology* (Elsevier, 2004).

[6] A. IIderem *et al., The emergence of SiGe:C HBT technology for RF applications,* in *Proc. International Conference on Compound semiconductors Mfg* (2006).

[7] A. Schüppen, *SiGe-HBTs for mobile communication,* Solid State electronics .

[8] J. S. Rieh *et al., SiGe-HBTs for millimeter-wave application with simultaneously optimized f_t and f_{max} of* 300 *GHz,* (IEEE. RF Integrated Curcuits Symposium, 2004).

[9] R. Gilmore and L. Besser, eds., *Practical RF Circuit Design for modern wireless systems* (Artech House, Inc, 2003).

[10] J. D. Cressler and G. Niu, eds., *Silicon Germanium Heterojunction Bipolar Transistors* (Artech House, Inc, 2002).

[11] J. L. Patton, D. L. Harame, J. M. C. Stork, B. S. Meyerson, G. J. Scilla, and E. Ganin, *SiGe-base, poly-emitter heterojunction bipolar transistors,* in *Proc. Symp. VLSI Technology* (1989) pp. 35–36.

[12] D. L. Harame *et al., The revolution in SiGe: impact on device electronics,* Elsevier B. V. , 9 (2003).

[13] H. Kroemer, *Theory of a wide-gap emitter for transistors,* in *in Proc. IRE* (1957) pp. 1535–1537.

[14] G. L. Patton *et al., Silicon germanium base heterojunction bipolar transistors by molecular beam epitaxy,* IEEE Electron Device. Lett .

1

[15] *Special issue on heterostructure transistors,* IEEE Trans. Electron Devices **36** (1989).

[16] P. Ashburn, *SiGe Heterojunction Bipolar Transistors* (John Wiley and Sons Ltd, 2003).

[17] M. Reisch, *High-Frequency Bipolar Transistors* (Springer-Verlag, 2003).

[18] B. S. Meyerson, *Silicon based semiconductor heterostructures: column IV bandgap engineering,* in *Proc. IEEE* (1992) pp. 571–587.

[19] B. S. Meyerson, *UVD/CVD growth of Si and Si:Ge alloys: chemistry, physics and device applications,* in *Proc. IEEE* (1992) pp. 1592–1608.

[20] R. D. Dupuis, *Preparation and properties of* $Ga_{1-x}Al_xAs$*-GaAs heterojunctions grown by metal-organic chemical vapour deposition,* Gallium Arsenide and Related Compounds 1978, Inst. Phys. Conf. Series .

[21] K. W. Wang, P. M. Asbeck, M. F. Chang, G. J. Sullivan, and D. L. Miller, *High speed circuits for lithtwave communication systems implemented with AlGaAs/GaAs heterojunction bipolar transistors,* in *BCTM. Digest* (1987) p. 142.

[22] P. A. H. Hart, ed., *Bipolar and Bipolar-MOS Integration* (Elsevier, 1994).

[23] M. Racanelli and P. Kempf, *SiGe BiCMOS technology for RF circuit applications,* IEEE Trans. Electron Devices .

[24] S. M. Sze and K. N. Kwok, *Physics of Semiconductor Devices*, 3rd ed. (Wiley, 2006).

[25] H. C. de Graaff and F. M. Klaassen, *Compact transistor modelling for circuit design* (Springer-Verlag, 1990).

[26] U. R. Pfeiffer and D. Goren, *An* 20*dBm fully-integrated* 60 *GHz SiGe power amplifier with automatic level control,* IEEE Journal of Solid-State Circuits .

[27] W. J. Kloosterman, J. C. J. Paasschens, and R. J. Havens, *A comprehensive bipolar avalanche multiplication compact model for circuit simulation,* in *Proc. BCTM* (2000) pp. 172–175.

[28] W. J. Kloosterman and H. C. de Graaff, *Avalanche multiplication in a compact bipolar transistor model for circuit simulation,* IEEE Trans. Electron Devices **36**, 1376 (1989).

[29] J. C. Paasschens, W. J. Kloosterman, R. J. Havens, and H. C. de Graaf, *Improved compact modeling of output conductance and cut-off frequency of bipolar transistor,* in *Proc. BCTM* (2000) pp. 62–65.

[30] S. M. Sze and G. Gibbons, *Avalanche breakdown voltages of abrupt linearly graded pn junctions in Ge, Si, GaAs, and GaP,* App. Phys. Lett .

[31] J. C. J. Paasschens, W. J. Kloosterman, and R. van der Toorn, *Model Derivation of Mextram 504, The physics behind the model,* Tech. Rep. Technical note TN-2004/01010 (Philips Research Eindhoven, 2005).

[32] S. M. Sze, *Physics of Semiconductor Devices*, 2nd ed. (John Wiley and Sons, INC, 1981).

[33] J. A. Yuan, *SiGe, GaAs, and InP Heterojunction Bipolar Transistors,* edited by K. Chang (John Wiley and Sons, Inc, 1999).

[34] C. T. Kirk, *A theory of transistor cut-off frequency (f_t) fall-off at high current densities,* IEEE Trans. Electron Devices .

[35] J.Lutz, H. Schlangenotto, U. Scheuermann, and R. D. Doncker, eds., *Semiconductor Power Devices: physics, characteristics, reliability* (Springer-Verlag, 2011).

[36] J. S. Yuan and J. J. Liou, *Semiconductor Device Physics and Simulation* (Plenum Press, New York, 1998).

[37] M. Schröter and A. Chakravorty, *Compact Hierarchical Bipolar Transistor Modeling with HICUM* (World Scientific Publishing Co. Pte. Ltd, 2010).

[38] R. van der Toorn, J. C. J. Paasschens, W. J. Kloosterman, and H. C. de Graaff, *Compact Modeling: Principles, techniques, and applications,* (Springer-Verlag, 2010) Chap. 7.

[39] R. van der Toorn, J. C. J. Paasschens, and W. J. Kloosterman, *The Mextram Bipolar Transistor Model: level* 504.10.1, Mextram definition document, (Delft University of Technology, 2012) http://www.mextram.ewi.tudelft.nl.

[40] J. C. J. Paasschens, W. J. Kloosterman, and R. J. Havens, *Parameter extraction for the Bipolar Transistor Model Mextram Level 504*, Unclassified report NL- UR 2001/801 (Philips Research Eindhoven, 2001).

[41] V. Milovanović, *Advanced Breakdown Modeling for Solid-State Circuit Design*, Phd. thesis, Delft University of Technology (2010).

[42] J. W. Slotboom, G. Streutker, M. J. v. Dort, P. H. Woerlee, A. Pruijmboom, and D. J. Gravesteijn, *Non-local impact ionization in silicon devices,* in *Proc. IEDM* (1991) pp. 127–130.

[43] H. C. de Graaf, W. J. Kloosterman, J. M. Geelen, and M. A. M. Koolen, *Compact bipolar transistor model for CACD with accurate description of collector behavior purposes,* in *Proc. BCTM*.

[44] H. C. de Graaf, W. J. Kloosterman, and T. N. Jansen, *Compact bipolar transistor model for CACD with accurate description of collector behavior purposes,* in *Proc. 18th Conf. Solid St. Dev. and Material, Tokyo, Japan.*

[45] W. Shockley, *Electrons and Holes in Semiconductors:* with applications to transistor electronics (D. Van Nostrand Company, Inc, 1950).

[46] M. Shur, *Introduction to Electronic Devices,* edited by S. M. Elliot (John Wiley and Sons, Inc, 1996).

[47] W. J. Kloosterman and H. C. de Graaff, *Avalanche multiplication in a compact bipolar transistor model for circuit simulation,* in *Proc. BCTM* (1988) pp. 103–106.

[48] G. B. Hong and J. G. Fossum, *Implementation of nonlocal model for implact-ionization current in bipolar circuit simulation and application to SiGe HBT design optimization,* IEEE Trans. Electron Devices **42**, 1166 (1995).

[49] M. M. Mahim and J. S. Yuan, *Modelling of the avalanche current including the non-local effect,* Int Journal Electronics **85**, 400 (1998).

[50] M. R. van den Berg, L. K. Nanver, and J. W. Slotboom, *Temperature dependence of avalanche multiplication in spiked electric fields,* in *Proc. IEDM* (2000) pp. 79–82.

[51] M. Lundstrom, *Fundamentals of Carrier Transport* (Cambridge University Press, 2000).

[52] N. Rinaldi and V. d'Alessandro, *Theory of electrothermal behavior of bipolar transistors: Part III- Impact-ionization,* IEEE Trans. Electron Devices **53**, 1683 (2006).

[53] L. L. Spina, V. d'Alessandro, S. Russo, N. Rinaldi, and L. K. Nanver, *Influence of concurrent electrothermal and avalanche effects on the safe operating area of multifinger bipolar transistors,* IEEE Trans. Electron Devices **56**, 483 (2009).

[54] R. Setekera, R. van der Toorn, and W. Kloosterman, *Local extraction of base and thermal resistance of bipolar transistors,* in *Proc. Bipolar Circuits and Technology Meeting* (IEEE, 2013) pp. 21–24.

[55] P. Ashburn, *Design and Realization of Bipolar Transistors*, edited by D. V. Morgan and H. R. Grubin (John Wiley and Sons, Inc, 1988).

[56] R. S. Muller, T. I. Kamins, and M. Chan, *Device Electronics for Integrated Circuits: Third Edition* (John Wiley and Sons Inc, 2003).

[57] C. M. Snowden, *Semiconductor Device modelling,* Tech. Rep. (Solid State Materials, Devices and Applications Group, Depertment of Electrical and Electronic Engineering, University of Leeds, 1985).

[58] M. Rudolph, *Introduction to Modeling HBTs* (Artech House, Inc, 2006).

[59] *TSUPREM*-4*: Two Dimensional Process Simulation Program, version* 2000.2, Avant Corporation, TCAD Business Unit (2000).

[60] *MEDICI: two dimensional device simulation program,* Tech. Rep. ed.2002.02 (Synopsys Inc, 2002).

[61] A. G. Chynoweth, *Ionization rates for electrons and holes in silicon,* Phys. Rev **109**, 1537 (1958).

[62] I. E. Getreu, *Modeling the Bipolar Transistor:* CAD of Electronic Circuits, vol.1 (Elsevier Scientific Publishing Company, 1978).

[63] R. van der Toorn, J. J. Dohmen, and O. Hubert, *Distributon of the collector resistance of planar bipolar transistor: Impact on small signal characteristics and compact modeling,* in *Proc. BCTM* (2007) pp. 184–187.

[64] T. Vanhoucke, H. M. J. Boots, and W. D. V. Noort, *Revised method for extraction of the thermal resistance applied to bulk and SOI SiGe HBTs,* IEEE Electron. Device. Lett Electron Devices **25**, 150 (2004).

[65] J. Steigerwald and P. Humphries, *TCAD assisted reflection on parameter extraction for compact modeling,* in *Proc. Bipolar Circuits and Technology Meeting* (2010) pp. 245–252.

2

Simultaneous local extraction of the base and thermal resistance of bipolar transistors

We present a method for simultaneously extracting the base and thermal resistance of advanced bipolar junction transistors. This method extends the category of parameter extraction methods that utilizes the weak avalanche current to vary the base current independently of the emitter current, while eliminating the voltage fluctuations across the emitter resistance by enforcing a fixed emitter current. This method consistently accounts for self-heating and Early effect so as to arrive at an accurate extraction of both base and thermal resistance. It involves only small variations in bias and temperature conditions, so that parameter extraction as a function of bias and temperature is enabled. The method is directly applicable to transistors – no dedicated test structures are needed – so that the method is cost effective. The method is demonstrated on modern RF SiGe HBTs, and the extracted results for the base and thermal resistance are compared with those determined from other independent DC- and AC-methods. In addition, the extracted results of the base resistance are compared with those determined from noise measurements. Our method is tested for self-consistency by using standard compact model simulations as input data for the extraction method.

This chapter is based on the published articles:
R. Setekera, L. F. Tiemeijer and R. van der Toorn, *WASET-IJECEECE*. 08 (2014) pp. 1456-1460.
R. Setekera, L. F. Tiemeijer, W. Kloosterman and R. van der Toorn, in *Proc. IEEE BCTM* (2014) pp. 215-218.
R. Setekera, R. van der Toorn and W. Kloosterman, in *Proc. IEEE BCTM* (2013) pp. 21-24.

2.1 Introduction

Both for characterization in a general sense, and for compact modeling purposes more specifically, the extraction of the base resistance R_B, and thermal resistance R_{TH}, of bipolar transistors is crucial (since both may degrade the device performance [1–3]). However, their extraction is far from trivial, even though a number of extraction techniques have been proposed in various publications. For example, a range of methods for base resistance extraction has been published [4–10], but the task is still recognized by experts as a challenge [11, 12].
A number of extraction techniques for R_{TH} have been published e.g., [13–22], but generally they are inconsistent. This inconsistence makes it difficult to choose the right method to use mostly in addressing internal self-heating of the highly scaled modern bipolar transistors. For example in [3], a commonly employed method [23, 24] for extracting R_{TH} was presented. This method is based on the good sensitivity of the base-emitter voltage V_{BE} to the temperature. Though this method is widely used in literature, it is inconsistent since it ignores the change in dissipated power during the measurements of V_{BE} as function of ambient temperature, for a fixed emitter current and constant collector-base voltage [25]. This in the end leads to an over estimate of the actual device temperature and the extracted device thermal resistance [25]. The impact of self-heating is of great concern in the current power transistors and high-speed heterojunction bipolar transistors (HBTs), which are fabricated in modern architectures such as those with deep trenches and buried oxides. Therefore, accurate extraction of thermal resistance has recently gained great attention [15, 25–27].

Published methods for extraction of base resistance (R_B) from *AC*-data have been shown [28, 29] to either *under*- or *over*-estimate the base resistance, depending on the topology of the small signal equivalent circuit on which the ac-method is based. DC-methods exploring the normal forward bias regime for extraction of R_B face the almost fundamental difficulty posed by the fact that with respect to the very current that actually traverses the base resistance, the emitter resistance is in series with the base resistance. This poses the problem of how to distinguish the two resistances given the observable I-V characteristics. This problem was first addressed by Verzellesi et al. [30], who utilized the collector-base avalanche current to vary the base and collector currents independently of the emitter current, while eliminating voltage fluctuations across the emitter resistance by enforcing a fixed emitter current. A refinement of this method [30] was carried out by Vanhoucke and Hurkx [27], who took effects of self-heating into account and as a by product offered an estimate for the thermal resistance.

We developed further an improved method in this category [31] that will be presented in this chapter. Compared to the method of Vanhoucke and Hurkx [27], this method includes a more sophisticated correction for the Early effect, thus enabling an accurate, unbiased extraction of the thermal resistance R_{TH}. A detailed derivation of this extraction method for R_B and R_{TH} will be presented in Section 2.2, where we will take the influence of both self-heating and Early effect into account. Our method will be demonstrated in Section 2.3 on the experimental data taken on advanced RF SiGe HBTs. Since the assumptions employed in the method derivation have observable consequences, we will present various methods that can be used to determine the right region for selection

of the method input data, in order to achieve accurate extracted values of R_B and R_{TH}. The extracted results for R_B and R_{TH} using our method will be compared in Section 2.4 to corresponding results from other independent DC- and RF-methods. From this comparison, we will show that our method yields reduced systematic error of the extracted values, especially the thermal resistance. Results for extracted base resistance using our method are also compared with those determined from noise measurements and the results show that the two methods yield results that are comparable within an acceptable error margin. By employing the simulation output data from the World Standard Compact Model MEXTRAM [2] as the input data for the extraction method, in Section 2.5 we carry out a self-consistency check of our extraction method over a wide range of bias and parameter values. The summary of the key achievements from this work will be presented Section 2.6.

2.2 Derivation of the R_B and R_{TH} extraction method

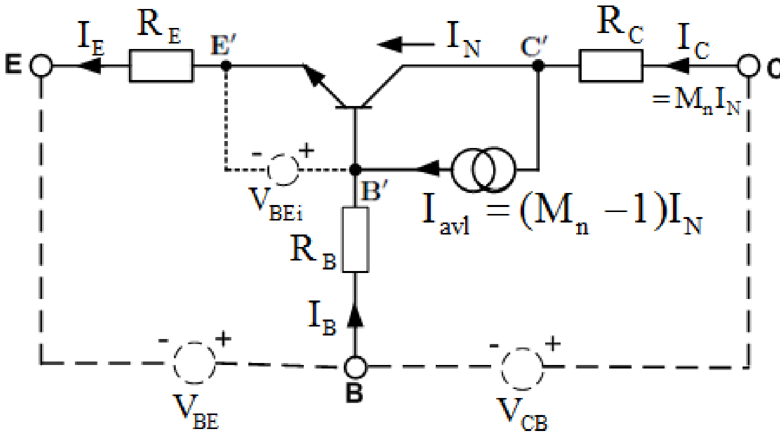


Figure 2.1: Schematic representation of the vertical bipolar junction transistor used for derivation of the extraction method of R_B and R_{TH}. Due to impact ionization in the collector-base space charge region of an NPN-transistor, the avalanche current $I_{\text{avl}} = (M_n - 1)I_N$ is generated and it leads to the reversal of the base current I_B. Here I_N is the main transistor current without avalanche and M_n is the impact-ionization induced multiplication factor. Externally applied voltages (i.e., V_{BE} and V_{CB}) are also shown.

For a bipolar transistor, the external base-emitter voltage (V_{BE}) equals the sum of the internal base-emitter junction voltage V_{BEi}, the voltage drops across the total effective base resistance R_B (due to the total base current I_B), and the emitter series resistance R_E (due to the emitter current I_E) [31]:

$$V_{BE} = V_{BEi} + I_B R_B + I_E R_E, \tag{2.1}$$

where the various employed quantities are also demonstrated in Fig. 2.1. Taking the differentials of V_{BE} in (2.1), we obtain

$$dV_{BE} = dV_{BEi} + R_B dI_B + I_B dR_B + dV_E, \tag{2.2}$$

where $dV_E = (I_E dR_E + R_E dI_E)$, the variation of voltage drop $(I_E R_E)$ across the emitter resistance. As stated earlier in Section 3.1, one of the key traditional difficulties in extracting the base resistance R_B is how to distinguish R_B from the emitter resistance R_E given the observable I-V characteristics; since in normal forward transistor operation regime, R_B and R_E are in series, thus the same current flows through them. This imposes a great standing block, on how to independently extract R_B. This challenge is addressed by adopting the idea employed by other extraction methods in the category our belongs to, which is to measure differentials in V_{BE} (2.2) while forcing a constant emitter current (I_E). This ensures that the voltage drop $(I_E R_E)$ across the emitter resistance will be constant, hence its variation (dV_E) is assured to be zero, provided the emitter resistance (R_E) can be assumed to be constant. Furthermore, we will assume that all measured data relevant to a single parameter extraction will be *local* in bias- and temperature space, i.e., only small variations in bias and temperature will be involved. Under such conditions we may assume that variations of R_E, R_B, and R_{TH} due to self-heating will be of higher order, i.e., negligible in the context of the method. Under these applied conditions, we may conclude from relation (2.2) that

$$dV_{BE} = dV_{BEi} - R_B dI_C, \tag{2.3}$$

where we applied that for fixed I_E we have $dI_B = -dI_C$. From this expression, we can see that the changes dV_{BE} are dependent, among others, on the changes in internal base-emitter voltage dV_{BEi}. This means that we need to determine the causes of the changes dV_{BEi}.

The methods in the category of extraction methods where our belongs to [27, 30], uses the *collector-base avalanche current* to vary the base current or the voltage drop across the total base resistance R_B, independently of the emitter current. As we can see from Fig. 2.2, by increasing the collector-base voltage V_{CB}, the base current I_B decreases and even changes sign at high values of V_{CB} due to avalanche [32, 33]. This increase of V_{CB} for constant emitter current, also has got observable effects on other device characteristics, e.g., on the terminal collector current I_C, which increases due to the generated electron carriers by avalanche multiplication [34], and on the external base-emitter voltage V_{BE}, which decreases as can be verified from Fig. 2.1 and expression (2.1).

At the same time, power is dissipated in the device, and this can be described by

$$P_{\text{diss}} = I_C V_{CB} + I_E V_{BE}. \tag{2.4}$$

This dissipated power, leads to self-heating of the device [20, 35]. Due to the device self-heating, the change dT_j of the junction temperature [1] defined as

$$T_j = T + R_{TH} P_{\text{diss}}, \tag{2.5}$$

where T denotes the ambient temperature, is given by

$$dT_j = dT + R_{TH} dP_{\text{diss}}, \tag{2.6}$$

where, for fixed I_E,

$$dP_{\text{diss}} = V_{CB}\, dI_C + I_C\, dV_{CB} + I_E\, dV_{BE}. \tag{2.7}$$

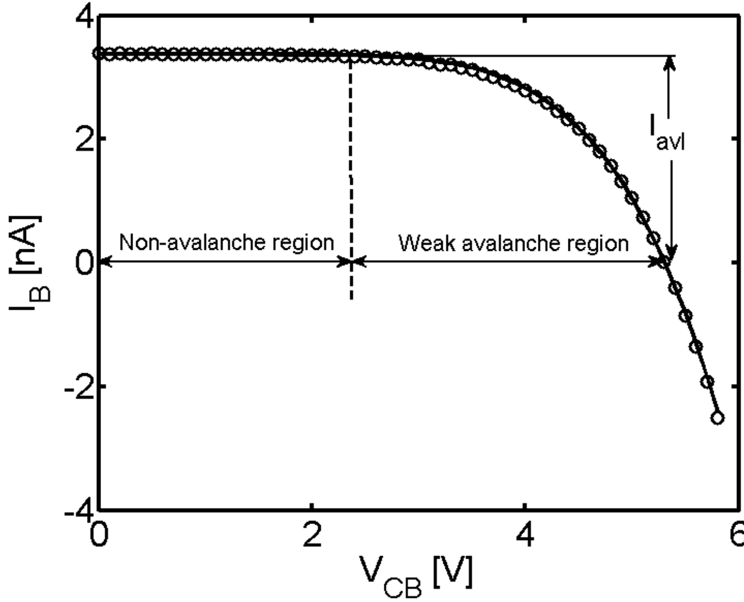


Figure 2.2: Plot for avalanche characteristics as a function of collector-base voltage V_{CB} for fixed non-zero base-emitter voltage V_{BE}. Shown is the base current (I_B) as a function of V_{CB}; due to avalanche, I_B decreases at high V_{CB} and it even changes sign. The generated weak avalanche current $I_{\rm avl}$ is the difference between I_B at very low and high V_{CB}.

We note that in expression (2.6), the term $dR_{TH}P_{\rm diss}$ is taken to be zero, since earlier we assumed that the variations in R_{TH} due to temperature effects will be of higher order (i.e., local in bias and temperature space).

Also, the increase of V_{CB} leads to an increase of the collector-base depletion width [36], which results into a modulation of the neutral base width, and thus the main current of the transistor I_N (in case of an NPN-transistor, the incoming electron current at the base side of the collector-base space-charge region, see Fig. 2.1) [37]. This effect of V_{CB} on I_N is commonly known as the Early effect [26, 38], and can be accounted for by adopting a simple model for the main forward current [2, 24, 39]

$$I_N = I_S(T_j)\exp\left(\frac{q_e V_{BEi}}{k_B T_j}\right)\left(1+\frac{V_{CB}}{V_A}\right), \tag{2.8}$$

where V_A is the forward Early voltage that will be assumed to be much greater than V_{CB} and $I_S(T_j)$ is a temperature dependent forward saturation current. In this expression, k_B is the Boltzmann constant, q_e the elementary charge, T_j is the junction temperature defined by (2.5).

From relation (2.8), the internal base-emitter voltage V_{BEi} can be derived as

$$V_{BEi} = \frac{k_B T_j}{q_e}\left(\ln(I_N) - \ln(I_S(T_j)) - \ln(1+\frac{V_{CB}}{V_A})\right). \tag{2.9}$$

Under weak avalanche conditions (i.e., the current due to the avalanche multiplication in the base-collector space charge region I_{avl}, is very small in comparison to the current due to incident carriers I_N [32, 40]), $I_N \approx I_C$. In this regime, the change of V_{BEi} (dV_{BEi}) due to the changes in the three independent variables I_C, T_j, and V_{CB} is given by

$$dV_{BEi} = \frac{\partial V_{BEi}}{\partial I_C} dI_C + \frac{\partial V_{BEi}}{\partial T_j} dT_j + \frac{\partial V_{BEi}}{\partial V_{CB}} dV_{CB}\,, \tag{2.10}$$

where the partial derivatives are computed as:

$$\frac{\partial V_{BEi}}{\partial I_C} = \frac{k_B T_j}{q_e} \frac{1}{I_C} = V_{T_j} \frac{1}{I_C}\,, \tag{2.11}$$

$$\frac{\partial V_{BEi}}{\partial T_j} = \frac{\partial V_{BEi}}{\partial T_j} - V_{T_j} \left(\frac{1}{I_S(T_j)} \frac{\partial I_S(T_j)}{\partial (T_j)} \right), \text{ and} \tag{2.12}$$

$$\frac{\partial V_{BEi}}{\partial V_{CB}} = V_{T_j} \left(\frac{1}{V_A + V_{CB}} \right), \tag{2.13}$$

with the thermal voltage $V_{T_j} = k_B T_j / q_e$. Substituting the partial derivatives (2.11) - (2.13) into (2.10) and simplifying further, we obtain

$$\begin{aligned} dV_{BEi} &= \left(\frac{V_{BEi}}{T_j} - V_{T_j} \frac{1}{I_S(T_j)} \frac{\partial I_S(T_j)}{\partial T_j} \right) dT_j \\ &+ V_{T_j} \left(\frac{1}{I_C} dI_C - \frac{1}{V_A + V_{CB}} dV_{CB} \right). \end{aligned} \tag{2.14}$$

By defining $\Delta T = R_{TH} P_{\mathrm{diss}}$, we can rewrite the junction temperature T_j, from (2.5) as

$$T_j = T(1 + \frac{\Delta T}{T}). \tag{2.15}$$

When we substitute (2.15) into the combination of (2.6) and (2.14), expand in a series expansion in $\delta_T = \Delta T / T$ and subsequently neglect higher order powers in δ_T and products of δ_T with other small quantities, we arrive at an approximate expression for the changes dV_{BEi} as:

$$\begin{aligned} dV_{BEi} &\approx \left(\frac{V_{BEi}}{T} - \frac{V_T}{I_S(T)} \frac{\partial I_S(T)}{\partial T} \right) (dT + R_{TH} dP_{\mathrm{diss}}) \\ &+ V_T \left(\frac{1}{I_C} dI_C - \frac{1}{V_A + V_{CB}} dV_{CB} \right). \end{aligned} \tag{2.16}$$

With this approximate expression for dV_{BEi}, we can now determine the changes dV_{BE}, as defined by (2.3), which after further simplification, results into

$$\begin{aligned} dV_{BE} &= -\alpha_T (dT + R_{TH} dP_{\mathrm{diss}}) \\ &+ \left((\frac{V_T}{I_C} - R_B) dI_C - \frac{V_T}{V_A + V_{CB}} dV_{CB} \right), \end{aligned} \tag{2.17}$$

where we introduced the following coefficient, the value of which will be important for our extraction method:

$$\alpha_T = -\left(\frac{V_{BEi}}{T} - V_T \frac{1}{I_S(T)} \frac{\partial I_S(T)}{\partial T}\right).$$

Substituting for dP_{diss} from (2.7) into (2.17), and then rearranging the resulting expression, we obtain

$$\begin{aligned}(1 \quad + \quad & \alpha_T R_{TH} I_E) dV_{BE} = -\alpha_T dT \\ + \quad & \left(-\alpha_T R_{TH} V_{CB} + \frac{V_T}{I_C} - R_B\right) dI_C \\ + \quad & \left(-\alpha_T R_{TH} I_C - \frac{V_T}{V_A + V_{CB}}\right) dV_{CB}\,. \end{aligned} \tag{2.18}$$

From this expression i.e., (2.18), we learn how we can measure the value of the introduced quantity α_T. Indeed, the last term on the right hand side (rhs) of (2.18) vanishes if we consider measurements taken at a constant value of V_{CB}. The second term on the rhs of (2.18) can be neglected under fixed I_E conditions on basis of $I_E = I_C + I_B = (1+\delta)I_C$, where $\delta = 1/\beta$, which implies that $dI_C = [-d\delta/(1+\delta)^2]I_E \approx -(d\delta)I_E$, which we shall assume to be negligible. Furthermore, it turns out in practice that at moderate I_E, $(1+\alpha_T R_{TH} I_E) \approx 1$; this can be verified after the thermal resistance R_{TH} has been extracted (see Subsection 2.3.3). Under these assumptions, (2.18) yields

$$\alpha_T \quad = \quad -\frac{dV_{BE}}{dT}. \tag{2.19}$$

This expression means that α_T can be determined from measurements of the external base-emitter voltage as a function of ambient temperature, at fixed emitter current and constant V_{CB} (this will be demonstrated in Subsection 2.3.1). Note that, α_T defines the (magnitude of the) sensitivity of the base-emitter voltage V_{BE} to temperature.

For α_T known and assuming a constant ambient temperature (implying that $dT = 0$) and $(1+\alpha_T R_{TH} I_E) \approx 1$, we can rewrite expression (2.18) as

$$\begin{aligned} -\frac{dV_{BE}}{dI_C} \quad = \quad & R_B + \alpha_T R_{TH}(V_{CB} + V_A^{\text{eff}}) \\ + \quad & \frac{V_T}{I_C} \frac{V_A^{\text{eff}}}{V_A + V_{CB}} \left(1 - \frac{V_A + V_{CB}}{V_A^{\text{eff}}}\right), \end{aligned} \tag{2.20}$$

where $\alpha_T > 0$. We defined the effective Early voltage

$$V_A^{\text{eff}} = I_C \left(\frac{dV_{CB}}{dI_C}\right), \tag{2.21}$$

which is computed directly from the measurements without any further approximations. Note that under weak avalanche current conditions, the quantity (dI_C/dV_{CB}) is relatively small [32, 36], which implies that V_A^{eff} is large. Based on this, we can assume that

$V_{CB} \ll V_A \ll V_A^{\text{eff}}$, then $(V_A + V_{CB})/V_A^{\text{eff}} \ll 1$ and $V_A^{\text{eff}} \approx (V_{CB} + V_A^{\text{eff}})$. Therefore, expression (2.20) can be approximated by

$$-\frac{dV_{BE}}{dI_C} = R_B + \left[\alpha_T R_{TH} + \frac{V_T}{V_A + V_{CB}} \frac{1}{I_C}\right] (V_{CB} + V_A^{\text{eff}}) \,. \tag{2.22}$$

Now assume it is possible to find bias regimes where $I_C = I_E - I_B \approx I_E$ and $V_{CB} \ll V_A$. Under such conditions, expression (2.22) reduces to

$$-\frac{dV_{BE}}{dI_C} \approx R_B + \left[\alpha_T R_{TH} + \frac{V_T}{V_A} \frac{1}{I_E}\right] (V_{CB} + V_A^{\text{eff}}) \,. \tag{2.23}$$

On basis of expression (2.23), we arrive at an extraction procedure for R_B and R_{TH} from DC-measurements. The intercept with the vertical axis of a plot of $-dV_{BE}/dI_C$ as a function of $(V_{CB} + V_A^{\text{eff}})$ gives R_B. The corresponding slope (S_{TOT}) from the same plot is equivalent to $(\alpha_T R_{TH} + V_T/(V_A I_E))$ and it depends on the emitter current (I_E). The thermal resistance (R_{TH}) is found by plotting the slope S_{TOT} as a function of $1/I_E$. Then the intercept (γ) with the vertical axis equals $\alpha_T R_{TH}$; where $\alpha_T = -dV_{BE}/dT$ can be determined separately, on basis of (2.19) [31]. The extracted thermal resistance is then given by $R_{TH} = \gamma/\alpha_T$.
We note that the assumptions behind expression (2.23) have observable consequences, thus careful selection of the input data is needed in order to achieve accurate extracted values of R_B and R_{TH}. Selection approaches of the right region where the input data must be taken from (i.e., extraction region or region of validity) will be discussed in the next section on the experimental results.

2.3 Experimental application and evaluation of the method

In this section, we will demonstrate the extraction method for the base resistance R_B and thermal resistance R_{TH} presented in Section 2.2, using measured data taken on a modern SiGe heterojunction bipolar transistor (HBT). The extraction steps described at the end of Section 2.2, will be followed during this process. In the first subsection, we will describe the measurement setup used to generate the required input measurement data for this method. We will also demonstrate how $\alpha_T = -dV_{BE}/dT$ is determined. The second subsection is dedicated to describing the various methods that can be used to determine the extraction region where the input data for the extraction method should be taken from, in order to achieve correct results for the extracted R_B and R_{TH}. Using the input data from the predetermined extraction region, the corresponding extraction results for R_B and R_{TH} will be presented in the last subsection.

2.3.1 Measurement setup

In order to demonstrate the above developed extraction method of R_B and R_{TH}, we took DC-(on wafer) measurements on a modern high speed (NPN) QUBiC4Xi (BNY-structures) [41] SiGe HBT. The selected bipolar transistor has the emitter area $A_E = W_E \times L_E =$

$0.40 \times 1.0\ \mu m^2$, $BV_{\rm CEO} = 0.8$ V, and $f_T/f_{\rm max}(V_{CE} = 1.5V) \approx 180/190$ GHz at T = 25 °C. Here, the quantities W_E and L_E are the emitter width and length, respectively. The measurement procedure is similar to that proposed by Verzellesi et al. [30] though we extended it to include measurements taken at a sequence of constant emitter currents (I_E). In summary (see also Fig. 2.1), by contacting the external base (B), emitter (E), collector (C), and substrate (S) contact-pads of this device, we measure the voltage across the base-emitter terminals (V_{BE}), the current at the base terminal (I_B), and the current at the collector terminal (I_C) as a function of the applied voltage across the collector-base terminals (V_{CB}), at a sequence of five constant currents at the emitter terminal and constant ambient temperature $T = 25$ °C. We note each of applied constant I_E, at $V_{CB} = 0$ V belongs to the normal forward bias regime of the forward Gummel plots, avoiding high current effects. This suggests that before taking the above stated measurements, first we take the forward Gummel measurements, i.e., for $V_{CB} = 0$ V, we measure the collector (I_C) and base (I_B) currents at respective terminals as a function of the V_{BE} at $T = 25$ °C; from which we select the employed sequence of fixed emitter currents $I_E = I_B + I_C$, over the ideal region.

Since the method employs the avalanche hole current to generate a voltage drop across the effective base resistance (R_B), V_{CB} is taken to be large, in order to generate sufficiently large avalanche current. The corresponding measurement results for I_B, I_C, and V_{BE} as a function of V_{CB} and I_E (taken on a QUBiC4Xi SiGe-HBT) are presented in Fig. 2.3. These measurements will be used for the verification of our R_B and R_{TH} extraction method as presented in the next subsections.

From the measurement results in Fig. 2.3, we observed that the base current I_B (in Figs. 2.3(a) and 2.3(b)) is constant for low V_{CB}, but then it decreases with increasing V_{CB} and even changes sign. Here, the constant behavior of I_B for low V_{CB} values is due to the fact that for such low V_{CB}, the electric field in the collector-base depletion region is very low, this results into low or no impact-ionization, thus low avalanche effect on the terminal characteristics. When V_{CB} is increased, impact-ionization increases (due to the increased field in the collector-base depletion region) [33], which results into a significantly generated avalanche current $I_{\rm avl}$ (see Fig. 2.1), thus a decreased in the measured terminal base current (I_B). These avalanche effects together with self-heating [2, 35, 42] leads to the observed decrease in the measured base-emitter voltage V_{BE} [43] in Fig. 2.3(c). The measured collector current I_C in Fig. 2.3(d) is constant over low and medium values of V_{CB}, but it increases at high values of V_{CB} due to the strong avalanche effects [26, 36]; i.e., at high V_{CB}, impact-ionization induced avalanche multiplication factor (M_n) is significantly high and this results into high collector current $I_C = M_n I_N$ (see Fig. 2.1).

To determine the coefficient $\alpha_T = -dV_{BE}/dT$, presented earlier in (2.19), we measured the extrinsic base-emitter voltage (V_{BE}) at different ambient temperatures (T), for constant $V_{CB} = 0$V and fixed $|I_E| = 4.309$ mA (see Fig. 2.3). The corresponding measured results (symbols) for V_{BE} as a function of ambient temperature are presented in Fig. 2.4; and by linear regression (dashed line), we compute the sensitivity of V_{BE} to temperature as $\alpha_T = -dV_{BE}/dT = 1.006$ mV/K.

Using the measured data for I_C, I_B, and V_{BE}, together with the computed value for α_T

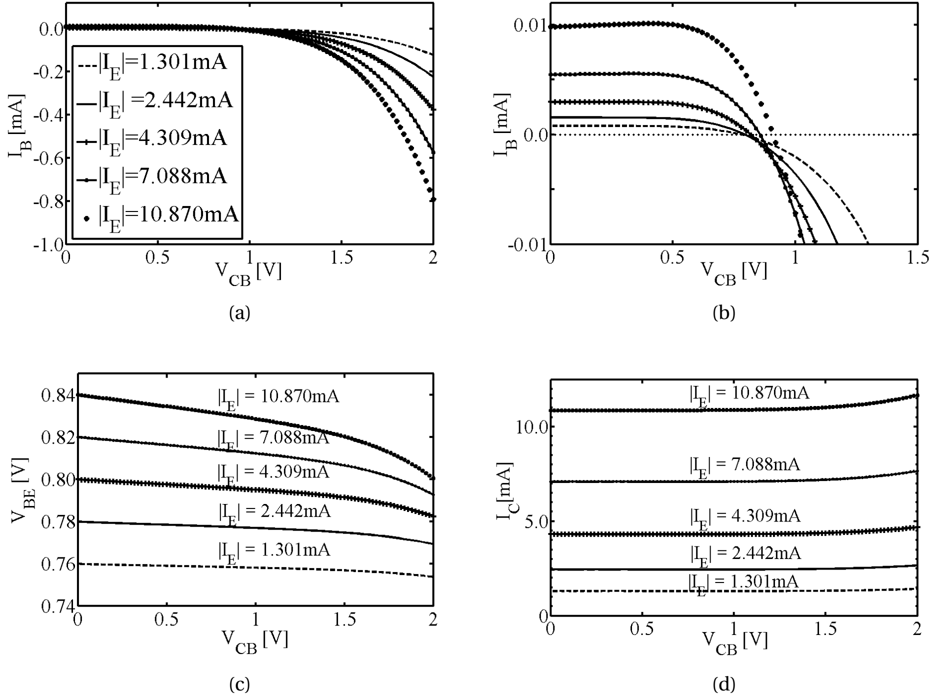


Figure 2.3: Measured base current I_B (*a*), base-emitter voltage V_{BE} (*c*), and collector current I_C (*d*) as a function of collector-base voltage V_{CB} for a sequence of five constant emitter currents I_E. In (*b*) is a zoom in of (*a*) showing that $I_B(I_E)$ is non-zero for lower values of V_{CB}. These measurements were taken on a high speed NPN QUBiC4Xi SiGe HBT (with emitter area $A_E = 0.40 \times 1.0\ \mu\text{m}^2$), at ambient temperature $T = 25\,°\text{C}$. Here, we observe that I_B and V_{BE} decreases with increasing V_{CB}, while I_C increases.

we can now determine the extraction region [regime of physical validity for expression (2.23)] for this employed method. This will be demonstrated in the next subsection.

2.3.2 Determination of the extraction region

We need to determine the extraction region for which the approximations that led to the final expression (2.23) from (2.22) actually hold. As stated earlier, for our extraction method to yield accurate results for R_B and R_{TH}, the input data, must be selected from this regime of physical validity of (2.23). Possible methods of determining this extraction region, are discussed in detail next.

One of the methods to determine the region of validity is to use the plot of the quantity $\partial/\partial X(-dV_{BE}/dI_C)$ as a function of V_{CB}, where $X = (V_{CB} + V_A^{\text{eff}})$, with V_A^{eff} defined by (2.21). From relation (2.23), we have

$$\frac{\partial}{\partial X}\left(-\frac{dV_{BE}}{dI_C}\right) = \alpha_T R_{TH} + \frac{V_T}{V_A}\frac{1}{I_E}\,. \tag{2.24}$$

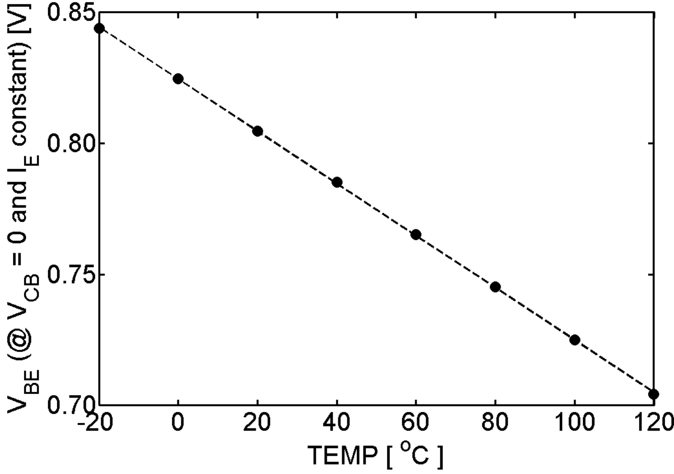


Figure 2.4: Measured (symbols) base-emitter voltage V_{BE} as a function of ambient temperature T. Measurements were taken at a fixed emitter current $|I_E| = 4.309$ mA (corresponding to $V_{BE} = 0.80V$ at $V_{CB} = 0$ V and $T = 25\,°\mathrm{C}$, see Fig. 2.3). From the slope (dashed line) we compute the sensitivity of V_{BE} to T as $\alpha_T = -dV_{BE}/dT = 1.006$ mV/K.

Using the measured data in Fig. 2.3, we determine $\partial/\partial X(-dV_{BE}/dI_C)$ over a sequence of the employed constant emitter currents I_E. The corresponding results for the quantity $[\partial/\partial X(-dV_{BE}/dI_C)]$ as a function of V_{CB} and I_E are presented in Fig. 2.5. From this plot, the extraction region can easily be recognized as the range of V_{CB} values for which $\partial/\partial X(-dV_{BE}/dI_C)$ is (near) constant as a function of V_{CB} [31]. In other words, the check for the approximation that validates the step from expression (2.22) to (2.23) is that for constant I_E, the slope of $-dV_{BE}/dI_C)$ is (near) constant as a function of X. Based on the results in Fig. 2.5 we can observe a dependence of the width of the required extraction region on the applied emitter currents $|I_E|$. The key setback of this method, is that the true extraction region can be affected by the stochastic errors in the measured data. These errors are likely to be magnified further by the differentiation carried out in this method. In order to reduce such stochastic effects, smoothing of the measurement data [44, 45] may be required.

Another possible method for determining the extraction region is based on the definition of the effective early voltage, which was introduced earlier in expression (2.20), i.e., $V_A^{\mathrm{eff}} = I_C(dV_{CB}/dI_C)$. From this definition, we can see that V_A^{eff} is proportional to the inverse of dI_C/dV_{CB}, which implies that low values of $(V_{CB} + V_A^{\mathrm{eff}})$ used in the final expression (2.23), corresponds to high values of V_{CB} as shown in Fig. 2.6. From this figure, we observe that for very low V_{CB} (i.e., $V_{CB} \ll BV_{CEO}$) an accurate extraction of R_B and R_{TH} is impossible, since $(V_{CB} + V_A^{\mathrm{eff}})$ values becomes infinite. The corresponding region is the non-avalanche region that must be excluded from the extraction region. The required extraction region is then the region where very small oscillations of $(V_{CB} + V_A^{\mathrm{eff}})$ as a function of V_{CB} are observed; then the corresponding lower limit of this region is where

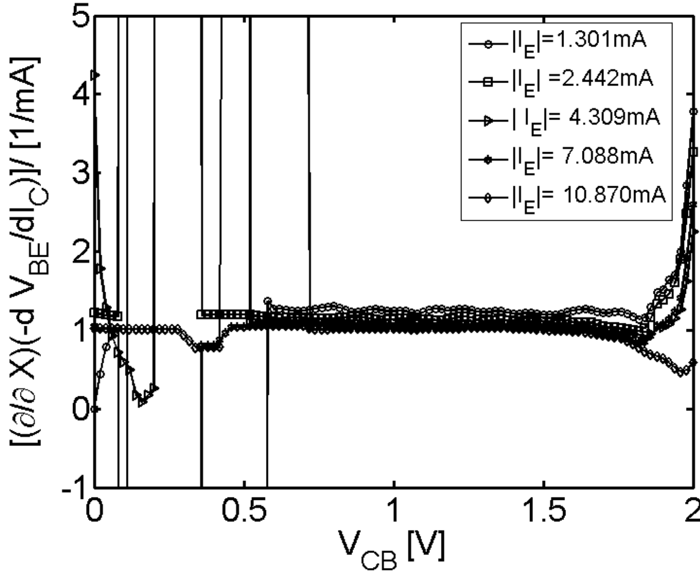
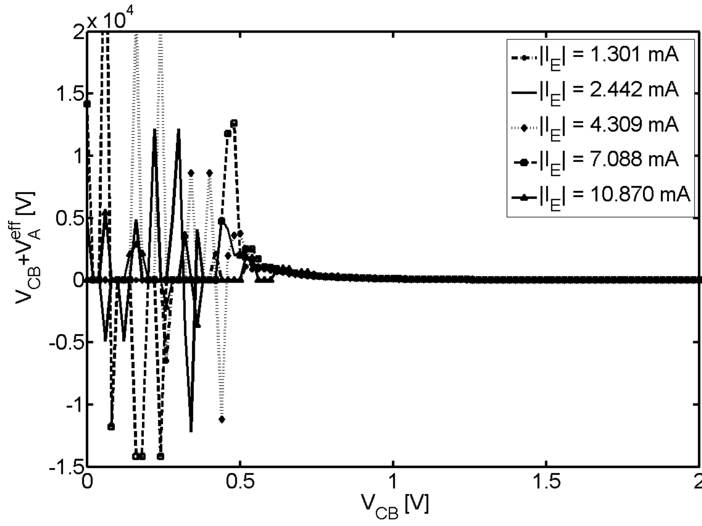


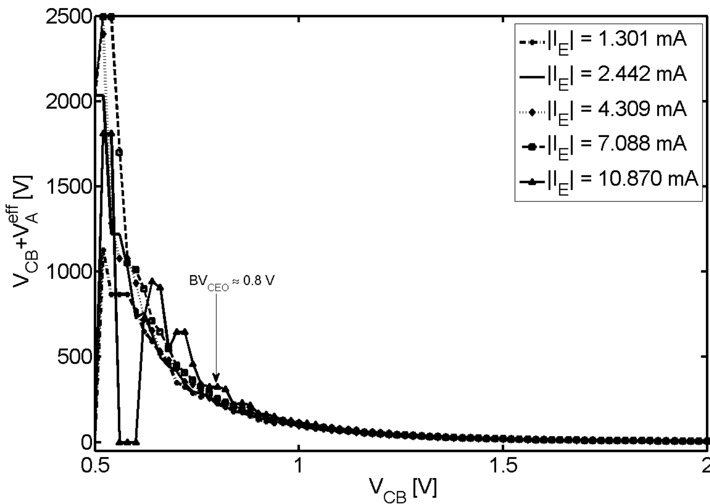
Figure 2.5: Determination of the extraction region for our proposed extraction technique using a plot of $[\partial/\partial X(-dV_{BE}/dI_C)]$ as a function of V_{CB}, for a sequence of constant emitter currents I_E. Large oscillations are observed for small V_{CB} ($V_{CB} < BV_{CEO}/2$) and for very large V_{CB}. This means that this extraction method cannot be used in these two regions as the approximations we made in the derivation of (2.23) are no longer valid. The region of validity is where $[\partial/dX(-dV_{BE}/dI_C)]$ is (near) constant.

small oscillations start from (see Fig. 2.6(a)). Though this method can be used to determine the lower limit for the extraction region, it is limited in terms of the upper limit, since it cannot be used to determine the exact boundary of the upper limit. In addition, by using this method, it becomes difficult to identify the dependence of the lower limit of the required extraction region on the applied emitter current I_E. This method gives an impression that the exact boundary of the lower limit can be uniquely determined (as a function of I_E), which is not true in general as was demonstrated earlier in Fig. 2.5.

A more robust method of determining the region of validity of our extraction method is to use contour plots showing the extracted R_B values against the V_{CB} lower limit (lower limit of the extraction region) and V_{CB} upper limit (upper limit of the extraction region) for all possible combinations of the lower limit and upper limit, containing contour lines of equal extracted R_B [46]. For example, a contour plot corresponding to a fixed emitter current $|I_E| = 7.088$ mA is presented in Fig. 2.7. From this figure, the suitable extraction region can be observed as the large flatland area in the top-right of the contour plot with very small variations (of about $\pm 1.0\ \Omega$) in R_B. From this figure, the boundary of the lower limit of the extraction region is approximately at $V_{CB} = 0.5$ V on the vertical axis. Over this possible extraction region the variation in R_B values is between 18.5 Ω to 19.7 Ω, which is relatively small in comparison to other complementary regions. From the vertical-axis of Fig. 2.7, the non-avalanche region can also be observed, and it corresponds to



(a)



(b)

Figure 2.6: Determination of the extraction region using the V_{CB} dependence of $(V_{CB} + V_A^{\text{eff}})$. In (a) we observed that lower values of V_{CB} yields very high values of $(V_{CB} + V_A^{\text{eff}})$. (b) A zoom in of (a) showing that for $(V_{CB} \ll BV_{CEO})$ accurate extraction of R_B and R_{TH} is impossible as $V_{CB} + V_A^{\text{eff}}$ becomes infinite.

V_{CB} (approximately) below 0.5 V, with variations of R_B between 1.0 Ω and 62.0 Ω. As stated earlier, this region is complementary to the required extraction region. We note that from Fig. 2.7, we cannot easily determine the *true boundaries* of the lower limit of

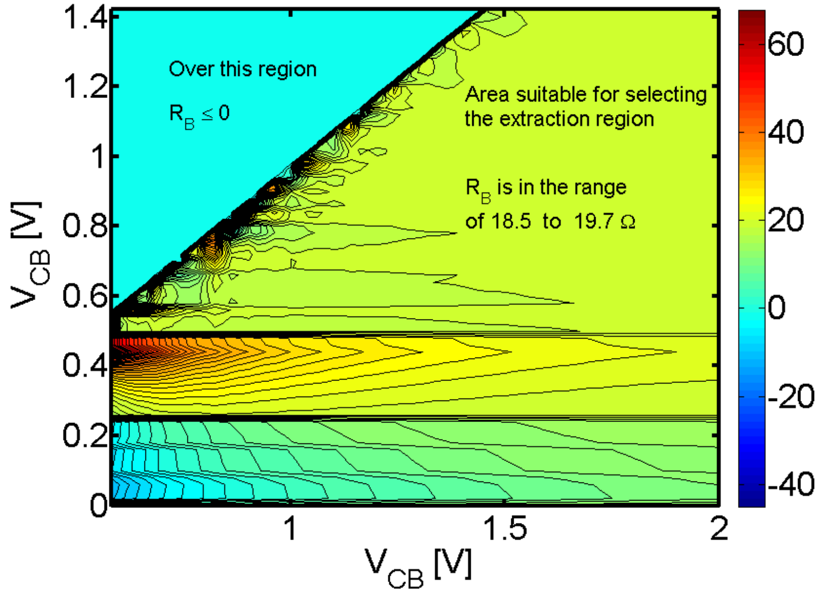


Figure 2.7: Determining the the limits of the extraction region using a contour plot showing lines of equal extracted R_B. These results correspond to a fixed $|I_E| = 7.088$ mA (see Fig. 2.3).

the extraction region, therefore another approach need to be employed; this is discussed next.

The *exact boundaries of the lower and upper limits of the extraction region* can be determined using 2D plots showing the extracted R_B values against the V_{CB} lower limit (lower limit of the extraction region) and against the V_{CB} upper limit (upper limit of the extraction region). Here one of the limits is fixed, and then all possible values of R_B are determined with the other limit being varied and vice verse. For example in Fig. 2.8, we present the results corresponding to the lower limit of the extraction region with (carefully chosen) fixed values of the upper limit. From the results in Fig. 2.8, we can clearly observe the exact boundary of the lower limit of the extraction region at $V_{CB} = 0.50$ V, where small variations in R_B values (of approximately 1.0%) start from, for all chosen fixed values of the upper limit. Below this limit, we observe large variations in R_B values (of approximately 65.0%). This region corresponds to the non-avalanche region and should be excluded from the extraction region as it will yield wrong results for the extracted R_B and R_{TH}. Note that from this figure, one can get an approximation of the exact value for the upper limit of the region of validity; which for this case is approximately at $V_{CB} = 1.30$ V.

A similar exercise as the one presented in Fig. 2.8 can be repeated to determined the exact value of the upper limit of the extraction region. This is done by careful choosing some fixed values of the lower limit of the extraction region, and then all possible R_B values are determined and plotted on the 2D-plot with a variable upper limit. A quick

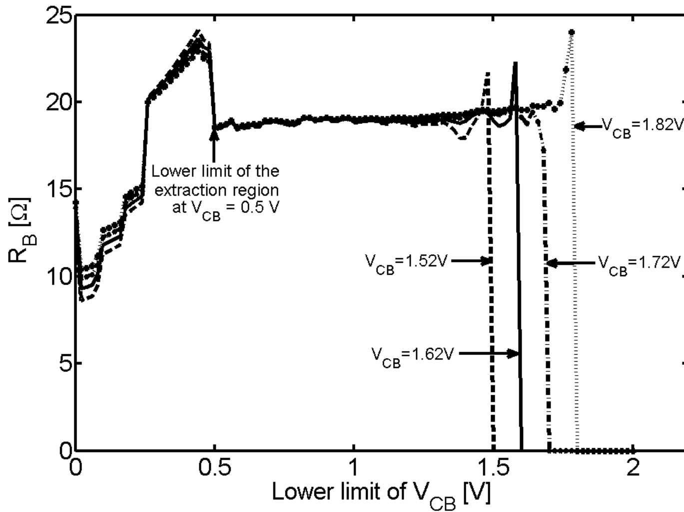


Figure 2.8: Determination of the exact boundary of the lower limit of the extraction region using a 2D-plot, for some fixed values of the upper limit. From the plot, we observe that the exact boundary of the lower limit of the extraction region is at $V_{CB} = 0.5$ V. These results correspond to a fixed $|I_E| = 7.088$ mA (see Fig. 2.3).

approach is to use the known value of the lower limit as the input fixed value, to determine R_B values for variable upper limits; then a similar exercise as the one we used for the lower limit is used to determine the true value for the upper limit. The major aim of this exercise is to exclude the strong avalanche region from the required extraction region, since the applied assumptions during the derivation of our extraction method are not applicable over this regime.

In the method described above for determining the extraction region (region of validity), we used DC-measurements corresponding to a single constant emitter current (i.e., $|I_E| = 7.088$ mA). But as we described in Section 2.2, to simultaneously extract R_B and R_{TH} we need DC-measurements taken on a sequence of constant emitter currents (see Fig. 2.3). Therefore, for each employed fixed emitter current, the corresponding extraction region has to be determined independently. Though this method of determining the extraction region is more involved/demanding compared to the previous two methods, it has a great advantage of determining the true lower and upper boundaries of the region of validity (as a function of I_E) in comparison to the other two methods. Therefore, it should be applied if one needs to determine accurately the right extraction region where the input data for the extraction method should be taken from, as this will results into accurate extracted values of R_B and R_{TH}. We note that, though we employed R_B in our demonstrations, but R_{TH} can as well be used to determine the required extraction region.

After determining the extraction regions corresponding to a sequence of the employed constant emitter currents, the extraction steps as described at the end of Section 2.2 are followed to extract both R_B and R_{TH} for a given bipolar transistor (BJT or

HBT). This is will be demonstrated in next subsection.

2.3.3 Extraction of both R_B and R_{TH}

By selecting the input measured data from the predetermined extraction regions (as a function of I_E), both the base resistance R_B and thermal resistance R_{TH} for the employed QUBiC4Xi SiGe-HBT can simultaneously be extracted, following a step-by-step process as described at the end of Section 2.2. Using the measured data presented in Fig. 2.3, we plot $(-dV_{BE}/dI_C)$ as a function of $(V_{CB}+V_A^{\mathrm{eff}})$ for different fixed values of I_E; the corresponding results are presented in Fig. 2.9(a). In this figure, open symbols correspond to

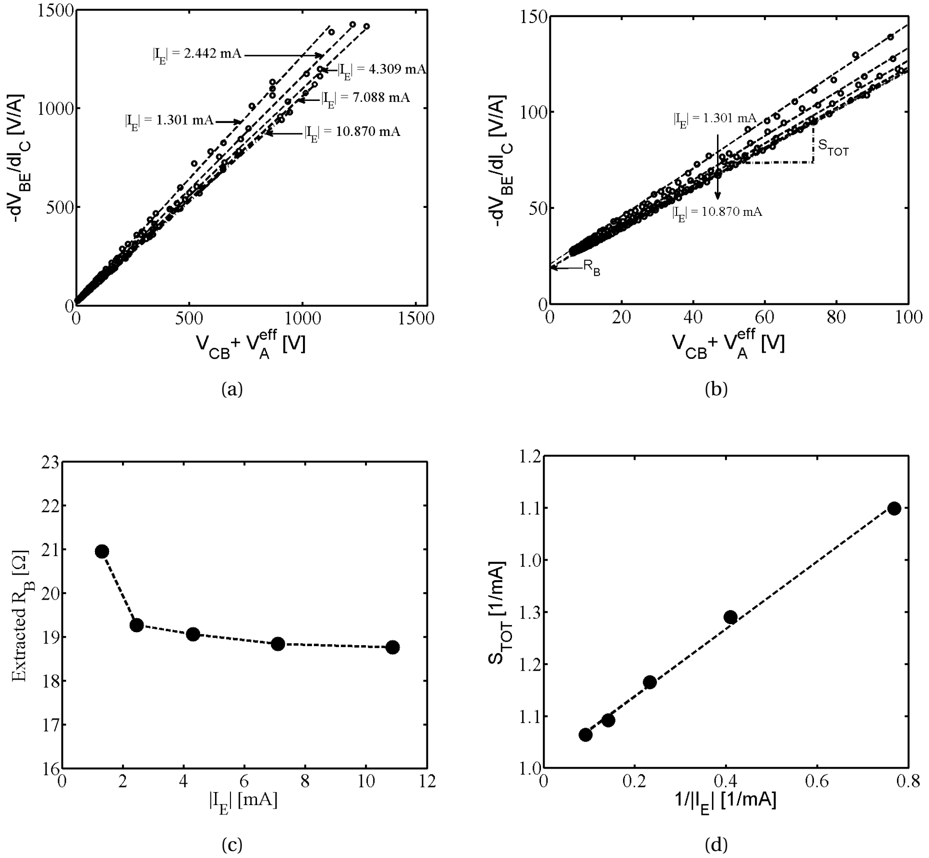


Figure 2.9: Extraction of R_B and R_{TH} using the slope and offset of (2.23) for a device with $A_E = 0.40 \times 1.0$ μm^2. In (*a*) is a plot of $-dV_{BE}/dI_C$ as a function of $(V_{CB}+V_A^{\mathrm{eff}})$ and $|I_E|$ [measurements (symbols) and linear fits (dashed lines)]. Figure (*b*) is a zoom-in of (*a*) for lower values of $(V_{CB}+V_A^{\mathrm{eff}})$ showing extracted R_B as the intercept on the vertical axis; the slope S_{TOT} is also represented. The extracted values of R_B [from vertical-axis intercept of (b)] as a function of $|I_E|$ are presented in (c). In (*d*), is the slope $\mathrm{S_{TOT}}$ [from (a)] as a function of $1/|I_E|$ [measurements (closed circles) and linear fit (dashed line)]; the vertical intercept $\gamma = \alpha_T R_{TH}$ and is used to extract R_{TH}.

the measured data, while the dashed lines correspond to linear fits, for the different employed emitter currents (but for each V_{CB} sweep fixed, i.e., $0 \leq V_{CB} \leq 2.0$). The intercept with the vertical axis (as shown in Fig. 2.9(b)) yields the extracted value of R_B for each employed constant I_E. Results for the extracted total base resistance R_B as a function of $|I_E|$ are presented Fig. 2.9(c); from which a gentle decrease of the extracted values of R_B with I_E is observed. This observed R_B behavior can be attributed to the increase of base diffusion charge with increasing I_E, which leads to a decrease in the intrinsic part of R_B [2, 34] (to be discussed further in Section 2.5). From Fig. 2.9(c), we observe that variations in the extracted R_B as a function of $|I_E|$ are small ($\approx \pm 2\ \Omega$) and for this RF-device we find $R_B \approx 19\,\Omega$.

A plot of the slope S_{TOT} (corresponding to linear fits in Fig. 2.9(a)) as a function $1/|I_E|$ is presented in Fig. 2.9(d), by the closed circles. The corresponding linear fit (dashed lines) yields the intercept with the vertical-axis $\gamma = 1.014$ [1/mA]. Using the earlier defined relation $\gamma = \alpha_T R_{TH}$, where $\alpha_T = 1.006$ mV/K, we get the extracted value for the thermal resistance of this RF-device as: $R_{TH} = 1.01 \times 10^3$ K/W.

Note that the observed linear behavior in Fig. 2.9, provides an a posteriori check on the previously applied assumptions [i.e., that R_B is constant or independent of V_{CB} (see (2.3)) and that $V_{CB} \ll V_A$ (see (2.8))]. In Fig. 2.10, we present results corresponding to $|\alpha_T R_{TH} I_E|$ as a function of the fixed emitter currents (I_E) employed in the previous results, e.g., in Fig. 2.9. From this figure, we can clearly see that for all employed values of I_E, $|\alpha_T R_{TH} I_E| \ll 1$. This justifies the assumption $(1 + \alpha_T R_{TH} I_E) \approx 1$, used in relation (2.20) [31].

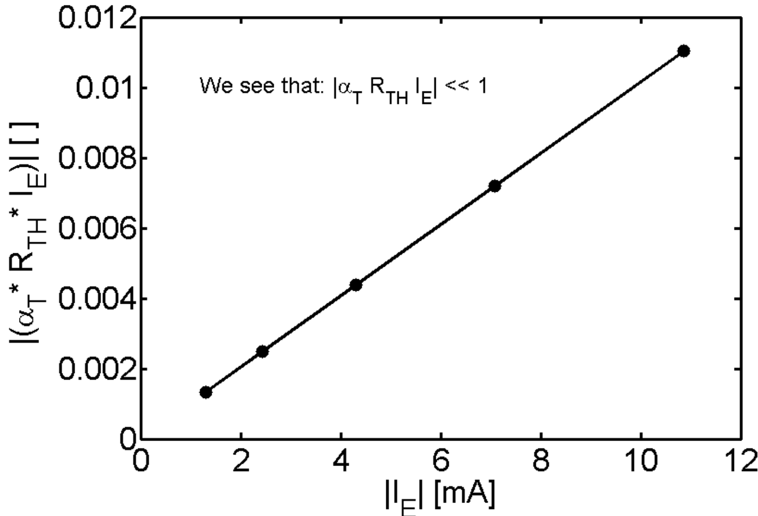


Figure 2.10: Plot for the computed values of $|\alpha_T R_{TH} I_E|$ as a function of previous employed values for $|I_E|$. From this plot, we can see that for all employed values of $|I_E|$, $|\alpha_T R_{TH} I_E| \ll 1$, which justifies the assumption $(1 + \alpha_T R_{TH} I_E) \approx 1$, used earlier in expressed (2.20).

2.4 Comparison to other R_B and R_{TH} extraction methods

We compare results from our extraction method for the base resistance R_B and thermal resistance R_{TH} described in Section 2.2 and demonstrated in Section 2.3 with results from other independently published methods. This can be seen as a verification of our method against the earlier published methods, and our extracted results for R_B and R_{TH} will be compared with alternative extraction methods. Firstly, our extracted results of R_B and R_{TH} will be compared with those from the DC-method proposed by Vanhoucke and Hurkx [27]. In addition, the extracted R_{TH} will be compared with corresponding results from another (DC-) extraction method for R_{TH} proposed by Vanhoucke et al [25]. Secondly, the extracted results of the total base resistance R_B will be compared with results from the traditional AC-methods (i.e., standard two port method and the standard circle impedance method [28, 29]). Lastly, the extracted R_B results from our method will be compared with alternative results extracted from noise measurements.

2.4.1 DC extraction methods

R_B and R_{TH} extraction method proposed in [27]

The extracted results for base resistance R_B and thermal resistance R_{TH} from our method are compared with similar results from the method that determines R_B and R_{TH} by considering the change of the internal base-emitter voltage V_{BEi} due to self-heating only [27]. In summarized form, this method is given by the expression [see equation (10) of [27]]

$$-\frac{dV_{BE}}{dI_C} \approx R_B + [\alpha_T R_{TH}]\,(V_{CB} + V_A^{\text{eff}})\,. \tag{2.25}$$

A comparison of this expression (2.25), with our expression (2.23), shows that (2.25) ignores the second term is the square brackets of (2.23). In other words, our expression (2.23) tends to (2.25) in the limit of extremely large Early voltage V_A. In the context of the present work, such can be interpreted as ignoring the influence of Early effect completely on the internal base-emitter voltage V_{BEi}. Following the same extraction procedure as that stated earlier, we can see that the slope ($S_{TOT} = \alpha_T R_{TH}$) of (2.25) is independent of the emitter current I_E; this is in contrast to what we observe from the measurements in Fig. 2.3 and Fig. 2.9.

Base resistance R_B

By considering a constant value for the emitter current I_E, the extraction method proposed in [27], i.e., expression (2.25) will yield the same extracted value of R_B as our extraction method defined by (2.23). This is a result of the fact that in both methods, a plot of $-dV_{BE}/dI_C$ as a function of $(V_{CB} + V_A^{\text{eff}})$, yields the intercept to the vertical-axis which is equivalent to R_B; thus when both methods are applied on the same measured data, the extracted R_B value will be the same, since they both (identically) take into account the influence of self-heating on V_{BEi}. But this is not the case for the extracted value of the thermal resistance R_{TH}, as we demonstrate next.

Thermal resistance R_{TH}
As stated earlier, the extraction method described by expression (2.25) [27], ignores the influence of the collector-base Early effect on internal base-emitter voltage V_{BEi}; this can result into an overestimate of the extracted R_{TH}, due to an underestimate of the dissipated power [34]. This stems from the fact that, neglecting the Early effect leads to an underestimate of the collector current, which results into an underestimate of the dissipated power $P_{\text{diss}} = I_C V_{CB} + I_E V_{BE}$, and thus an overestimate of $R_{TH} = (T_j - T)/P_{\text{diss}}$ [see (2.5)].

In Fig. 2.11, the solid line with closed circles are the results from this work (i.e., our proposed extraction method) and the dashed line with open squares are results from the method proposed in [27]. From this figure, we observe that for a sequence of constant emitter currents I_E, the method proposed in [27], yields extracted values of R_{TH} that are consistently larger than those from our proposed extraction method [31], for all employed values of emitter width (i.e., $W_E = 0.15,\ 0.30, 0.40\ \mu$m), with emitter length $L_E = 1.0\ \mu$m fixed. Here, the relative difference in the extracted R_{TH} values as a function of $1/W_E$ is approximately between 10% and 25%. We note that, this plot by itself does not specify which of the two methods yields the correct value of the extracted R_{TH}, this will be discussed in Section 2.5.

Summary
Here we compared the extracted values of R_B and R_{TH} using the method presented in this work, with those from the DC-method proposed in [27], which is summarized in (2.25). From these results, we see that for a fixed emitter current I_E, both methods would yield the same for the extracted value of the base resistance R_B (since they both (identically) take into account the influence of self-heating on V_{BEi}), but different values of the thermal resistance R_{TH}. The key reason for the variations in the extracted R_{TH} values is that the method proposed in [27] does not contain second term in the square brackets of expression (2.23), which is responsible for addressing the influence of Early effect on V_{BEi}. From the plot of $-dV_{BE}/dI_C$ as a function of $(V_{CB} + V_A^{\text{eff}})$, we see that both methods yields the same value for the intercept of the vertical axis, which is equivalent to R_B. But since our method takes the Early effect into account, while the method in [27] does not, then both methods yield different values for the R_{TH} as demonstrated in Fig. 2.11. Therefore, in relation to the method proposed by Vanhoucke and Hurkx [27], our method presented in Section 2.2, yields improvements in the extracted value for the thermal resistance R_{TH}, but both methods yields the same value of the base resistance R_B.

R_{TH} extraction method proposed in [25]

The extracted results for the thermal resistance R_{TH} from our new extraction method, are compared with those from another DC-method proposed by Vanhoucke et al [25]. This method is a modification of another method proposed in [3], which uses the base-emitter voltage V_{BE} as a temperature-sensitive electrical parameter. In Vanhoucke et al [25] the same extraction technique for R_{TH} is used, but a revised compensation is employed in order to correctly take into account the self-heating of the device during the

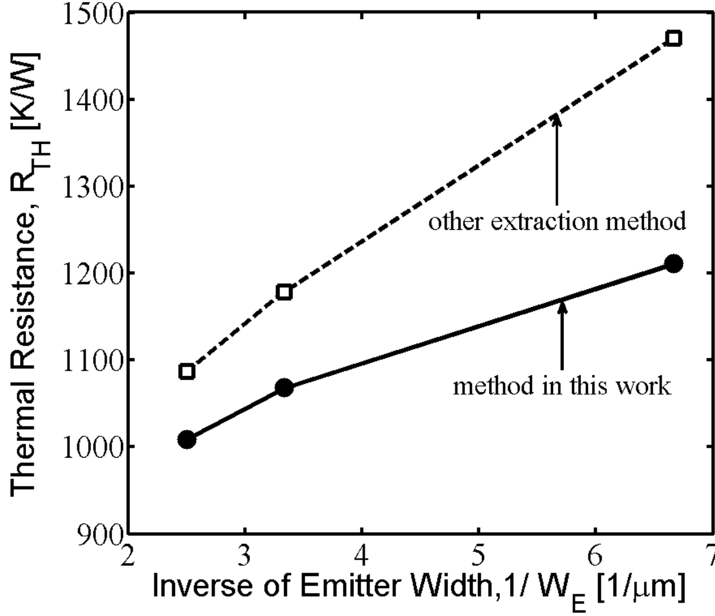


Figure 2.11: Extracted thermal resistance R_{TH} as a function of the inverse of the emitter width W_E (with constant emitter length $L_E = 1.0$ μm), using the method described in this work (solid line with closed circles) and verification using extraction method proposed by Vanhoucke and Hurkx [27] (dashed line with open squares). We observe that the extraction method proposed in [27] yields systematically higher values for the extracted R_{TH} in relation to the extraction method presented in this work.

measurements. The same type of measurements as those presented in Subsection 2.3.1 are employed in this method, though with a different type of data treatment.

With the emitter current I_E fixed and collector-base voltage V_{CB} taken to be constant, the base-emitter voltage V_{BE} is measured for different ambient temperatures T_{amb}. This results into a linear relation between V_{BE} and T_{amb}, which can be approximated [25] by

$$V_{BE} = A + BT_{\text{amb}}, \tag{2.26}$$

where $B = -\alpha_T$, with α_T defined by (2.19). In this measurement setup, since V_{BE} changes with the applied ambient temperature (see Fig. 2.4), it implies that there is a change in dissipated power $P_{\text{diss}} = I_E V_{BE}$ (where V_{CB} is considered to be constant and equal to zero). In the other measurement, the same fixed I_E is used, but now T_{amb} is taken to be constant, then the dissipated power P_{diss} is increased by means of increasing V_{CB}. According to [25], in the first measurement setup, due to the dissipated power $P_{\text{diss}} = I_E V_{BE}$, the increase in junction temperature T_j is defined as in (2.5), i.e.,

$$T_j = T_{\text{amb}} + R_{TH} P_{\text{diss}} = T_{\text{amb}} + R_{TH} I_E V_{BE}. \tag{2.27}$$

A combination of (2.27) with (2.26), yields

$$V_{BE} = A + B(T_j - R_{TH} I_E V_{BE}). \tag{2.28}$$

By assuming $V_{BE} = \alpha + \beta P_{\text{diss}}$, with α and $\beta < 0$ as fitting parameters [19], the above relations can be used to obtain the relation between T_j and P_{diss} as [25]

$$T_j = B^{-1}\left[-A + \alpha(1 + BR_{TH}I_E) + \beta(1 + BR_{TH}I_E)P_{\text{diss}}\right], \tag{2.29}$$

and the thermal resistance ($R_{TH} = dT_j/dP_{\text{diss}}$) as

$$R_{TH} = \frac{\beta}{B(1-\beta I_E)}. \tag{2.30}$$

Since $B = -\alpha_T = -1.006$ mV/K, we can see from relation (2.30) that we only need the value of parameter β in order determine R_{TH}. This parameter can be determined from the relation $V_{BE} = \alpha + \beta P_{\text{diss}}$, by plotting V_{BE} as a function of the dissipated power P_{diss} (using the second measurement setup data) and the corresponding slope gives β. By considering the same fixed emitter current value as that used in Fig. 2.4 when determining α_T, i.e., $|I_E| = 4.309$ mA, β is computed from the linear fit in Fig. 2.12 as $\beta = -1.1008$ mV/W. Employing this value of β together with (2.30), we obtain $R_{TH} = 1.10 \times 10^3$ K/W. A comparison of this value for R_{TH}, with the one extracted using our extraction method, i.e., $R_{TH} = 1.01 \times 10^3$ K/W (see Subsection 2.3.3), we obtain a percentage error (of the deviations) of approximately 8.90%; which shows that the two method yields results for R_{TH} that are fairly close one another. The inset in Fig. 2.12, shows the junction temperature T_j as a function of the dissipated power P_{diss}, computed from relation (2.29).

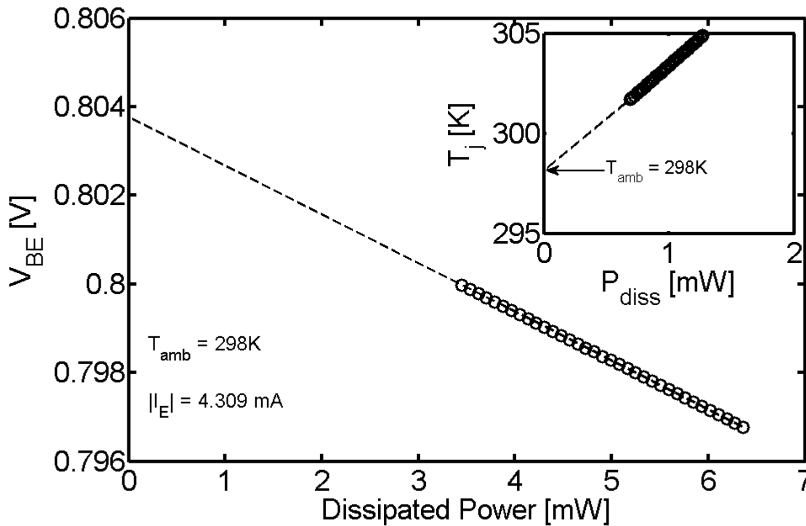


Figure 2.12: Measured base-emitter voltage V_{BE} as a function of dissipated power P_{diss}. Measurements were taken at a fixed emitter current $|I_E| = 4.309$ mA and constant ambient temperature $T_{\text{amb}} = 25\,°\text{C}$. From the line fit we compute $\beta = dV_{BE}/dP_{\text{diss}} = -1.099$ mV/W. The inset presents the corresponding increase in junction temperature T_j as a function of P_{diss}.

Discussion

Though we compared the extraction method presented in this chapter with a method proposed by Vanhoucke et al [25], and a relatively acceptable deviation in the extracted R_{TH} value was achieved, we cannot comment much on these results. This stems from the fact that we have got some doubts on the correctness of the method proposed in [25], because in the derivation of this method, two relations that seem to be different were equated to one another, and these led to the final expression (2.29), which is then used to compute R_{TH} from (2.30). Lets elaborate further how (2.29) comes about. We consider

$$V_{BE} = \alpha + \beta P_{\text{diss}} = A + B(T_j - R_{TH} I_E V_{BE}). \quad (2.31)$$

By making T_j the subject, we obtain

$$T_j = \frac{(\alpha + \beta P_{\text{diss}} - A)}{B} + R_{TH} I_E V_{BE}. \quad (2.32)$$

Now we again substitute $V_{BE} = \alpha + \beta P_{\text{diss}}$ into (2.32), which yields

$$T_j = \frac{(\alpha + \beta P_{\text{diss}} - A)}{B} + (R_{TH} I_E)(\alpha + \beta P_{\text{diss}}) = \frac{(1 + B R_{TH} I_E)(\alpha + \beta P_{\text{diss}}) - A}{B}. \quad (2.33)$$

This expression is simplified further obtain (2.29).

We note that the first temperature measurements for V_{BE} that led to expression (2.26) were carried out at a fixed I_E and constant V_{CB} (= 0 V for this case). The corresponding dissipated power $P_t = I_E V_{BE}$, and is only applicable under these measurement conditions. Putting this together, we get

$$V_{BE} = A + B(T_j - R_{TH} I_E V_{BE}) = (A + B T_j) + (-B R_{TH}) P_t = \alpha_1 + \beta_1 P_t. \quad (2.34)$$

where $\alpha_1 = (A + B T_j)$ and $\beta_1 = (-B R_{TH})$. In the second measurements, same fixed I_E is used, but now T_{amb} is constant and V_{CB} is varied (leading to variation of $P_{\text{diss}} = I_C V_{CB} + I_E V_{BE}$). From this, we can write (assuming a linear relation between V_{BE} and P_{diss})

$$V_{BE} = \alpha + \beta P_{\text{diss}}. \quad (2.35)$$

Basing on the above relations, we can see that expression (2.31) is only valid when: (1) $P_{\text{diss}} = P_t$ (meaning at $V_{CB} = 0$),(2) $\alpha = \alpha_1$, and (3) $\beta = \beta_1$. Also taking V_{BE} in (2.32) to be equal to $(\alpha + \beta P_{\text{diss}})$, is unconvincing, since these seem to be totally different voltages. Such observations leaves a number of unanswered questions on the validity of the extraction method for R_{TH} proposed by Vanhoucke et al [25]. We will not pursue this further as it is not relevant for the present work.

2.4.2 RF extraction methods

In this subsection, we compare the extracted total base resistance R_B using our extraction method presented in Section 2.2 with the results from two traditional *AC-methods*, i.e., the standard two port method and the standard circle impedance method [28, 29], that are both based on small-signal high frequency admittance measurements. The corresponding results are presented in Fig. 2.13. Since in the *standard two port method* the

external base-collector capacitance is assumed to be connected to the external base, the base current partially flows through the capacitance, and this causes the extracted R_B to be overestimated [28]. On the other hand, the *standard circle impedance method* does not account for the AC current which passes through the external base-collector capacitance and as a result, the extracted R_B is underestimated [28].

In [28] and [29], it has been experimentally shown that the true value of the total base resistance lies between the value extracted using the standard two port method and that of the standard circle impedance method. According to Fig. 2.13, our extracted value of R_B using the method described in Section 2.2 (solid line with closed circles) indeed nicely falls between the standard two port (dashed lines with open squares) and standard circle impedance (dashed lines with open triangles) methods, which is supportive for the validity of our method [31].

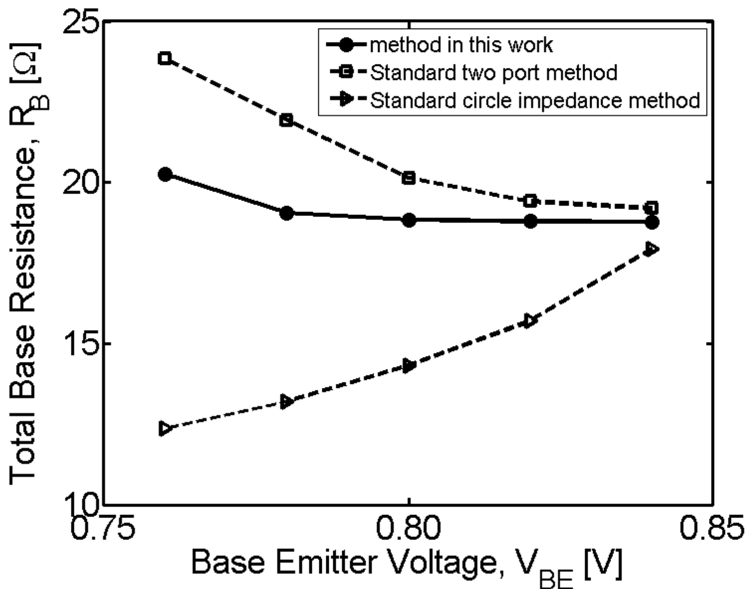


Figure 2.13: Extracted base resistance R_B as a function of V_{BE} using the method described in this work (solid line with closed circles), and verification using small signal standard two port method (dashed lines with open squares) and standard circle impedance method (dashed lines with open triangles), which are described in [28] and [29]. Results corresponding to our method, falls nicely in between the results from the other two small signal methods, which is supportive of the validity of our extraction method.

2.4.3 Noise measurements method

To enable comparison of the results for the the base resistance R_B from the method presented in this work, with those determined from the noise measurements, we carried out R_B extraction using DC measurements taken on a modern QUBiC4mmW SiGe HBTs [47].

The advanced QUBiC4mmW process is similar to the advanced QUBiC4Xi process (used earlier) but with a reduced total base resistance. The extracted R_B results corresponding to different QUBiC4mmW devices (with different geometry and base layout) are presented in Table 2.1.

Noise parameter measurements were also taken on the same QUBiC4mmW devices and R_B was determined from these noise measurements for each employed device [48]. This method of determining the total base resistance (R_b) from the noise measurements is based on the observation that the noise parameter R_n is approximately [49]:

$$R_n \approx R_b + \frac{1}{2g_m}. \tag{2.36}$$

To get good accuracy, the extended transistor equivalent circuit depicted in Fig. 2.14 was assumed. In this circuit, R_b, R_c, and R_e are the base, emitter, and collector resistances,

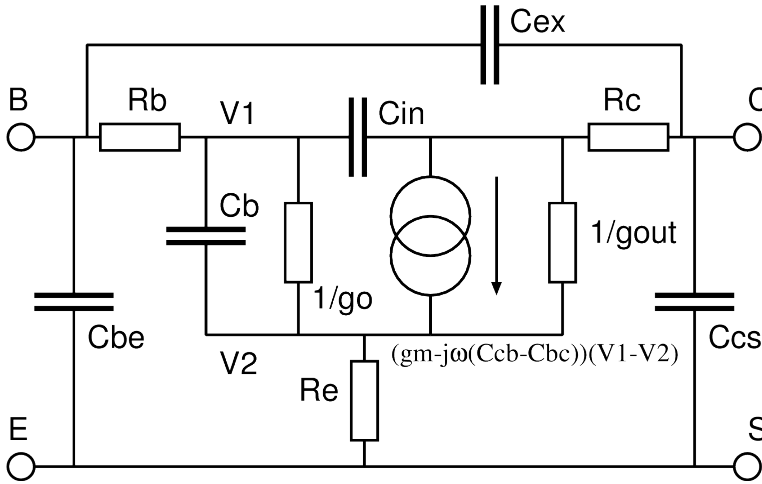


Figure 2.14: Small-signal equivalent circuit assumed for base resistance R_b extraction from S-parameter and noise parameter data.

respectively, C_{ex} and C_{in} are the external and internal base-collector capacitances, respectively, C_b is the sum of the base-emitter depletion and diffusion capacitances, g_m is the transistor's forward transconductance, g_o and g_{out} are the conductances of the base-emitter and base-collector junctions respectively, C_{be} is the base-emitter overlap capacitance, and C_{cs} is the collector-substrate capacitance. S-parameters and noise parameters [1, 2, 43] were measured in the 1 GHz to 50 GHz range at $V_{BE} = 0.8$ V and $V_{CE} = 1.0$ V.

In a first extraction step the values of g_m, g_o, g_{out}, R_b, R_e, R_c, $C_{cb} - C_{bc}$, C_{in}, C_b, C_{be}, C_{ex}, and C_{cs} were optimized to fit the measured Y-parameters after de-embedding. After, the equivalent circuit depicted in Fig. 2.14 was embedded between the measured test-structure parasitics, as inferred from the open and short dummy structures, a second comparison between measured and simulated S-parameters at the probe tips was made.

Finally shot noise currents given by [48]

$$\begin{aligned} S_{i,bc} &= 2kTj\omega(C_{cb} - C_{bc}), \\ S_{i,cb} &= -S_{i,bc}, \\ S_{i,cc} &= 2qI_c, \\ S_{i,bb} &= 2qI_b + S_{i,bc}S_{i,cb}/S_{i,cc}, \end{aligned} \tag{2.37}$$

were added to the intrinsic transistor, and the usual thermal noise voltages were added to R_b, R_e, and R_c, and to the parasitic resistances found in the test-pads. Then in a second extraction step the values of R_b, R_e, R_c, $C_{cb} - C_{bc}$, C_{in}, C_b, C_{be}, C_{ex}, and C_{cs} are readjusted for a best fit of both the simulated S-parameters and the simulated noise parameters to the measured ones. In this extraction step deviations between measured and simulated S-parameters and Γ_{opt} values are weighted equally, whereas deviations between measured and simulated $F_{\min}$ and R_n values are weighted stronger. Typical, after the fit to the de-embedded Y-parameters, the extracted R_b is roughly the same as obtained from the circle impedance method [28, 29], which does not come as a big surprise, as the methods are very similar. As soon as we take the measured noise parameters into account we find that R_b needs to be increased by about 50% to get a proper fit to the measured R_n value [48].

The bipolar devices used in the analysis have 20 emitter fingers ($MULT$) with the first two having emitter area $A_E = 0.30 \times 1.0\ \mu\text{m}^2$ either with a standard or a reduced(*) base resistance layout, and the third one with a square emitter of area $A_E = 0.40 \times 0.40\ \mu\text{m}^2$. In Table 2.1 the extracted R_B for each device using our extraction method (see Section 2.2) are compared with those determined from the noise parameter measurements. From the results in Table 2.1, we can observe that the extracted total base resistance R_B

Table 2.1: Comparison of our extracted values of R_B with those determined from the noise measurements

$W_E \times L_E \times MULT$	R_B [Ω] (our method)	R_B [Ω] (noise method)
0.3 × 1.0 × 20	12.85	10.63
0.3 × 1.0 × 20*	10.58	8.76
0.4 × 0.4 × 20	20.65	17.19

using the method presented in this work is larger (by ≈ 22.0%) than that extracted from the noise measurements [48], for all employed devices. The actual cause of this offset is not well known, but it can be related to the fact that employed measurements are practically taken on test structures with different layout, i.e., DC-devices and RF-devices. Such variation in device layouts may lead to number of effects (such as modulation of contact resistances), which may transfer to the final extracted values. Though these methods have different fundamental basis, i.e., noise is a result of macroscopic effects induced by fluctuations occurring inside the transistor (microscopic noise sources), whereas our method is a result of variation of the base and collector currents (independently of the emitter current) using the collector-base weak avalanche current, the two methods yield

extracted R_B that is closely related. Therefore, these results suggests that our proposed DC-extraction method can be used as a viable alternative to such very expensive noise measurement method for extracting the total base resistance R_B of bipolar transistors.

2.5 Self consistency check of the extraction method

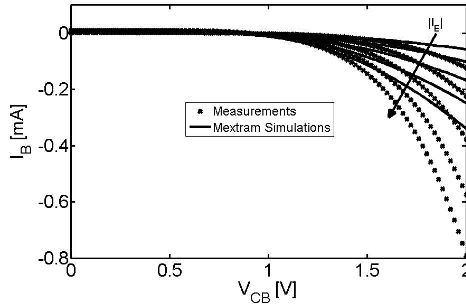
To study the accuracy levels of our proposed extraction method for R_B and R_{TH}, we carried out a self-consistency check of the method using *simulations results* from the World Standard Compact model for bipolar junction transistors MEXTRAM [2, 34] as the input data for the extraction method (instead of the measured data). In the model simulations, the Mextram 504.10 version with self-heating was used [2]. A QUBiC4Xi RF SiGe HBT parameter set was used to generate the simulation results which are presented in Fig. 2.15 by the solid curves. In this figure, the symbols correspond to the measurement results presented earlier in Fig. 2.3, they are presented here just to show that a practical model parameter set was used in the Mextram model simulations. The relevant parameters used in the Mextram model simulations are shown in Table 2.2, were the parameter R_B “comprises" of the variable and constant parts of the total base resistance [26, 48, 50].
In Fig. 2.15(a), the decrease of the base current (I_B) with collector-base voltage (V_{CB})

Table 2.2: Base and thermal resistance parameter values used in the Mextram model simulations

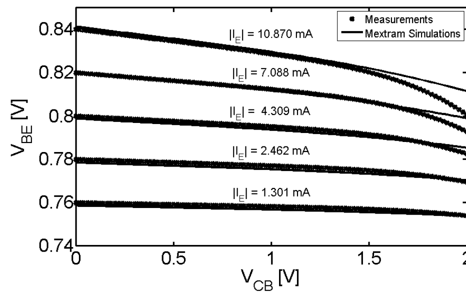
Model parameter name	Employed parameter value
R_B	20 [Ω]
R_{TH}	1027 [K/W]

is observed, with the Mextram model simulations (solid curves) showing good fits with the measurements (symbols) over the non-avalanche and weak avalanche regions (see Fig. 2.2 for the description of these regions). In Fig. 2.15(b) and Fig. 2.15(c), the decrease of base-emitter voltage (V_{BE}) and an increase of the collector current (I_C), respectively, with V_{CB} is observed. Other than a small offset at high V_{CB} and $|I_E|$ (over the strong avalanche region), the decrease of V_{BE} and increase of I_C with V_{CB} is captured quite well by the Mextram simulations. For low $|I_E|$, there is overlap between the Mextram model simulations and the measurements. We note that, for self-consistency check, the accuracy of the Mextram model simulations with respect to the measurements is of no relevancy, what matters is having a practical set of model parameters, which is then used to generate the required input for the extraction method.

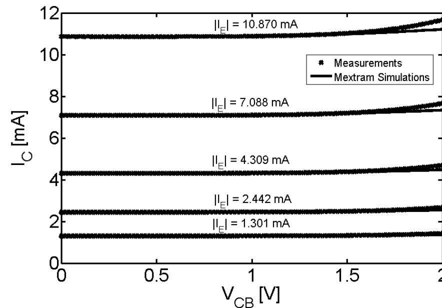
In order to carry out a self consistency check of our extraction method, the Mextram model simulation results are used as the input data of the extraction method instead of the measurement data [48] used earlier. The self consistency check steps used in this section are summarized in Fig. 2.16. The idea is that the *exact* values of R_B and R_{TH} are known before hand (from the operating point information (OP-info) [2] of the Mextram model) and we want to analyze (or determine) how accurate our extraction method can



(a)



(b)



(c)

Figure 2.15: Measurements (symbols) and Mextram model simulations (solid curves) of base current I_B (*a*), base-emitter voltage V_{BE} (*b*), and collector current I_C (*c*) as a function of the base-collector voltage (V_{CB}). In (*a*), over the non-avalanche and the weak avalanche regions, the Mextram model simulations describe quite well the base current. In this region there is an overlap between the measurements and the Mextram model simulations. In (*b*) and (*c*), a decrease of the base-emitter voltage V_{BE} and an increase of the collector current I_C with V_{CB} is well described with the Mextram model simulations except an offset at high $|I_E|$ and V_{CB} (strong avalanche region). For low $|I_E|$, the Mextram simulations overlap with the measurements. These simulation results are used as input data for our extraction method during the self consistency check.

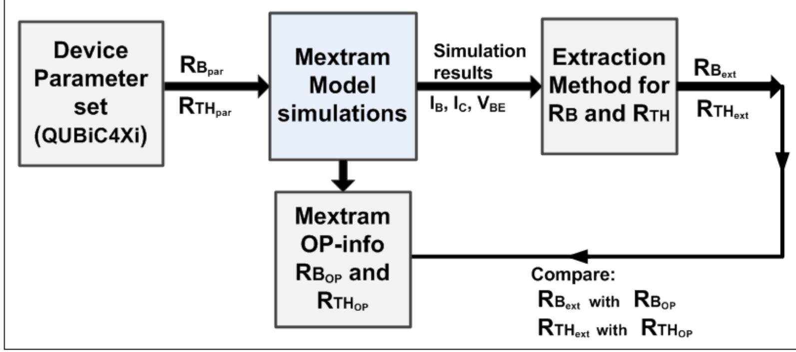


Figure 2.16: Summarized steps used for self consistency check of our extraction method for R_B and R_{TH}. Here $R_{\mathrm{B_{par}}}$ and $R_{\mathrm{TH_{par}}}$ are the relevant model parameters presented in Table 2.2, $R_{\mathrm{B_{OP}}}$ and $R_{\mathrm{TH_{OP}}}$ are the exact values from the Mextram model operating point information (OP-info), and $R_{\mathrm{B_{ext}}}$ and $R_{\mathrm{TH_{ext}}}$ are the extracted values using our method.

approximate these known values. This will be done by comparing the extracted values of R_B and R_{TH} from our method, with the known input *exact* values. In the Mextram model and other compact models [26, 51], the operating point information, is a list of quantities that describe the internal state of the transistor [2]. This information, once provided by the circuit simulator, can help the circuit designer to understand the behavior of the transistor and the circuit as a whole. In Fig. 2.16, $R_{\mathrm{B_{par}}}$ and $R_{\mathrm{TH_{par}}}$ are the input model parameters as presented in Table 2.2, $R_{\mathrm{B_{ext}}}$ and $R_{\mathrm{TH_{ext}}}$ are the extracted values from our method, and $R_{\mathrm{B_{OP}}}$ and $R_{\mathrm{TH_{OP}}}$ are the OP-info from the Mextram model. We note that these OP-info values take the temperature scaling of the model parameters into account, so they differ from the original input model parameter values. The goal of this self-consistency check is to study how accurate the extracted values using our method, i.e., $R_{\mathrm{B_{par}}}$ and $R_{\mathrm{TH_{par}}}$ approximate those from the OP-info, i.e., $R_{\mathrm{B_{OP}}}$ and $R_{\mathrm{TH_{OP}}}$. This will be discussed in detail (for each quantity) in the next subsections.

2.5.1 Self consistency check for R_{TH}

To carry out a self consistency check of our extraction method for R_{TH}, we compare the extracted value using our method presented earlier, with the exact value from the Mextram OP-info which takes into account the temperature scaling of R_{TH}. In the Mextram model, the OP-info exact value of R_{TH} is computed basing on the fact that the thermal conductivity decreases with temperature, which implies that the thermal resistance will increase with temperature; this results into the employed Mextram model temperature scaling rule [20]

$$R_{TH,T_{amb}} = R_{TH} \cdot \left(\frac{T_{\mathrm{amb}}}{T_{\mathrm{ref}}} \right)^{A_{\mathrm{TH}}}. \tag{2.38}$$

In this expression, T_{amb} is the ambient temperature, T_{ref} is the reference temperature, and A_{TH} the temperature coefficient of the thermal conductivity, a material coefficient.

Note that the temperature dependence of R_{TH} is given in terms of the T_{amb}, a chosen temperature in the Mextram compact model in order to avoid non-linearity. But in reality, R_{TH} depends on the temperature distribution due to the distributed dissipated power in the device [26]. In the simulations we took $R_{TH} = 1027$ K/W (see Table 2.2), and $A_{TH} = 1.702$. In Table 2.3, we compare the extracted value of R_{TH} with the exact value (Mextram OP-info value), and the value extracted using the method proposed by Vanhoucke and Hurkx [27]. From Table 2.3, we observe that our extraction method yields a

Table 2.3: Comparison of our extracted value of R_{TH} with the exact value and that from the method in [27]

	Exact value	Extracted using our method	Extracted using method in [9]
R_{TH} [K/W]	1048	1039	1212
Percentage error [%]	-	0.86	15.65

smaller percentage error ($\approx 0.9\%$) in the value for R_{TH} in comparison to the method proposed by Vanhoucke and Hurkx ($\approx 16.0\%$). These results provide an explanation for the deviations in the extracted values of R_{TH} observed earlier in Fig. 2.11, when the two extraction methods were applied to the same measurement data. In addition, these results confirm the importance of consistently accounting for the influence of the collector-base Early effect on the internal base-emitter voltage V_{BEi}. We recall that the extraction method proposed by Vanhoucke and Hurkx [27] ignores the influence of collector-base Early effect on V_{BEi}, which as demonstrated in Table 2.3 leads to a larger extracted value for the thermal resistance R_{TH} [48]. In general, the results in Table 2.3 demonstrate the key modifications provided by our new proposed extraction method, i.e., by consistently accounting for both self-heating and Early effect, we arrive at an accurate extraction of thermal resistance R_{TH}.

2.5.2 Self consistency check for R_B

In order to carry out a self consistency check of our extraction method for the total base resistance R_B, we need the exact value (Mextram OP-info value). This exact value differs from the model parameter value (presented in Table 2.2) due to the variations in the variable part of the base resistance (R_{Bv}) [2] with base-emitter voltage (V_{BE}), and variations in the constant part of the base resistance (R_{Bc}) due to internal transistor self-heating [2, 43, 48]. We note that in the Mextram model, the variable part of the base resistance is a modulated resistance accounting for the bias dependency due to current crowding and reversal of base current [34]. For the model simulations, realistic practical values of $R_{Bv} = 11.47\ \Omega$ and $R_{Bc} = 8.53\ \Omega$ are employed, these contribute to the total model parameter $R_B = 20\ \Omega$ presented earlier in Table 2.2. The temperature dependence of R_{Bc} according to the Mextram model [2] is given by

$$R_{BcT} = R_{Bc} \cdot \left(\frac{T}{T_{\text{ref}}}\right)^{A_{\text{ex}}}, \tag{2.39}$$

2

and that of R_{Bv} is described by

$$R_{BvT} = R_{Bv} \cdot \left(\frac{T}{T_{\text{ref}}}\right)^{A_B - A_{Q_{B0}}}, \tag{2.40}$$

where A_B and A_{ex} are temperature coefficients of resistivity (model parameters) of the base and extrinsic base, respectively, and T is the device temperature. $A_{Q_{B0}}$, also a model parameter, is the temperature coefficients of the zero-bias base charge Q_{B0} [2, 50]. In Fig. 2.17, we present the constant R_{BcT} (Fig. 2.17(a)) and variable R_{BvT} (Fig. 2.17(b)) parts (of the OP-info value) of R_B (i.e., $R_{\text{B}_{\text{OP}}}$) as a function of the base-emitter voltage (V_{BE}). In Fig. 2.17(c), we present results for the total exact base resistance $R_B = R_{BcT} + R_{BvT}$, which shows a decrease for low and medium values of V_{BE} (caused by a decrease in R_{BvT} due to base bias variation), and then an increase at high values of V_{BE} (caused by an increase in R_{BcT} due to self-heating).

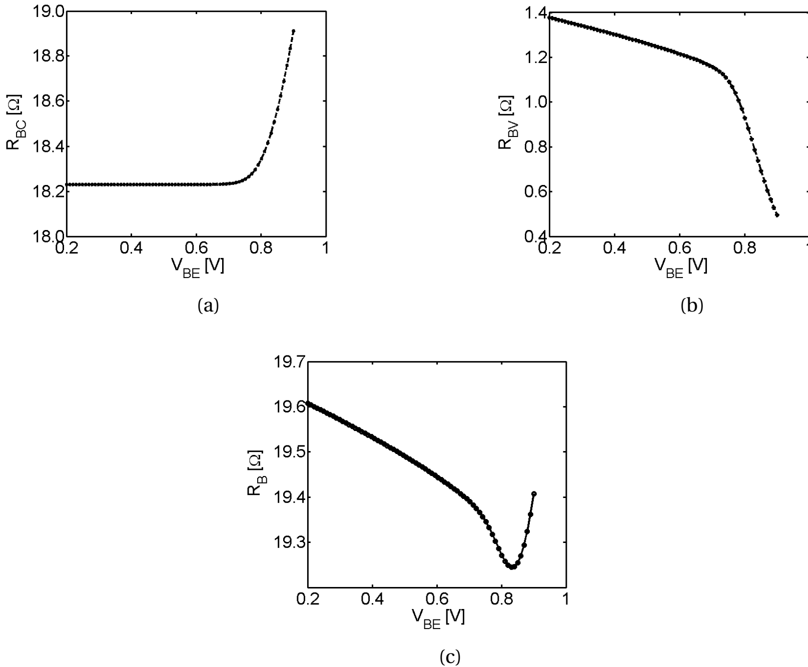


Figure 2.17: Exact value (from operation point information of the Mextram model) of the constant part (*a*), variable part (*b*), and total (*c*) R_B as a function of V_{BE}. The decrease in total R_B over the low and moderate V_{BE} is due to the decrease in R_{BV} caused by variation in the base diffusion charge with V_{BE} and the increase over high V_{BE} is due to the increase of R_{BC} caused by self-heating.

A comparison of the extracted total base resistance R_B using our method with the exact value from the Mextram OP-info is presented in Fig. 2.18, as a function of the external base-emitter voltage (V_{BE}) [48]. From the top figure of Fig. 2.18, we observe that

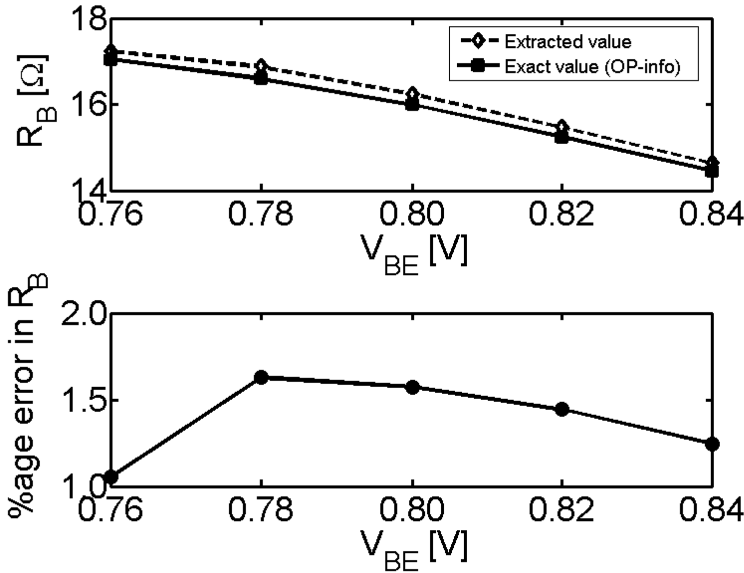


Figure 2.18: Exact value (from operation point information of the Mextram model) and extracted value of R_B as a function of V_{BE} (top figure). From the top figure, we observe a gentle decrease in both extracted and exact R_B values with V_{BE}. The bottom figure shows the percentage error in the extracted R_B as a function of V_{BE}. This error is within a 2.0% error margin. Here, the model parameter R_B = 20 Ω (with realistically R_{BV} = 11.47 Ω and R_{BC} = 8.53 Ω).

both the exact (solid curve with closed squares) and extracted (dashed curve with open diamonds) values of R_B decreases gently with increasing V_{BE}, with a consistent deviation in between them. Here our proposed extraction method, yields consistently (and slightly) larger R_B values with respect to the exact values. The offset between the extracted and exact values of R_B yields a percentage error in the extracted R_B that is less than 2.0% (as can be seen from the bottom plot in Fig. 2.18).

2.5.3 Summary

From the self consistency check for thermal resistance R_{TH} and base resistance R_B discussed above, we see that our extraction method yield results that are consistent with exact values from the Mextram model operating point information [2]. From the self consistency check for R_{TH}, we observed that our extraction method yields more accurate results (that are within an error margin of 1.0%) than the method proposed by Vanhoucke and Hurkx [27] (with an error margin of around 16.0%). The method proposed in [27] yields larger values of the extracted thermal resistance (as seen earlier in Fig. 2.11), since it ignores the influence of collector-base Early effect on the internal base-emit-

ter voltage V_{BEi}. This clearly shows that with respect to the current generation of SiGe HBT's, it is significant to consistently account for the effect of the collector-base Early effect and self-heating on V_{BEi}, as addressed by our proposed extraction method. These results clearly demonstrate the impact of the modifications provided by the extraction method presented in this work, i.e., a more accurate extracted R_{TH} value. For R_B, our method yield results that are within an error margin of 2.0% over the whole chosen bias-range. Therefore, our proposed extraction method for R_{TH} and R_B is self-consistent in both R_{TH} and R_B [48]; hence, it can be used to accurately extract both R_{TH} and R_B for Si/SiGe bipolar transistors (from DC-measurements only).

2.6 Summary

In this chapter, we presented a method to extract both the base resistance (R_B) and the thermal resistance (R_{TH}) of Si and SiGe bipolar transistors. Our approach extends the category of methods [27, 30] that utilizes fixed emitter current conditions to distinguish between emitter (R_E) and base (R_B) resistance. These two series resistances are known to pose a fundamental difficulty in DC-methods that explore the normal forward bias regime [11, 12]. The key idea is to measure the differentials in the external base-emitter voltage (V_{BE}) while forcing a constant emitter current. This ensures that the voltage drop across the emitter resistance is constant, hence its variation is zero for constant R_E. Small (local) variations in bias and temperature space are considered [31]. Then the resulting expression for the differentials in V_{BE} would depend on the differentials in the internal base-emitter voltage (V_{BEi}) and in the base current (I_B). Since methods in this category employs the collector-base avalanche current to vary the base current (I_B) independently of the emitter current (I_E), then the differentials in V_{BEi} with respect to the collector-base voltage (V_{CB}), would strongly be influenced by Early effect and self-heating. Here self-heating is due to dissipated power (P_{diss}) in the device, which leads to an increase in the device (junction) temperature (T_j). In contrast to previous methods in this category i.e., in [27, 30], for the extraction method in this work, we consistently account for the influence of both self-heating and Early effect on V_{BEi}. This is achieved by employing the model for the main transistor current (I_N) [2], which depends on the junction temperature (T_j), Early voltage (V_A), and V_{BEi}. From this relation together with weak avalanche current conditions, we derive an approximate expression for the differentials in V_{BEi}, which is then used to determine the approximate expression for the differentials in V_{BE}. Also in this process, we learn how to measure the sensitivity (α_T) of V_{BE} to temperature. The final expression is achieved by making a few further assumptions which restricts the method to the weak avalanche regime and to devices with relatively high Early voltages; these approximations play an essential role in the definition of the extraction method. Using this final expression, the extraction procedure for the R_B and R_{TH} from the DC-measurements is defined. Since the employed assumptions have observable consequences, careful selection of the input data is needed in order to achieve accurate extracted values of R_B and R_{TH}.

This derived extraction method was demonstrated on DC-measurements corresponding to a RF QUBiC4Xi SiGe-HBT [41]. By employing this measured data, we presented three different approaches that can be used to the determine the region where the input

data can be selected from. This region can be viewed as a regime of applicability of the method, or alternatively an experimental check of the validity of the presupposed underlying assumptions. Among these methods, the method that employs the contour plots in combination with 2D-plots [46] is more robust, since it determines the exact boundaries of the required extraction region as a function of the employed constant emitter currents. With the input data taken from the appropriate regime of applicability, the extracted base resistance R_B is determined from the intercept with the vertical axis of the plot of $-dV_{BE}/dI_C$ as a function of $(V_{CB} + V_A^{\mathrm{eff}})$, where V_A^{eff} is determined directly from the same measured data [31]. The corresponding slope together with the predetermined value for α_T are used to determine the extracted thermal resistance R_{TH}. These results are also used to verify the various presupposed underlying assumptions.

2

We compared extracted results of R_B and R_{TH} from our method, with corresponding results from other DC- and AC-methods. The extracted R_B and R_{TH} from our method, are compared with results from the DC-method proposed by Vanhoucke and Hurkx [27]; which determines both R_B and R_{TH} by considering the change of V_{BEi} due to *self-heating*, but ignores the influence of the Early effect on V_{BEi}. A comparison to the theoretical results [1, 36, 37, 43] suggest that this neglect of Early effect, can lead to an underestimation of the device collector current and thus the dissipated power; which leads to an overestimation of R_{TH}. This was indeed observed as the method in [27] yielded higher values for R_{TH} in comparison to our modified extraction method, for a fixed sequence of constant emitter currents. This is the key result (i.e., different values of R_{TH}) provided by our extraction method in relation to the method proposed in [27]. On the other hand, these two methods yielded the same extracted R_B, since they both take the influence of self-heating on V_{BEi} into account.

Our extraction results for R_{TH} were also compared with corresponding results from another DC-method proposed by Vanhoucke et al [25]. This method is a revised version of the method proposed in [3] for extraction of R_{TH} from DC-measurements similar to those employed in our earlier work, though now with the self-heating effect taken into account during the measurements for α_T. A deviation of $\approx 9.0\%$ was observed between the extracted R_{TH} from our method and that from the method proposed in [25]. This could be related to the unconvincing derivation of the method in [25], since two different base-emitter voltages were equated to one another, thus the correctness of this method is somehow questionable.

We compared our extracted total base resistance R_B with results from the standard two port method and the standard circle impedance method that are both based on the small-signal high frequency admittance measurements. In [28] and [29] it was experimentally shown that the true value of the base resistance lies between the value extracted using the standard two port method and that of standard circle impedance method. This was indeed the case for our method, since our extracted R_B fell nicely between the value from these two AC-methods; this is supportive of the validity of our proposed method.

The extracted total base resistance R_B from our method was also compared with results determined independently from the noise measurements, on the same devices with different geometry. Results from noise measurements introduce a fully independent ingredient, namely the employment of intrinsic distributed noise sources as internal signal sources, so as to probe the intervals of a device. The extraction of R_B from the noise mea-

surements is generally a complex process that requires expensive AC-measurements, extensive de-embedding of the parasitics, and application of a small-signal compact bipolar transistor model. The deviation between the extracted R_B from these two methods is ≈ 21.0% for all employed devices [48]. Given the complexity of the noise measurements method, this error margin is considered quite acceptable within the semiconductor device physics domain, indeed the results presented here can be taken as mutually supportive for both confronted approaches.

To enable quantification of the level of accuracy of the extraction method we presented in this chapter, we carried out a self-consistency check, where we used Mextram model simulation results as input data for the extraction method. Employing the temperature scaling rules of the Mextram model, effect of self-heating on the relevant model parameters is taken into account. The results corresponding to the operating point information (OP-info) of the Mextram model [2] are taken as the known exact values, and they are compared with extracted results from our method and those from the method proposed in [27]. A comparison of the exact value of R_{TH} with results extracted from our method, yielded an error margin of ≈ 1.0%, while with that from the method in [27] yielded an error margin of ≈ 16.0%. These results explain the systematic difference observed in the extracted R_{TH} values when both extraction methods were applied on the same measured data. It also demonstrates the importance to accounting for the influence of Early effect on V_{BEi}, which was ignored in the method proposed in [27]. In other words, these results show that by accounting for the influence of both self-heating and Early effect on V_{BEi}, a more accurate extracted value of R_{TH} is obtained than when only self-heating is accounted for. The observed deviation between the exact value of R_B and that extracted from our extraction method is less than 2.0%. These results show that our extraction method is self-consistent in both extracted R_B and R_{TH}, with sufficiently low error margins.

In general, the extraction method presented in this chapter is local in bias- and temperature space: measured data involved for a single parameter extraction includes only small variations in emitter-base bias conditions and temperature. With respect to collector-base biasing, only a traverse of the *weak* avalanche regime is involved. This enables parameter extraction as a function of bias and temperature. The method is also local in device space: it is applicable to individual transistors; no additional (DC-) test structures are needed. This makes the method cost efficient, e.g., with respect to parameter extraction as a function of geometry.

2

References

[1] P. Ashburn, *SiGe Heterojunction Bipolar Transistors* (John Wiley and Sons Ltd, 2003).

[2] J. C. J. Paasschens, W. J. Kloosterman, and R. van der Toorn, *Model Derivation of Mextram 504, The physics behind the model,* Tech. Rep. Technical note TN-2004/01010 (Philips Research Eindhoven, 2005).

[3] J.-S. Rieh, D. Greenberg, B. Jagannathan, and S. Subbanna, *Measurement and modeling of thermal resistance of high speed SiGe heterojunction bipolar transistors,* in *Proc. Topical Meeting on Silicon Monolithic Integrated Circuits in RF Systems* (2001) pp. 110–113.

[4] W. M. C. Sansen and R. G. Meyer, *Characterization and measurement of the base and emitter resistance of bipolar transistors,* IEEE J. Solid-State Circuits **6**, 1969 (2005).

[5] T. H. Ning and D. D. Tang, *Method for determining the emitter and base resistance of bipolar transistors,* IEEE Trans. Electron Devices **31**, 409 (1984).

[6] M. Linder, F. Ingvarson, K. Jeppson, J. Grahn, and S. Z. M. Östling, *Extraction of emitter and base series resistances of bipolar transistors from a single dc measurement,* IEEE Trans. Electron Devices (2000).

[7] T. Zimmer *et al.*, *Method of determining the effective base resistance of bipolar transistors,* in *Proc. Bipolar Circuits and Technology Meeting* (1996) pp. 122–125.

[8] F. Hébert and D. Roulston, *Base resistances of bipolar transistors from layout details including two dimensional effects at low currents and low frequencies,* IEEE Solid-State Electron **31**, 283 (1988).

[9] J. Weng, J. Holz, and T. Meister, *New method of determine the base resistances of bipolar transistors,* IEEE Electron Device Lett **13**, 158 (1992).

[10] G. Hong, J. Gong, and S. Lu, *A general method of characterizing base resistance of bipolar junction transistors,* Jpn J. Appl. Phys **39**, 428 (2000).

[11] J. Steigerwald and P. Humphries, *TCAD assisted reflection on parameter extraction for compact modeling,* in *Proc. Bipolar Circuits and Technology Meeting* (2010) pp. 245–252.

[12] S. Maas, *Why I hate base resistance,* in *Microw. Mag* (2004) pp. 54–60.

[13] D. Walkey *et al.*, *Extraction and modelling of thermal behavior in trech isolated bipolar structures,* in *in Proc. Bipolar Circuits and Technology Meeting* (1999) pp. 97–100.

[14] M. Pfost, V. Kubrak, and P. Brenner, *A practical method to extract the thermal resistance for heterojunction bipolar transistors,* in *Proc. Bipolar Circuits and Technology Meeting* (IEEE, 2003) pp. 335–338.

[15] P. S. Marsh, *Direct extraction technique to derive the junction temperature of hbts under high self-heating,* IEEE Trans. Electron Devices **47**, 288 (2000).

[16] D. E. Dawson, A. K. Gupta, and M. L. Salib, *CW measurement of HBTs thermal resistance under high self-heating,* IEEE Trans. Electron Devices **39**, 2235 (1992).

[17] F. F. Oettinger *et al.*, *Thermal characterization of power transistors,* IEEE Trans. Electron Devices **23**, 831 (1976).

[18] M. G. Alderstein *et al.*, *Thermal resistance measurements for AlGaAS/GaAs heterojunction bipolar transistors,* IEEE Trans. Electron Devices **38**, 1533 (1991).

[19] W. D. van Noort and R. Dekker, *Thermal resistance of (H)BTs on bulk, SOI and glass,* in *Proc. IEEE* (2003) pp. 129–132.

[20] J. C. J. Paasschens, S. Harmsma, and R. van der Toorn, *Dependence of thermal resistance on ambient and actual temperature,* in *Proc. Bipolar/BiCMOS Circuits and Technology Meeting* (2004) pp. 96–99.

[21] P. Palestri, A. Pacelli, and M. Mastrapasqua, *Thermal resistance in $Si_{1-x}Ge_x$ hbts on bulk-Si and SOI substrates,* in *Proc. Bipolar Circuits and Technology Meeting* (2001) pp. 98–101.

[22] D. Williams and P. Tasker, *Thermal resistance extraction techniques using DC I-V data for HBT transistors,* in *High Frequency Postgraduate Student Colloquium* (2000) pp. 71–75.

[23] H. Hanington, C. Chang, P. Zampardi, and P. Asbeck, *Thermal effects in HBT emitter resistance extraction,* Electron. Lett. **32**, 1515 (1996).

[24] S. M. Sze, *Physics of Semiconductor Devices*, 2nd ed. (John Wiley and Sons, INC, 1981).

[25] T. Vanhoucke, H. M. J. Boots, and W. D. V. Noort, *Revised method for extraction of the thermal resistance applied to bulk and SOI SiGe HBTs,* IEEE Electron. Device. Lett Electron Devices **25**, 150 (2004).

[26] M. Reisch, *High-Frequency Bipolar Transistors* (Springer-Verlag, 2003).

[27] T. Vanhoucke and G. A. M. Hurkx, *Simultaneous extraction of the base and thermal resistances of bipolar transistors,* IEEE Transactions on Electron Devices **52**, 1887 (2005).

[28] T. Nakadai and K. Hashimoto, *Measuring the base resistance of bipolar transistors,* in *Proc. Bipolar Circuits and Technology Meeting* (1991) pp. 200–203.

[29] W. J. Kloosterman, J. C. J. Paasschens, and D. B. M. Klaasen, *Improved extraction of base and emitter resistance from small signal high frequency admittance measurements,* in *Proc. BCTM* (1999) pp. 93–96.

[30] G. Verzellesi, R. Turetta, P. Pavan, A. Collini, A. Chantre, A. Marty, and C. Canali, *Extraction of dc base parasitic resistance of bipolar transistors based on impact-ionization-induced base current reversal,* IEEE Electron Device Letters **14**, 431 (1993).

[31] R. Setekera, R. van der Toorn, and W. Kloosterman, *Local extraction of base and thermal resistance of bipolar transistors,* in *Proc. Bipolar Circuits and Technology Meeting* (IEEE, 2013) pp. 21–24.

[32] W. J. Kloosterman and H. C. de Graaff, *Avalanche multiplication in a compact bipolar transistor model for circuit simulation,* in *Proc. BCTM* (1988) pp. 103–106.

[33] W. J. Kloosterman, J. C. J. Paasschens, and R. J. Havens, *A comprehensive bipolar avalanche multiplication compact model for circuit simulation,* in *Proc. BCTM* (2000) pp. 172–175.

[34] R. van der Toorn, J. C. J. Paasschens, W. J. Kloosterman, and H. C. de Graaff, *Compact Modeling: Principles, techniques, and applications,* (Springer-Verlag, 2010) Chap. 7.

[35] N. Rinaldi and V. d'Alessandro, *Theory of electrothermal behavior of bipolar transistors: Part III- Impact-ionization,* IEEE Trans. Electron Devices **53**, 1683 (2006).

[36] H. C. de Graaff and F. M. Klaassen, *Compact transistor modelling for circuit design* (Springer-Verlag, 1990).

[37] P. A. H. Hart, ed., *Bipolar and Bipolar-MOS Integration* (Elsevier, 1994).

[38] P. Ashburn, *Design and Realization of Bipolar Transistors*, edited by D. V. Morgan and H. R. Grubin (John Wiley and Sons, Inc, 1988).

[39] H. C. de Graaf and W. J. Kloosterman, *The Mextram Transistor Model*, Unclassified Report NL-UR 006/94 (Koninklijke Philips Electronics N. V., Nat. lab, 1994).

[40] H. C. de Graaf, W. J. Kloosterman, J. M. Geelen, and M. A. M. Koolen, *Compact bipolar transistor model for CACD with accurate description of collector behavior purposes,* in *Proc. BCTM*.

[41] W. D. Noort, A. Rodriguez, H. Sun, F. Zaato, N. Zhang, T. Nesheiwat, F. Neuilly, J. Melai, and E. Hijzen, *BiCMOS technology improvements for microwave application,* in *Proc. BCTM* (2008) pp. 93–96.

[42] L. L. Spina, V. d'Alessandro, S. Russo, N. Rinaldi, and L. K. Nanver, *Influence of concurrent electrothermal and avalanche effects on the safe operating area of multifinger bipolar transistors,* IEEE Trans. Electron Devices **56**, 483 (2009).

[43] S. M. Sze and K. N. Kwok, *Physics of Semiconductor Devices*, 3rd ed. (Wiley, 2006).

[44] S. Jeffrey, *Smoothing methods in statistics,* Spring (1996).

[45] V. Blobel, *Fittings without a parametrization*, Tech. Rep. (University of Hamburg, 2005).

[46] R. Setekera, L. F. Tiemeijer, W. J. Kloosterman, and R. van der Toorn, *Analysis of the local extraction method of base and thermal resistance of bipolar transistors,* in *Proc. Bipolar Circuits and Technology Meeting* (IEEE, 2014) pp. 215–218.

[47] D. Leenaerts, *mmwave activities within NXP,* Compus Technology Seminar (2012).

[48] R. Setekera, L. F. Tiemeijer, and R. van der Toorn, *Verification of the simultaneous local extraction method of base and thermal resistance of bipolar transistors,* WASET International Journal of Electrical, Computer, Energetic, Electronic and Communication Engineering **08**, 1456 (2014).

[49] G. Niu *et al.*, *Noise-gain tradeoff in RF SiGe HBT's,* Solid-State Elec. **46**, 1445 (2002).

[50] J. C. J. Paasschens, W. J. Kloosterman, and R. J. Havens, *Parameter extraction for the Bipolar Transistor Model Mextram Level 504,* Unclassified report NL- UR 2001/801 (Philips Research Eindhoven, 2001).

[51] C. McAndrew *et al.*, $VBIC95$ *the vertical bipolar inter-company model,* IEEE J. Solid-State Circuits **31**, 476 (1996).

3

Compact model for the non-local avalanche effect in Si/SiGe bipolar transistors

In this chapter, we derive a physics based compact model for non-local avalanche effects in bipolar transistors. This model is based on the approximate energy balance equation and Chynoweth's empirical law for impact-ionization. With this physical basis, we get an expression for the impact-ionization rate as a function of the electron temperature; which turns out to be a sharply peaked function about the maximum position. The approximate integral of such functions is developed on basis of a Taylor series expansion of the integral in terms of the relative width of the function peak. This is employed to determine an approximate expression for the multiplication factor (i.e., the integral of the impact-ionization rate over the collector-epilayer width). The product of the multiplication factor with the collector current results into the non-local avalanche current model. This non-local avalanche model is explicit and in terms of elementary functions, so it can readily be implemented in existing complete compact bipolar transistor models. The formulation has essentially only two new model parameters, both well known material constants: the energy relaxation length and its corresponding temperature coefficient. An example of how this non-local avalanche compact model can be implemented in a complete compact model for bipolar transistors will be presented.

This chapter is based on the work submitted in the article:
R. Setekera and R. van der Toorn, *Solid-State Electronics*, manuscript submitted (July 2015).

3.1 Introduction

In order to meet the globally increasing requirements of the emerging RF applications, in modern bipolar transistors, breakdown voltages are aggressively traded off against speed performance. Fortunately, RF SiGe hetero-junction bipolar transistors (HBTs) explore highly doped collector-base junctions in which the electric field $E(x)$ is sufficiently peaked to relax this trade-off, to some extent, by enhancing the breakdown voltages through the so-called non-local avalanche effect [1, 2]. Under such non-local effect, the electron temperature $T_e(x)$ [1] lags $E(x)$ due to the finite energy relaxation length of the electrons [2, 3]. Hence, the commonly used local electric field model [4, 5] over estimates the impact-ionization rate α and the avalanche current, and underestimates the breakdown voltage [2, 3]. This was demonstrated earlier in Fig. 1.3(b), where the measured (symbols) and Mextram's [6, 7] standard *local* (electric field based) avalanche model [8–10] simulated (solid curves) results for the avalanche characteristics were presented as a function of collector-base voltage (V_{CB}) and ambient temperature. From these results, it was clear that the standard local avalanche model overestimates avalanche, and thus underestimates the breakdown voltage. In addition, it was observed that the local avalanche model is insufficient in fitting the measured avalanche characteristics as a function of V_{CB} and ambient temperature, even though appropriate temperature scaling of Cynoweth's empirical law was applied [10, 11].

In practical applications, circuit designers are forced to exploit transistor operation above the common-emitter breakdown voltage with open base (BV_{CE0}) limits, and thus the need for simulation capabilities with respect to transistor characteristics beyond BV_{CE0}. At the same time, either because of application determined requirements (e.g., tolerances with respect to possibly extreme environmental conditions, such as in automotive applications) or because of significant self-heating in SiGe HBT applications, modeling of temperature dependencies with respect to all transistor characteristics requires undiminished attention. Hence, the need [12] for adequate compact model formulations of non-local avalanche, and the motivation to include in these adequate modeling of temperature dependencies [13].

To address precisely the combination of these needs, in this chapter we shall derive a *physics based* compact model for non-local avalanche, including its temperature dependence. We recall that, having a physics based compact model is vital for scalability (over temperature and geometry) and statistical modeling. This is also important for model verification of the employed physics, since the extracted model parameters can be related to the available device physics literature. Our final non-local avalanche compact model formulation is explicit and in terms of elementary functions, so it can be readily incorporated in standard compact bipolar transistor models such as Mextram [6, 7]; in this our solution is distinct from computationally more demanding approaches such as those presented in e.g., [2, 3].

Our modeling approach as described in Section 3.2, results into an expression for the non-local avalanche impact-ionization rate α as a function of the electron temperature. The integral of this expression over the epilayer width yields an approximate expression for the multiplication factor. To enable implementation of this non-local avalanche model in a complete compact transistor model such as Mextram [6], the model is modi-

fied further in order to take into account cases where the model may become subject to evaluation for parameter and bias values outside its range of physical validity, as will be discussed in Section 3.3. In the last Section 3.4, we will give a summary of the compact model developments.

3.2 Compact model for non-local avalanche current

In this section, we will derive a physics based compact model formulation for non-local avalanche effects in bipolar transistors. This model is based on the approximate energy balance equation (Subsection 3.2.1) and Chynoweth's empirical law, which are used to determine the non-local impact-ionization rate (Subsection 3.2.2) as a function of the electron temperature. We will extensively derive (Subsections 3.2.4 and 3.2.5) an approximate expression for the integral of the impact-ionization rate as a function of the electron temperature over the collector width (i.e., multiplication factor), for a given local electric field distribution. This expression for the electron temperature dependent multiplication factor is then used to determine the non-local avalanche current. We will address the temperature dependence of the new introduced model parameter, i.e., the energy relaxation length λ_e (Subsection 3.2.6), so as to support a direct comparison of the model parameters to the well known material coefficients.

3.2.1 Derivation of the electron temperature

When the electrons in the conduction band of a semiconductor device carry a drift current under the influence of an electric field $E(x)$ (throughout this work, we shall use a sign convention $E(x) < 0$), the elevation $\Delta T_e(x) = (T_e(x) - T_0)$ of their temperature $T_e(x)$ with respect to the equilibrium lattice temperature T_0 is governed by the following approximate energy balance equation [1, 14, 15]

$$\frac{d\Delta T_e(x)}{dx} + \frac{1}{\lambda_e}\Delta T_e(x) + \frac{2q_e}{5k_B}E(x) = 0\,. \tag{3.1}$$

In this expression, k_B is the Boltzmann's constant, q_e is the elementary charge, and λ_e the energy relaxation length. Under the assumption of constant λ_e [1], the solution of the differential equation (3.1) is given by

$$\Delta T_e(x) = -\frac{2q_e}{5k_B}\int_0^x \exp\left(\frac{u-x}{\lambda_e}\right)E(u)\,du\,. \tag{3.2}$$

From (3.2), we see that in order to determine $\Delta T_e(x)$, the local electric field distribution $E(x)$ have to be defined. For spatially homogenous collector doping concentration N_{epi}, $E(x)$ throughout the depletion region in the collector of the bipolar transistor takes the form [2, 6, 16]

$$E(x) = ax + E_0\,, \tag{3.3}$$

where E_0 is the electric field at position $x = 0$, and a is the gradient of the local electric field that depends on the collector doping concentration N_{epi} and the collector current I_{epi} as [7]

$$a = \frac{dE(x)}{dx} = \frac{q_e}{\epsilon}N_{\text{epi}}\left(1 - \frac{I_{\text{epi}}}{I_{\text{hc}}}\right), \tag{3.4}$$

which is based on Poisson's equation [17–19]. In relation (3.4), ϵ is the permittivity and $I_{hc} = q_e A_{em} N_{epi} v_{sat}$ is the hot carrier current [6, 20], with A_{em} being the emitter area and v_{sat} the carrier saturation drift velocity [21]. Substituting for $E(x)$ from (3.3) into (3.2), and applying the initial condition $T_e(x) = T_0$, (3.2) then yields

$$\Delta T_e(x) = -\frac{2q_e\lambda_e}{5k_B}\left[ax - (a\lambda_e - E_0)\left(1 - \exp\left(-\frac{x}{\lambda_e}\right)\right)\right]. \tag{3.5}$$

This expression defines the spatial dependence of the electron carrier temperature in terms of the local electric field gradient (a), the field ($E_0 = E(x = 0)$), and the energy relaxation length (λ_e) of the electrons. In the next subsection, this expression will be used to determine the non-local impact-ionization rate as a function of the electron temperature.

If we take the first derivative of expression (3.5) with respect to the spatial coordinate x, we obtain

$$\frac{d\Delta T_e(x)}{dx} = -\frac{2q_e\lambda_e}{5k_B}\left[a - (a\lambda_e - E_0)\left(\frac{1}{\lambda_e}\exp\left(-\frac{x}{\lambda_e}\right)\right)\right]. \tag{3.6}$$

Now equating this gradient of ΔT_e, i.e., relation (3.6) to zero, the position (x_{max}) where the extreme value of the electron temperature (3.5) is located can be found as

$$x_{max} = -\lambda_e \ln\left[\frac{a\lambda_e}{a\lambda_e - E_0}\right] = \lambda_e \ln\left[1 - \frac{E_0}{a\lambda_e}\right]. \tag{3.7}$$

This expression for x_{max}, will be used later in this section, to determine the approximate expression for the multiplication factor.

3.2.2 Non-local impact-ionization rate

Due to the (bias dependent) high electric field $E(x)$ in the collector-base space charge region, the incident electrons (from the emitter) gain high kinetic energy to knock a bound electron out of its bound state (in the valence band) and promote it to a state in a conduction band, thus creating an electron-hole pair [8, 18]. In this process (impact-ionization), the generated charge carrier pairs are separated by the electric field. The rate of carrier generation due to the local electric field induced impact-ionization has been subject of an extensive literature, e.g., [5, 22–24], but commonly used relation is the Chynoweth's empirical law [11], which provides an empirical relation for the ionization coefficient

$$\alpha = A_n \exp\left(-\frac{B_n}{|E|}\right). \tag{3.8}$$

In this expression, A_n and B_n are the original coefficients of Chynoweth's empirical law; the values of these material parameters are experimentally well-established [8–10]. The quantity $|E|$ denotes the electric field strength, which can be considered to have been spatially constant in the experiments, over sufficiently long distances such as to guarantee that the electrons reached a stationary value of thermal kinetic energy. Therefore, we may assume that the electron temperature ΔT_e, during the experiments performed to determine the avalanche coefficients in (3.8), took a stationary value $\Delta T_e^{(S)}$. This value

$$\Delta T_e^{(S)} = \frac{2q_e\lambda_e}{5k_B}|E|^{(S)} \tag{3.9}$$

follows from the approximated energy balance equation (3.1) in combination with the steady state condition (i.e., $d\Delta T_e(x)/dx = 0$). On rearranging (3.9), we get a value of the electric field that would under stationary conditions correspond to a given value of $\Delta T_e^{(S)}$ as

$$|E|^{(S)} = \frac{5k_B \Delta T_e^{(S)}}{2q_e \lambda_e}\,. \tag{3.10}$$

Since the dependence of α on ΔT_e is physically more direct than that of α on $|E|$, following [1] we shall assume that after a substitution of $|E|^{(S)}$ for $|E|$ in relation (3.8), one may replace $\Delta T_e^{(S)}$ by $\Delta T_e(x)$ and subsequently assume that the relation thus obtained between α and $\Delta T_e(x)$ is applicable under circumstances in which the electron temperature is position dependent. The resulting expression for impact-ionization rate as a function of the electron temperature is

$$\alpha[\Delta T_e(x)] \;=\; A_n \exp\left[-\frac{2q_e\lambda_e}{5k_B}\frac{B_n}{\Delta T_e(x)}\right]. \tag{3.11}$$

This expression for the impact-ionization rate, will be used in the next subsection to determine the avalanche multiplication factor of the ionized charge carriers across the collector-epilayer region.

3.2.3 Multiplication factor: $(M_n - 1)$

The weak avalanche current $I_{\text{avl}} = (M_n - 1)I_C$ [10] follows from the collector current I_C through the global multiplication factor $(M_n - 1)$ [6], which is defined as the integral of the electron temperature dependent impact-ionization rate (3.11) over the collector-epilayer width W_{epi}:

$$(M_n - 1) \;=\; \int_0^{\mathrm{W_{epi}}} \alpha[\Delta T_e(x)]\,dx\,. \tag{3.12}$$

We note that, in any realistic regime of parameters, as valid for conditions in the collector region of bipolar transistors, the function $\alpha[\Delta T_e(x)]$ (3.11), turns out to be a sharply peaked function about position x_{max} given by relation (3.7), with the maximum value

$$\alpha[\Delta T_e(x_{\text{max}})] = A_n \exp\left[\frac{B_n}{a x_{\text{max}} + E_0}\right]. \tag{3.13}$$

Such sharply peaked functions cannot be integrated directly, and in most cases, integration of such functions can be done only numerically, by taking advantage of the known position of the peak [25]. Coordinate transform can be done to obtain a more favorable integrand and then a fixed quadrature rule or even an iterative routine can be used to evaluate the transformed integral. Another option is to perform the integral in three segments, so that the sharp peak can be integrated over a narrow interval [26]. Though accurate numerical solutions can be achieved using such methods, none of them is suitable for standard compact bipolar transistor models, such as Mextram [7]. This means that new approximations of the integral for sharply peaked functions such as $\alpha[\Delta T_e(x)]$ in (3.11) need to be developed. In the next subsection, we will derive a new method for integrating sharply peaked functions. The resulting relation will be used to determine

an approximate expression for the multiplication factor $(M_n - 1)$, from relation (3.12), which is suitable for implementation in standard compact bipolar transistor models.

3.2.4 Approximate analytical integration of sharply peaked functions

Consider the class f_β of smooth but sharply peaked functions, which consist of real valued functions that depend on a set of parameters β. Suppose that f_β, as a function of spatial variable x is significantly non-zero only in a narrow spatial interval about a central point s, such that the integral over x along all of the real axis of any of these functions can be approximated by the integral over a finite spatial interval between $s-\delta$ and $s+\delta$, i.e.,

$$\int_{-\infty}^{\infty} f_\beta(x)dx \approx \int_{s-\delta}^{s+\delta} f_\beta(x)dx, \tag{3.14}$$

where δ is the width of the peak from the central point s. Furthermore, if the peak can formally be assumed to be narrow, such that δ can be considered to be a small parameter, then the integral of such a function f_β along the real axis can be expanded in a Taylor series in δ (about the central point s), this is elaborated below.

The function $f_\beta(x)$ can be expanded in a Taylor series about a point $x = s$ as

$$f_\beta(x) = f_\beta(s) + \frac{f_\beta^{'}(s)}{1!}(x-s) + \frac{f_\beta^{''}(s)}{2!}(x-s)^2 + \frac{f_\beta^{'''}(s)}{3!}(x-s)^3 + \frac{f_\beta^{''''}(s)}{4!}(x-s)^4 + \cdots. \tag{3.15}$$

On combining (3.14) and (3.15), we have

$$\begin{aligned}
\int_{-\infty}^{\infty} f_\beta(x)dx &\approx \int_{s-\delta}^{s+\delta} f_\beta(x)dx \qquad (3.16)\\
&= \int_{s-\delta}^{s+\delta} f_\beta(s)dx + \int_{s-\delta}^{s+\delta} \frac{f_\beta^{'}(s)}{1!}(x-s)dx + \int_{s-\delta}^{s+\delta} \frac{f_\beta^{''}(s)}{2!}(x-s)^2 dx\\
&+ \int_{s-\delta}^{s+\delta} \frac{f_\beta^{'''}(s)}{3!}(x-s)^3 dx + \int_{s-\delta}^{s+\delta} \frac{f_\beta^{''''}(s)}{4!}(x-s)^4 dx + \cdots.
\end{aligned}$$

A computation of the first few terms on the right hand side of (3.16), yields

$$\begin{aligned}
\int_{s-\delta}^{s+\delta} f_\beta(s)dx &= 2f_\beta(s)\delta, \qquad (3.17)\\
\int_{s-\delta}^{s+\delta} \frac{f_\beta^{'}(s)}{1!}(x-s)dx &= 0,\\
\int_{s-\delta}^{s+\delta} \frac{f_\beta^{''}(s)}{2!}(x-s)^2 dx &= 2f_\beta^{''}(s)\frac{\delta^3}{3!},\\
\int_{s-\delta}^{s+\delta} \frac{f_\beta^{'''}(s)}{3!}(x-s)^3 dx &= 0,\\
\int_{s-\delta}^{s+\delta} \frac{f_\beta^{''''}(s)}{4!}(x-s)^4 dx &= 2f_\beta^{''''}(s)\frac{\delta^5}{5!},
\end{aligned}$$

which can be expressed as:

$$\int_{s-\delta}^{s+\delta} \frac{f_{\beta}^{(n)}(s)}{n!}(x-s)^n dx = \begin{cases} 2f_{\beta}^{(n)}(s)\frac{\delta^{\mathrm{n}+1}}{(n+1)!}, & \text{if n is even} \\ 0, & \text{if n is odd} \end{cases} \tag{3.18}$$

where $f_{\beta}^{(n)}$ represents the nth order derivative of f_{β} with respect to the spatial variable x and $\delta^{\mathrm{n}+1}$ represents δ raised to power $(n+1)$.
Using relation (3.18), (3.16) can be rewritten as

$$\int_{-\infty}^{\infty} f_{\beta}(x)dx \approx \int_{s-\delta}^{s+\delta} f_{\beta}(x)dx = 2\sum_{n=0}^{\infty} f_{\beta}^{(n)}(s)\frac{\delta^{\mathrm{n}+1}}{(n+1)!}, \tag{3.19}$$

where the factor 2 is due to symmetry of the chosen interval of integration about the central point s. If we take the first two terms of the expansion, we have

$$\int_{-\infty}^{\infty} f_{\beta}(x)dx \approx \int_{s-\delta}^{s+\delta} f_{\beta}(x)dx = 2f_{\beta}(s)\delta + O(\delta^3), \tag{3.20}$$

where the second term in δ in this expansion vanishes so that, while this approximation is linear in δ, it is accurate up to third order in δ. We can use as a measure for the value of δ in (3.20) the position of a point of inflexion, i.e., a solution of the equation

$$f_{\beta}''(s+\delta_i) = 0, \tag{3.21}$$

where f_{β}'' represents the second order derivative of f_{β} with respect to variable x. Expanding this equation (3.21) as a Taylor series up to third order in δ_i [consistently with the accuracy of expression (3.20)], we obtain

$$f_{\beta}''(s+\delta_i) = f_{\beta}''(s) + \delta_i f_{\beta}'''(s) + \frac{1}{2}\delta_i^2 f_{\beta}''''(s) + O(\delta_i^3). \tag{3.22}$$

Now truncating the expansion (3.22) and solving for δ_i [by combining (3.21) and (3.22)], we find as an estimate for the distance δ_i from the point s to the *right hand side* point of inflexion

$$\delta_i \approx \frac{-f_{\beta}'''(s) + \sqrt{f_{\beta}'''(s)^2 - 2f_{\beta}''(s)f_{\beta}''''(s)}}{f_{\beta}''''(s)}, \tag{3.23}$$

where f_{β}''' and f_{β}'''' denote the third and fourth order derivative of f_{β} with respect to variable x, respectively.

For the appropriate value of δ in (3.20), we consider $\delta = \gamma_f \delta_i$, where γ_f is a constant that can in practice be chosen to calibrate the resulting expression (3.24) for a class of functions f_{β} as a whole; in other words, γ_f denotes a fixed real number that does not depend on the parameters β. As an illustration, the value for γ_f suitable for integration of Gaussian distributions is given in Appendix A.1. By combining (3.20) and (3.23), we arrive at the following general approximate expression for the integral over all the real

x-axis of a function $f_\beta(x)$ that is sharply peaked and significantly non-zero only in the neighborhood of a central point s:

$$\int_{-\infty}^{\infty} f_\beta(x)dx \approx \frac{2\gamma_f f_\beta(s)\left(-f_\beta'''(s)+\sqrt{f_\beta'''(s)^2-2f_\beta''(s)f_\beta''''(s)}\right)}{f_\beta''''(s)} . \tag{3.24}$$

The practical use of this approximate expression (e.g., for purposes in compact modeling), lies in the fact that the right hand side depends on the function $f_\beta(x)$ and its derivatives only. This means that the integral approximation can be evaluated and expressed in terms of the elementary functions of any analytical function $f_\beta(x)$, without any need for an explicit expression of the primitive function of $f_\beta(x)$ in terms of elementary functions. Furthermore, in practical applications such as compact modeling, the function $f_\beta(x)$ usually depends on parameters β. In such cases, expression (3.24) offers an explicit expression for the integral of $f_\beta(x)$ in terms of elementary functions and as a function of the parameters β of f_β. Hence, as opposed to (computationally expensive) numerical integration methods, our method enables explicit approximate computation of the integral of sharply peaked functions, while taking parameter dependence into account. This final expression (3.24) can be used to determine the approximate expression for the integral of the electron temperature dependent impact-ionization coefficient over the collector-epilayer width [i.e., the multiplication factor (M_n-1)], this will be presented in the next subsection.

3.2.5 Approximate expression for the multiplication factor (M_n-1)

Since the function $\alpha[\Delta T_e(x)]$ in (3.11) is sharply peaked and rapidly approaches zero away from maximum position x_{max} in (3.7), the integral (3.12) of $\alpha[\Delta T_e(x)]$ over the collector-epilayer width [i.e., the multiplication factor (M_n-1)], can now be approximated using the method described in the previous subsection. The aim is to come up with an explicit expression for (M_n-1), that can be incorporated in a complete compact transistor model. For this purpose, we assume that an expression for the relevant local electric field of the form (3.3) is readily available, and at position x_{max} (3.7), this electric field has the value E_{avl} given by

$$E_{\text{avl}} = a x_{\text{max}} + E_0 . \tag{3.25}$$

Employing this E_{avl} together with approximation (3.24), we arrive at the following expressions for the sharply-peaked real valued function $\alpha[\Delta T_e(x)]$ and its corresponding

derivatives at position $x = x_{\max}$:

$$
\begin{aligned}
\alpha[\Delta T_e(x_{\max})] &= A_n \exp\left[\frac{B_n}{E_{\text{avl}}}\right], \\
\frac{d}{dx}\left(\alpha[\Delta T_e(x)]\right)\Big|_{x=x_{\max}} &= 0, \\
\frac{d^2}{dx^2}\left(\alpha[\Delta T_e(x)]\right)\Big|_{x=x_{\max}} &= -B_n\left[\frac{1}{E_{\text{avl}}^2}\cdot\frac{a}{\lambda_e}\right] A_n \exp\left[\frac{B_n}{E_{\text{avl}}}\right], \\
\frac{d^3}{dx^3}\left(\alpha[\Delta T_e(x)]\right)\Big|_{x=x_{\max}} &= B_n\left[\frac{1}{E_{\text{avl}}^2}\cdot\frac{a}{\lambda_e^2}\right] A_n \exp\left[\frac{B_n}{E_{\text{avl}}}\right], \\
\frac{d^4}{dx^4}\left(\alpha[\Delta T_e(x)]\right)\Big|_{x=x_{\max}} &= \left[-B_n\left(\frac{-6a^2}{E_{\text{avl}}^3\lambda_e^2}+\frac{a}{E_{\text{avl}}^2\lambda_e^3}\right)+B_n^2\left(\frac{1}{E_{\text{avl}}^2}\cdot\frac{a}{\lambda_e}\right)^2\right] A_n \exp\left[\frac{B_n}{E_{\text{avl}}}\right].
\end{aligned}
$$

Substituting these expressions for $\alpha[\Delta T_e(x)]$ and its corresponding derivatives at $x = x_{\max}$ into (3.24), and then simplifying further, we arrive at an approximation for the integral of the impact-ionization coefficient in (3.12) or the multiplication factor $(M_n - 1)$ as

$$
(M_n - 1) \approx \frac{4\gamma_f E_{\text{avl}} A_n \lambda_e \exp(B_n/E_{\text{avl}})}{E_{\text{avl}} - \sqrt{6a\lambda_e(2E_{\text{avl}} + B_n) - E_{\text{avl}}^2}}. \tag{3.26}
$$

We note that in this expression, except the electric field variable E_{avl} and field gradient a, the rest of the quantities (in the right-hand side) can be treated as model constants or model parameters, whose temperature dependence will be discussed in the next subsection. In practice, a value of the constant coefficient γ_f can be found, once and for all, by comparing the results for $(M_n - 1)$ obtained by (3.26) with those obtained by numerical integration of relation (3.12). With respect to integration of the class of functions (3.11) and in the context of our present application, we found $\gamma_f = 1.645$ to be adequate. From expression (3.26), we can see that once the local electric field distribution is provided by the complete compact transistor model, then (3.26) provides an *explicit* formulation for the non-local avalanche multiplication factor, which is based on the non-local impact-ionization rate. With more care on the denominator term (i.e., to avoid cases of this term becoming zero, as well as the square root term becoming negative), this expression can easily be incorporated in a full compact transistor model, thus enable the simulation of non-local weak avalanche current $I_{\text{avl}} = (M_n - 1)I_C$.

3.2.6 Temperature scaling of model parameters

In the context of compact modeling of bipolar transistors, a lot of experience has been built up in the past by applying the Chynoweth's empirical law for ionization coefficient [11] and experimental results for representative values as well as the temperature dependence of the coefficients A_n and B_n (3.8) have been incorporated in compact models, e.g., [6, 9, 10]. The only new physical parameter introduced by expression (3.26) is the energy relaxation length λ_e. Following [13], we model the dependence of this parameter

λ_e on ambient temperature T by a power law as

$$\lambda_e(T) = \lambda_e(T_{\text{ref}}) \left(\frac{T}{T_{\text{ref}}} \right)^{\mathrm{A}_{\lambda_\mathrm{e}}}, \tag{3.27}$$

where T_{ref} denotes the reference temperature. The quantities $\lambda_e(T_{\text{ref}})$ and A_{λ_e} are new constant non-local avalanche model parameters. This temperature dependence of the relaxation length λ_e (3.27) is inspired by the relation [1]

$$\lambda_e = \frac{5}{3} \tau_e v_{\text{sat}}, \tag{3.28}$$

where τ_e denotes the energy relaxation time and v_{sat} the carrier saturation drift velocity; it is related to the critical electric field E_c by the relation $v_{\text{sat}} = \mu E_c$, where μ is electron ohmic mobility. We note that both E_c and μ strongly depend on temperature, and according to [21, 27], this dependence is defined as: $E_c(T) = E_c(T_{\text{ref}})[T/T_{\text{ref}}]^{\gamma}$ and $\mu(T) = \mu(T_{\text{ref}})[T/T_{\text{ref}}]^{\xi}$, respectively. Here, γ and ξ are temperature coefficients, and according to [21], $\gamma \approx 1.55$ and $\xi \approx -2.42$. Since the dependence of τ_e on ambient (i.e., lattice) temperature [28] is much weaker than that of v_{sat} [21, 27], the temperature scaling of λ_e will be similar to the temperature scaling of v_{sat}, hence the form (3.27). Therefore, the value of the temperature coefficient $\mathrm{A}_{\lambda_\mathrm{e}}$ can be expected to closely approximate the corresponding parameter for the temperature dependence of v_{sat}.

By comparing relation (3.28) with (3.27), we see that the model parameter for the relaxation length $\lambda_e(T_{\text{ref}})$ can be estimated as

$$\lambda_e(T_{\text{ref}}) = \frac{5}{3} \tau_e v_{\text{sat}}(T_{\text{ref}}). \tag{3.29}$$

With the approximate values of $\tau_e = 0.4$ psec and $v_{\text{sat}}(T_{\text{ref}}) = 1.0 \times 10^7$ cm/sec, for electrons at high fields reported in, e.g., [27] at $T_{\text{ref}} = 300$ K, (3.29) yields an approximate value of $\lambda_e(T_{\text{ref}}) = 66.7$ nm. This value is in agreement with $\lambda_e(T_{\text{ref}}) = 65.0$ nm, as published in [1] for pure Si bipolar devices; more results in this direction will be presented in the next chapter.

3.3 Implementation of the non-local avalanche compact model in Mextram

In the previous section, we derived an expression for the non-local avalanche current I_{avl} as a product of the multiplication factor $(M_n - 1)$ and the collector current I_C, where the multiplication factor is given by the approximate relation (3.26). Also the temperature dependence of the new introduced physical model parameter, i.e., λ_e was provided using a power law. In industrial applications, i.e., as part of a complete compact transistor model, expression (3.26) is likely to become subject to evaluation for parameter and bias values outside its range of physical validity, for example during automated parameter optimization procedures. One possible instance of this would be evaluation in the high current regime, in which the slope of the electric field, represented in (3.26) by a, would

vanish or even charge sign due to Kirk effect [6, 18, 20, 29, 30]. This would cause the argument of the square root function in (3.26) to become negative.

To address this issue, we follow e.g., [12] in the viewpoint that any straightforward application of (3.26), in combination with reversal of the sign of a is actually unrealistic, because in the targeted high current regime, spatial current spreading effects [20] will likely occur in practical bipolar transistors. In the present work therefore, we shall restrict ourselves to developing a counterpart of (3.26) that will provide a good approximation in the low and intermediate current regimes, while it will be such that the avalanche current I_{avl}, will become negligible in the high current limit. In other words, the modifications we shall apply will be such that they preserve accuracy of the formulation for low and intermediate collector currents, but suppress the avalanche current in the high current limit, so as to make the model formulation globally robust; such can be achieved as follows.

In practical circumstances, except in the limit for vanishing a where we consider (3.26) to be inaccurate anyhow, the term $-E_{\text{avl}}^2$ contributes only a percentage of the full argument of the square root function. Then the first step is to drop the term $-E_{\text{avl}}^2$ and replace E_{avl} by a bounded counterpart $\tilde{E}_{\text{avl}}$ (see Appendix A.2), which yields the new expression for (3.26) as:

$$(M_n - 1) \quad \approx \quad \frac{4\gamma_f \tilde{E}_{\text{avl}} A_n \lambda_e \exp(B_n/\tilde{E}_{\text{avl}})}{\tilde{E}_{\text{avl}} - \sqrt{6a\lambda_e(2\tilde{E}_{\text{avl}} + B_n)}}. \tag{3.30}$$

The second step to be taken is to make sure that the incorporation of this expression in the context of a complete compact transistor model is such that:

(a) the electric field gradient a approaches zero asymptotically in the high current limit (so that a will never vanish nor change sign due to Kirk effect); this will ensure robustness of (3.30), and

(b) the bounded function $\tilde{E}_{\text{avl}}$ is an approximation of E_{avl} that is smoothly bounded from below by a limiting value $-0.5B_n$; such implementation of smoothly bounded functions is a standard technique in the context of compact modeling (see Appendix A.2).

Inclusion of expression (3.30) in the avalanche dedicated part of a complete compact model like Mextram is straightforward. For example, with respect to Mextram 504 [6, 16], the value of quantity E_0 in (3.25) is equal to the negative of Mextram's model (current dependent) variable E_M (see Section 1.3.3). The quantity a in (3.25) would then be given by E_M/λ_D. Here, λ_D is another Mextram model (current and bias dependent) variable, representing the extrapolated width of the depletion layer, disregarding its limitation by e.g., a buried layer (see Chapter 1). Note that, both E_M and λ_D are strictly positive quantities. Employing these definitions for quantities a and E_0, together with relations (3.7) and (3.25), E_{avl} can now be calculated as

$$E_{\text{avl}} \quad = \quad E_M\left(\frac{\lambda_e}{\lambda_D}\ln\left[1+\frac{\lambda_D}{\lambda_e}\right]-1\right). \tag{3.31}$$

Since both E_M and λ_e are strictly positive, it follows from (3.31) that $E_{\text{avl}} < 0$, so that the denominator of (3.30) will be negative and thus the expression as a whole will be numerically well-defined. In the Mextram model, the multiplication factor $(M_n - 1)$ corresponds

to the *model variable* for the generation factor G_{EM} [7]. All other quantities in expression (3.30) correspond to existing Mextram model parameters or model constants [6].
We note that relation (3.30) can easily be incorporated in a full compact bipolar transistor model, with only two extra introduced model parameters, i.e., the relaxation length λ_e and its temperature coefficient A_{λ_e}. The extraction of the parameter values for these new non-local avalanche model parameters will be discussed in the next chapter.

3.4 Summary

In this chapter, we presented a physics based compact model for non-local avalanche in bipolar transistors. This model is based on the approximate energy balance equation [1], which yielded the relation for the electron temperature $\Delta T_e(x)$ in terms of the energy relaxation length λ_e and the spatial local electric field $E(x)$. Using the Chynoweth's empirical law [11], together with the steady state condition for the electron temperature, we derived an expression for the impact-ionization rate as a function of the electron temperature, i.e., $\alpha\,[\Delta T_e(x)]$. In any realistic regime of parameters, as valid for conditions in the collector region of bipolar transistors, the function $\alpha\,[\Delta T_e(x)]$, turns out to be a sharply peaked function about the maximum position $x_{\max}$. The integral of $\alpha\,[\Delta T_e(x)]$ over the collector epilayer width yielded an expression for the multiplication factor (M_n-1), which was then used to determine the non-local weak avalanche current, as $I_{\text{avl}} = (M_n-1)I_C$. By taking advantage of the fact that the function $\alpha\,[\Delta T_e(x)]$, is sharply peaked and approaches zero away from maximum position $x_{\max}$, we derived an approximate expression for (M_n-1), by means of a Taylor series expansion of the integral in terms of the width of the peak. The resulting formulation is explicit in nature and in terms of elementary functions, so it can be readily implemented in existing complete compact bipolar transistor models, such as Mextram [6], whenever the model provides an explicit relation for the electric field. We took into account the temperature dependence of the new introduced physical parameter λ_e, using the power law (which is based on the product of the energy relaxation time and the electron saturation drift velocity [1]), and this resulted into another model parameter, i.e., the temperature coefficient A_{λ_e}. Thus, in total only two new model parameters were introduced; as we will discuss in the next chapter, these parameters correspond to independent material coefficients, so that their values can be found in the semiconductor device literature.

For purposes of implementation in a complete compact transistor model, like Mextram [7], the approximate expression for the multiplication factor (M_n-1) was modified further, in order to address cases where it becomes subject to evaluation for parameter and bias values outside its range of physical validity. The resulting expression is globally robust, and for low and intermediate current levels it closely approximates the more fragile but more physical version. For the case of the Mextram model, the existing local electric field (bias and current dependent) variables are used to compute the input electric field variable E_{avl} corresponding to the maximum position $x_{\max}$. The multiplication factor (M_n-1) corresponds to the Mextram model variable for the generation factor G_{EM} [6], while other quantities correspond to existing Mextram model parameters or model constants [16]. Thus, this new non-local avalanche compact model can easily be incorporated in the standard Mextram model. This will be demonstrated in the next

chapter, where we will employ an experimental variant of Mextram model (i.e., Mextram model extended with our new non-local avalanche compact model or extended version of Mextram) together with the measurements taken on both Si and SiGe industrial bipolar transistors, to carry out an experimental assessment of the relevant and significant physical mechanisms on the electrical characteristics in the weak collector-base breakdown regime.

References

[1] J. W. Slotboom, G. Streutker, M. J. v. Dort, P. H. Woerlee, A. Pruijmboom, and D. J. Gravesteijn, *Non-local impact ionization in silicon devices,* in *Proc. IEDM* (1991) pp. 127–130.

[2] G. B. Hong and J. G. Fossum, *Implementation of nonlocal model for implact-ionization current in bipolar circuit simulation and application to SiGe HBT design optimization,* IEEE Trans. Electron Devices **42**, 1166 (1995).

[3] M. M. Mahim and J. S. Yuan, *Modelling of the avalanche current including the non-local effect,* Int Journal Electronics **85**, 400 (1998).

[4] H. C. de Graaf and W. J. Kloosterman, *The Mextram Transistor Model*, Unclassified Report NL-UR 006/94 (Koninklijke Philips Electronics N. V., Nat. lab, 1994).

[5] J. A. Yuan, *SiGe, GaAs, and InP Heterojunction Bipolar Transistors*, edited by K. Chang (John Wiley and Sons, Inc, 1999).

[6] J. C. J. Paasschens, W. J. Kloosterman, and R. van der Toorn, *Model Derivation of Mextram 504, The physics behind the model*, Tech. Rep. Technical note TN-2004/01010 (Philips Research Eindhoven, 2005).

[7] R. van der Toorn, J. C. J. Paasschens, W. J. Kloosterman, and H. C. de Graaff, *Compact Modeling: Principles, techniques, and applications,* (Springer-Verlag, 2010) Chap. 7.

[8] W. J. Kloosterman and H. C. de Graaff, *Avalanche multiplication in a compact bipolar transistor model for circuit simulation,* in *Proc. BCTM* (1988) pp. 103–106.

[9] W. J. Kloosterman and H. C. de Graaff, *Avalanche multiplication in a compact bipolar transistor model for circuit simulation,* IEEE Trans. Electron Devices **36**, 1376 (1989).

[10] W. J. Kloosterman, J. C. J. Paasschens, and R. J. Havens, *A comprehensive bipolar avalanche multiplication compact model for circuit simulation,* in *Proc. BCTM* (2000) pp. 172–175.

[11] A. G. Chynoweth, *Ionization rates for electrons and holes in silicon,* Phys. Rev **109**, 1537 (1958).

[12] M. Schröter and A. Chakravorty, *Compact Hierarchical Bipolar Transistor Modeling with HICUM* (World Scientific Publishing Co. Pte. Ltd, 2010).

[13] M. R. van den Berg, L. K. Nanver, and J. W. Slotboom, *Temperature dependence of avalanche multiplication in spiked electric fields,* in *Proc. IEDM* (2000) pp. 79–82.

[14] G. Streutker and J. W. Slotboom, *Non-local avalache calculation for deep sub-micron devices,* Microelectronic Engineering **19**, 299 (1992).

[15] F. Lin, B. Chen, T. Zhou, B. L. Ooi, and P. S. Kooi, *Characterization and modeling of avalanche multiplication in HBTs,* in *Proc. SPIE* (2000) pp. 158–163.

[16] M. Reisch, *High-Frequency Bipolar Transistors* (Springer-Verlag, 2003).

[17] G. A. M. Hurkx, *Bipolar and bipolar-MOS integration,* (Elsevier, Amsterdam, 1994) Chap. 3.

[18] H. C. de Graaff and F. M. Klaassen, *Compact transistor modelling for circuit design* (Springer-Verlag, 1990).

[19] S. M. Sze and K. N. Kwok, *Physics of Semiconductor Devices,* 3rd ed. (Wiley, 2006).

[20] J. C. Paasschens, W. J. Kloosterman, R. J. Havens, and H. C. de Graaf, *Improved compact modeling of output conductance and cut-off frequency of bipolar transistor,* in *Proc. BCTM* (2000) pp. 62–65.

[21] C. Canali, G. Majni, R. Minder, and G. Ottaviani, *Electron and hole drift velocity measurements in silicon and their empirical relation to electric field and temperature,* IEEE Trans. Electron Devices **22**, 1045 (1975).

[22] P. A. H. Hart, ed., *Bipolar and Bipolar-MOS Integration* (Elsevier, 1994).

[23] H. C. de Graaf, *Electrical Behavior of Lightly Doped Collectors in Bipolar Transistors,* Phd. thesis, Eindhoven University of Technology (1975).

[24] H. Poon and J. Meckwood, *Avalanche effects in integral control model,* IEEE Trans. Electron Devices **19** (1972).

[25] H. H. H. Hometer and E. O. Steinborn, *Numerical integration of functions with a sharp peak at or near one boundary using möbius transformations,* J. Computational Physics **87**, 61 (1988).

[26] H. H. H. Hometer and E. O. Steinborn, *Numerical integration of functions with a sharp peak at or near one boundary using möbius transformations,* J. Computational Physics **87**, 61 (1988).

[27] C. Jacoboni, C. Canali, G. Ottaviani, and A. A. Quaranta, *A review of the some charge transport properties of silicon,* IEEE Trans. Electron Devices **20**, 77 (1977).

[28] G. Sasso, N. Rinaldi, G. Matz, and C. Jungemann, *Accurate mobility and energy relaxation time models for SiGe HBTs numerical simulation,* in *SISPAD* (IEEE, 2009).

[29] C. T. Kirk, *A theory of transistor cut-off frequency (f_t) fall-off at high current densities,* IEEE Trans. Electron Devices .

[30] P. Ashburn, *SiGe Heterojunction Bipolar Transistors* (John Wiley and Sons Ltd, 2003).

4

Experimental assessment of the physical mechanisms in the weak collector-base breakdown regime

A detailed experimental assessment of the relevant and significant physical mechanisms that exist in the weak collector-base breakdown regime and result into the deviations between the measured and Mextram-simulated family of characteristics in present day Si/SiGe bipolar transistors will be presented in this chapter. A Mextram model extended with our physics based non-local avalanche compact model (i.e., an extended version of Mextram) will be used; other than following the traditional approach of carrying our model verification on only the fitting capabilities against the measured characteristics, we will also verify the physical content of this model. Here, we will demonstrate that the model is on the one hand sufficiently complete and accurate while on the other hand its parameters are sufficiently independent, for the parameter extraction to be taken as an in situ measurement for both the newly introduced model parameters (the electron energy relaxation length and its temperature coefficient): values obtained correspond to values published earlier in the semiconductor device literature. This approach is ambitious, since there exists a number of coupled physical effects that interfere with the weak avalanche characteristics and their temperature dependence; these must independently be taken into account. Our model verification will be carried out on measured data taken on both Si and SiGe industrial bipolar transistors, over a sequence of ambient temperatures. The variation in the magnitude of the non-local effects in relation to the employed device technologies will be discussed.

This chapter is based on the work submitted in the article:
R. Setekera and R. van der Toorn, *Solid-State Electronics*, manuscript submitted (July 2015).

4.1 Introduction

This chapter is fully dedicated to experimental assessment of the relevant and significant physical mechanisms in the weak collector-base breakdown regime of present day Si BJT's and SiGe HBT's that cause the deviations between the measured and simulated characteristics. For this purpose, we adopted the ambition to not only demonstrate adequate model fitting capabilities regarding the targeted electrical characteristics (a traditional approach), but also to demonstrate that our model actually captures the presupposed underlying semiconductor device physics. This later aspect shall be demonstrated by showing that the values obtained for the two relevant non-local avalanche model parameters (i.e., the relaxation length λ_e and its temperature coefficient A_{λ_e}), by optimization against the measured data, correspond to values that have been previously published in independent sources (i.e., semiconductor device literature). We note that, demonstrating such correspondence requires estimation of the accuracy of our extracted values of λ_e and A_{λ_e}. This includes a demonstration of the relative uniqueness of the extracted parameter values. Comparing extracted values to independently published values for these quantities is possible because they are material coefficients; thus our extraction for λ_e and A_{λ_e} can be viewed as an in situ measurement of these quantities. This requires adequate accounting for all interfering physical effects and independent parameter extraction for these, including their temperature dependencies. From a slightly different point of view, the results presented in this chapter can be viewed as an experimental analysis of the physical mechanisms that are of significance in the weak avalanche regime, in present day industrial bipolar devices. These physical mechanisms need to be adequately addressed before hand, to enable verification of the physical nature of our non-local avalanche model.

To achieve this, we incorporated our new non-local avalanche compact model in an experimental variant of the Mextram compact bipolar transistor model (version 504.10) with self-heating [1], i.e., we shall use a private extended version of Mextram. This enables us to profit from:

(a) Mextram's internal representation of the relevant local electric field distribution,

(b) Mextram's physics based coverage of all other possibly interfering physical effects, and

(c) Mextram's built-in temperature scaling rules [2–4] for the temperature dependence of Chynoweth's empirical law, (3.8).

For comparison purposes, we have implemented both expression (3.26) (which with some care can practically be operated by taking the absolute value of the square root argument) and its robust counterpart (3.30).

The verification of the semiconductor device physics, as presented in this chapter is based on (3.26); this choice avoids compromising the physical content of the model. The model fits obtained can also be achieved by (3.30), at the cost of compromising the extracted parameter values by only a small percentage. By employing Mextram's extensive capabilities, we sufficiently account for independent (relevant) physical mech-

anisms that may interfere with the weak avalanche characteristics; this includes their temperature dependence (Section 4.2), through self-heating effect [5].

To best serve our ambition to verify the physical content of our model, we chose to first verify (Sections 4.3 and 4.4) our non-local avalanche compact model against data taken from a transistor that will likely have the least alternative physical effects interfering with avalanche effects. For this reason, we chose to firstly verify our new non-local avalanche compact model against measurements taken on a pure Si (NPN) device. This choice avoids SiGe-specific interfering effects that may be present in e.g., high speed SiGe devices, such as enhanced Early effect and effects of the Ge-profile on temperature dependencies [6]. The selected device has a constant doping concentration in the collector epilayer; this was verified from the process technology data and also by the standard physics based C(V) technique [7]. Hence, the device matched the key assumption of constant doping level that underlines Mextram's description of the electric field [1, 8]. The accuracy levels for the extracted non-local avalanche parameter (i.e., λ_e and A_{λ_e}) for this device, will as well be studied (Section 4.5), as this is essential for quantification of the correspondence of these parameter values to the published literature values.
Note that, for comparison purposes, the standard Mextram avalanche compact model [4, 8] will also be simulated against the same measured data for the avalanche characteristics. This will enable us to observe the limited capabilities of the *local* model in fitting the temperature dependence of the avalanche current. In addition, it will help in the verification process of the physical content of the extracted electric field avalanche model parameters, as well as to measure the magnitude of the non-local effects in the employed devices.

This model verification strategy, i.e., testing the model fitting capabilities against the measured avalanche characteristics, will be extended to modern SiGe-HBT industrial devices (Section 4.6), from different bipolar technologies. The employed SiGe devices will be anonymous and denoted as: *high speed*, *high voltage*, and *very high voltage* devices. These devices will also be used to analyse (Section 4.7) the variation in the magnitude of non-local effects in relation to device technology.

Carrying out the model verification on both fitting capabilities and physical nature of the extracted model parameters is not straight forward and that is why it's "traditionally not" carried out in the available semiconductor device literature. This may be caused by the fact that most of the physical effects addressed by such compact models are coupled, thus analyzing such effects independently is complicated. In most available published literature, e.g., [4, 9, 10], the verification of the fitting capabilities of the developed avalanche models is carried out against the measurements of the multiplication factor $(M_n - 1)$. This tacitly presupposes avalanche is the sole significant physical mechanism; but for our case, we are also interested in addressing all interfering physical effects in the weak avalanche regime. Also due to the fact that the multiplication factor is a concept that describes the avalanche process alone, it is not possible to obtain a true estimate for $(M_n - 1)$ from the measured data over the avalanche regime; this is due to the fact that the measured data over the avalanche regime is also significantly affected by other temperature effects, through self-heating. Thus estimating $(M_n - 1)$ from the measured data would in general require an assumption of other effects being negligible, but this would suppress our goal of taking into account all these physical mechanisms so as to

4

arrive at a more detailed physical interpretation of the data. Therefore, in our model verification process, the various results will be presented directly in terms of the measured and simulated terminal current characteristics.

We note that, in all the model simulations, the ICCAP simulation software [11], in combination with the Levenberg-Marquardt optimization algorithm [12–14], will be used.

4.2 Experimental measurement setup

For the case of model verification, we took on-wafer DC-measurements of the relevant electrical characteristics as a function of both bias and ambient temperature. These measurements were taken directly on the devices (on-wafer), by contacting the various contact pads, i.e., the base (B), collector (C), and emitter (E) pads, as shown in the schematic representation of the transistor in Fig. 4.1. We note that for some employed devices with a collector-substrate, the substrate (S) pad (not indicated in Fig. 4.1) was also contacted during the measurement process. The applied boundary conditions to these measurements are the external biases, i.e., base-emitter (V_{BE}), collector-base (V_{CB}), and collector substrate (V_{CS}) voltages, and the ambient temperatures; while the observed quantities are the terminal currents, i.e., collector (I_C), base (I_B), emitter (I_E), and substrate (I_S) currents, and the junction depletion capacitances. Some of these input terminal voltages and output currents are indicated in Fig. 4.1, mostly for a transistor in forward operation mode (i.e, $V_{BE} > 0$ and $V_{CB} < 0$). The input bias conditions largely depend on the electrical characteristics to be measured and the limits for such biases depend largely on the breakdown voltages for the employed devices. The input ambient temperatures are adjusted by changing the global chuck temperature during the measurement process. Due to the applied external biases on the transistor terminals and the generated terminal currents, power is dissipated in the transistor; this leads to an increase in the local transistor temperature due to self-heating [5, 15, 16]. Thus, depending on the magnitude of the applied terminal biases, the local device temperature may differ from the applied ambient (chuck) temperature, due to the generated device self-heating; this will be addressed further in the next section.

The essential DC-measurements for the relevant electrical characteristics, together with the applied measurement bias conditions for this model verification process are summarized in Table. 4.1 of the next section. Since one of our key goals is to study the temperature dependence of the (weak) avalanche characteristics, we carried out measurements of the base current (I_B) over a sequence of ambient temperatures (i.e., $T = 25, 50, 75, 100$, and $125°$C), for a fixed base-emitter voltage (V_{BE}) and variable collector-base voltage (V_{CB}), which is increased until base current I_B has significantly decreased due to avalanche multiplication. The base-emitter bias is set to a low value, i.e., $V_{BE} = 0.65$ V (see Fig. 4.2), in order to minimize the influence of high current effects such as self-heating [5, 17], series resistance [18], and Kirk effect [3, 19, 20] on avalanche. These I_B measurements (i.e., forward Early measurements) were carried out on both pure Si and SiGe industrial bipolar transistors; the corresponding measurement data will be presented (by the symbols) in the various sections of this chapter. Note that in order to study the effect of small variations in V_{BE} on the simulated avalanche characteristics, we carried out similar forward Early measurements, but now with different fixed input

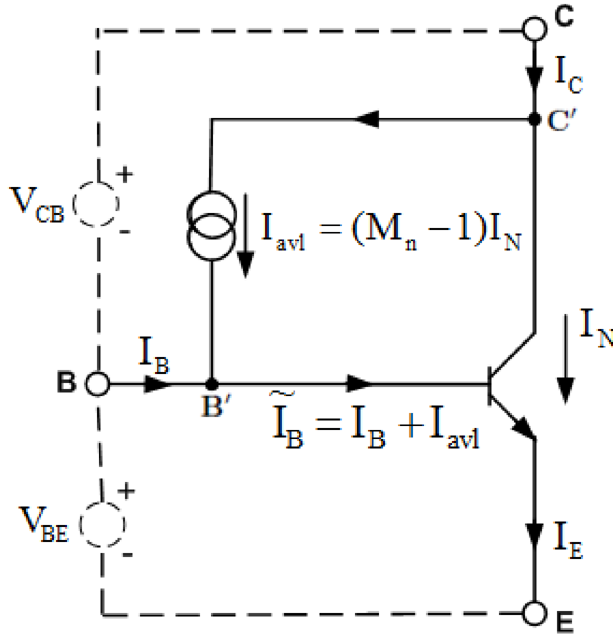


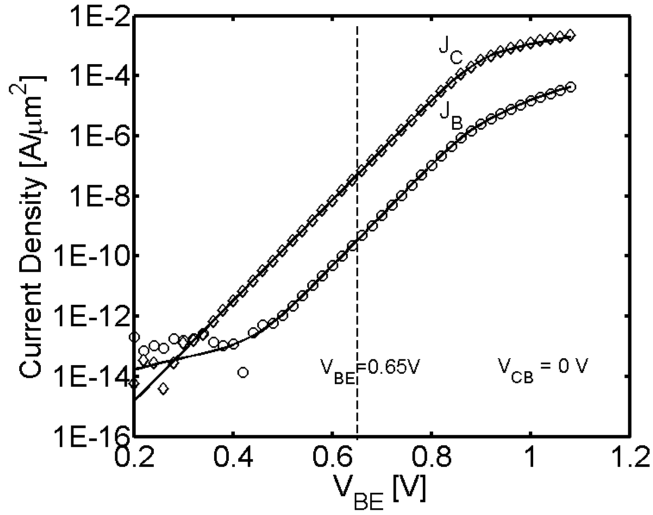
Figure 4.1: Schematic representation of the vertical NPN bipolar junction transistor showing the different current components and voltage sources of our measurement setup. Due to the generated avalanche current $I_{\mathrm{avl}} = (M_n - 1)I_N$, the terminal base current I_B will be decreased with increasing collector-base voltage V_{CB}. Here I_N is the main transistor current without avalanche and V_{BE} is the base-emitter voltage, which is kept fixed during the measurements for the avalanche characteristics.

4

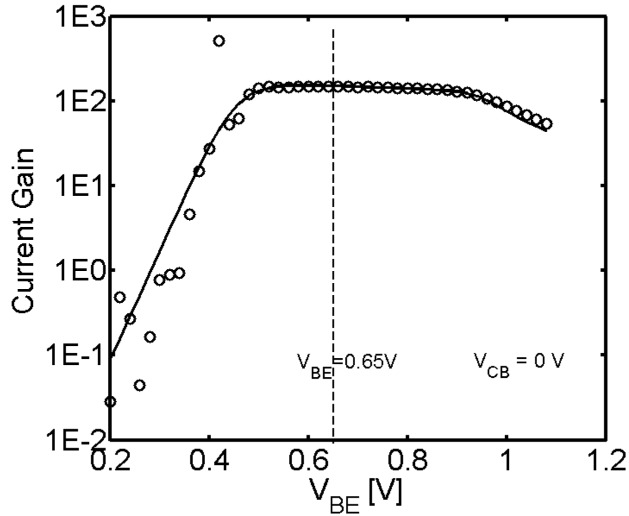
values of V_{BE}, e.g., with $V_{BE} = 0.55$ V.

Other relevant DC-measurements, i.e., base-emitter capacitance, collector-base capacitance, collector-substrate capacitance, forward Gummel, reverse Gummel, reverse Early, and output characteristics were also carried out over the same sequence of ambient temperatures, and on the same bipolar devices. These measurements will enable extraction of all relevant (input) Mextram model [8] parameters (indicated in Table. 4.1), as well as addressing all other relevant physical mechanisms that may interfere with the weak avalanche characteristics and their temperature dependence (as discussed in Section 4.3).

To elaborate more on these measurements, the forward Gummel measurements were carried out for variable base-emitter voltage V_{BE} and fixed collector-base voltage, i.e., $V_{CB} = 0$ V, over a sequence of the applied ambient temperatures; the sample results at $T = 25^{\circ}$C are presented (symbols) in Fig. 4.5(a), while sample measurements over all ambient temperatures are presented later in Fig. 4.4 of Section 4.3. This choice of $V_{CB} = 0$ V, minimizes the influence of avalanche, (forward) Early effect, and self-heating on the measured terminal collector and base currents [21–24]. As can be seen from the schematic representation in Fig. 4.1, when the external junction bias $V_{CB} = 0$, then the generated internal avalanche current $I_{\mathrm{avl}} = (M_n - 1)I_N \approx 0$, due to the minimized car-



(a)



(b)

Figure 4.2: Measured (symbols) and standard Mextram model simulated (solid curves) forward Gummel characteristics as function of base-emitter voltage (V_{BE}) and fixed collector-base voltage V_{CB} = 0 V, for a Si-BJT with effective emitter area $A_E = 0.50 \times 20.7\ \mu\text{m}^2$. (a) Shows the base current density J_B and collector current density J_C as a function of V_{BE}. (b) Shows the DC-forward current gain (= J_C/J_B) as a function of V_{BE}. From (a) and (b), various bias regimes of the transistor, i.e., non-ideal, ideal, and high-current, can be observed. To minimize the influence of high-current effects on avalanche, V_{BE} is taken from the ideal region as shown in (a) and (b), for our case we took V_{BE} = 0.65 V.

rier multiplication process (thanks to the small junction electric field) [2]. In such a case, the measured terminal collector current I_C, will approximate the main transistor current I_N (mostly in the normal forward operation mode). For the reverse Gummel measurements, the base (I_B) and emitter (I_E) terminal currents are measured as a function of collector-base voltage (V_{CB}) and ambient temperature, for fixed emitter-base voltage, i.e., $V_{EB} = 0$ V. This choice of $V_{EB} = 0$ V, minimizes the effects of (reverse) Early effect on the measured terminal currents. Measurements for the base-emitter depletion junction capacitance C_{BE} were carried out over all ambient temperatures and variable base-emitter junction bias, with $V_{CB} = 0$ V; while those of the collector-base junction depletion capacitance were carried out as a function of V_{CB} and temperature, with $V_{EB} = 0$ V. Note that for such depletion capacitance measurements, the applied junction bias must be taken moderate in order to avoid interference of the diffusion charges [25].

4.3 Accounting for independent physical mechanisms that interfere in the weak avalanche regime

As we discussed earlier, there are a number of independent physical mechanisms that may interfere with the weak avalanche characteristics. These effects need to be adequately addressed independently in order to enable verification of the presupposed underlying device physics for our non-local avalanche compact model. This includes in-depth extraction of the essential model parameters for these effects and their temperature dependence taken into account. In order to fully carter for these, one needs to incorporate our non-local model into a complete transistor model, which fully addresses these physical effects and their temperature dependence. In our case, we use an extended version of the standard Mextram compact model for bipolar transistors [8]. Using this experimental variant of Mextram model, together with the temperature measurements of relevant electrical characteristics taken on a QUBiC4+ (NPN) Si-BJT (with emitter area $A_E = 0.50 \times 20.7\ \mu\text{m}^2$, $\text{BV}_{\text{CE0}} = 5.3$ V, and $f_T / f_{\text{max}}(V_{CE} = 2.0V) = 35/80\,\text{GHz}$ at T = 25 °C), we exhaustively account for these relevant physical effects and their temperature dependence.

Notably, our measured data for the base current I_B, the collector current I_C, and the depletion capacitances (base-emitter and collector-base capacitances), in the avalanche regime is also significantly affected by other temperature effects through self-heating [26–28]; thus adequate extraction of the device thermal resistance R_{TH} (see Chapter 2) [5, 16, 24, 29, 30] is vital so as to account for self-heating. To demonstrate this, we consider the measured and Mextram model simulated results for electrical characteristics (over a sequence of ambient temperatures) corresponding to the employed pure Si-BJT.

Overview of involved temperature dependencies

According to Mextram (and other standard compact models for bipolar transistors, e.g., [31, 32]), self-heating is taken into account by means of an external thermal equivalent circuit shown in Fig. 4.3. Here, the value of the voltage v_{dT} at the temperature node dT

gives the increase in the local device temperature [8]. Then the corresponding device temperature T_{dev} is calculated from

$$T_{\text{dev}} = \text{TEMP} + \text{DTA} + 273.15 + v_{dT}, \tag{4.1}$$

and the reference temperature is

$$T_{\text{ref,K}} = \text{T}_{\text{ref}} + 273.15, \tag{4.2}$$

where T_{ref} (in °C) is the temperature at which the parameters were determined, TEMP (also in °C) is the actual ambient simulation temperature, and DTA is difference between the local ambient and global ambient temperatures. The normalized temperature is then defined as $t_N = T_{\text{dev}}/T_{\text{ref,K}}$. The Mextram self-heating network (Fig. 4.3), contains the

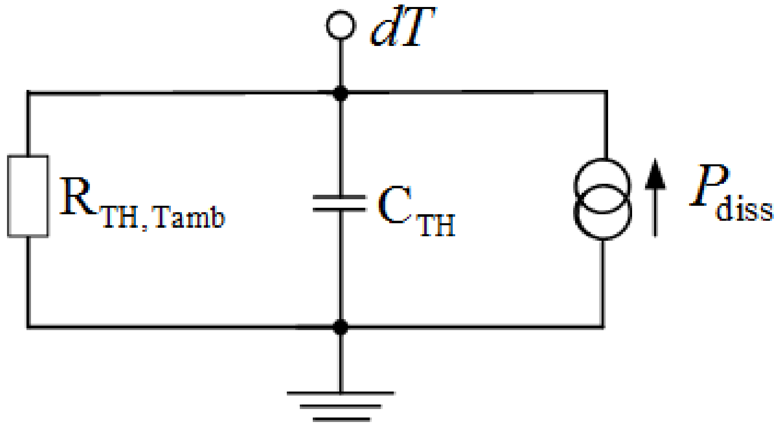


Figure 4.3: Mextram's self-heating network [8]. This circuit contains the thermal resistance $R_{TH,T_{amb}}$, which corresponds R_{TH}, but now after temperature scaling. C_{TH} is the capacitance and P_{diss}, is the dissipated power, which is taken as a current source. Node dT is the temperature node, which is available to the model user, in order to enhance flexibility.

self-heating resistance $R_{TH,T_{\text{amb}}} = \mathbf{R_{TH}}\left(T_{\text{amb}}/T_{\text{ref,K}}\right)^{\mathbf{A}_{\text{TH}}}$, where the ambient temperature $T_{\text{amb}} = \text{TEMP} + \text{DTA} + 273.15$, $\mathbf{R_{TH}}$ is the thermal resistance, and $\mathbf{A}_{\text{TH}}$ is the temperature coefficient of the thermal resistance (≈ 1.3 for Si and ≈ 1.15 for Ge). This same circuit also contains the capacitance $\mathbf{C}_{\text{TH}}$, which (as well as $R_{TH,T_{\text{amb}}}$) is connected between the ground and the temperature node dT [8]. The dissipated power P_{diss} is given by the product of all the terminal currents and voltages. Here, the **bold quantities** are Mextram model parameters [8].

Using the normalized temperature t_N, the temperature dependence of the various characteristics is taken into account via the temperature scaling rules for the various model parameters. For example, the main transistor current is defined by

$$I_N = \frac{I_{\text{sT}}}{q_B^I}\left[\exp\left(\frac{q_e V_{B_2E_1}}{k_B T_{\text{dev}}}\right) - \exp\left(\frac{-q_e V^*_{B_2C_2}}{k_B T_{\text{dev}}}\right)\right], \tag{4.3}$$

where $V_{B_2E_1}$ is the internal base-emitter voltage, $V^*_{B_2C_2}$ is the internal base-collector voltage that takes quasi-saturation effects in the epilayer into account, and

$$I_{\mathrm{sT}} = \mathbf{I_s}\, t_N^{4-\mathbf{A_B}-\mathbf{A_{QB0}}} \exp\left(\frac{-\mathbf{V_{gB}}}{V_{\Delta T}}\right), \tag{4.4}$$

with $\mathbf{I_s}$ the collector-emitter saturation current, $\mathbf{A_B}$ the temperature coefficient of the base, $\mathbf{V_{gB}}$ the bandgap voltage of the base,

$$V_{\Delta T} = \frac{k_B}{q_e}\left(T_{\mathrm{dev}} - T_{\mathrm{ref,K}}\right), \tag{4.5}$$

and $\mathbf{A_{Q_{B0}}}$ the temperature coefficient of the zero-base charge

$$Q_{\mathrm{B0}} = \tau_{BT} I_{kT} = \left(\tau_B\, t_N^{\mathbf{A_{QB0}}+\mathbf{A_B}-1}\right)\left(I_k\, t_N^{1-\mathbf{A_B}}\right), \tag{4.6}$$

with τ_B the transit time of the base stored charge and I_k the collector-emitter high injection knee current. In relation (4.3), the quantity q^I_B is the normalized base Gummel number [25, 33, 34]), and is given by the expression

$$q^I_B = q^I_0\left(1 + \frac{1}{2}n_0 + \frac{1}{2}n_B\right), \tag{4.7}$$

where

$$n_0 = \frac{\frac{4I_{\mathrm{sT}}}{I_{kT}} \exp\left(\frac{qV_{B_2E_1}}{K_B T_{\mathrm{dev}}}\right)}{1 + \sqrt{1 + \frac{4I_{\mathrm{sT}}}{I_{kT}} \exp\left(\frac{qV_{B_2E_1}}{K_B T_{\mathrm{dev}}}\right)}}, \tag{4.8}$$

and

$$n_B = \frac{\frac{4I_{\mathrm{sT}}}{I_{kT}} \exp\left(\frac{qV^*_{B_2C_2}}{K_B T_{\mathrm{dev}}}\right)}{1 + \sqrt{1 + \frac{4I_{\mathrm{sT}}}{I_{kT}} \exp\left(\frac{qV^*_{B_2C_2}}{K_B T_{\mathrm{dev}}}\right)}}. \tag{4.9}$$

These (i.e., n_0 and n_B) are also used to compute the base diffusion charges as: for the base-emitter part

$$Q_{BE} = \frac{1}{2} Q_{B0}\, n_0\, q^I_0, \tag{4.10}$$

and for the base-collector part

$$Q_{BC} = \frac{1}{2} Q_{B0}\, n_B\, q^I_0, \tag{4.11}$$

where

$$q^I_0 = 1 + \frac{\mathrm{V_{t_E}}}{\mathrm{V_{erT}}} + \frac{\mathrm{V_{t_C}}}{\mathrm{V_{efT}}}. \tag{4.12}$$

In this relation, i.e., (4.12), the temperature scaling for the forward Early voltage ($\mathbf{V_{ef}}$) is given by

$$V_{efT} = \mathbf{V_{ef}}\, t_N^{\mathbf{A_{QB0}}} \left[(1 - \mathbf{X_p})\left(\frac{\mathbf{V_{d_C}}}{V_{d_C T}}\right)^{\mathbf{p_C}} + \mathbf{X_p}\right]^{-1}, \tag{4.13}$$

where $V_{d_C T}$ is the temperature scaling of the collector-base diffusion voltage $\mathbf{V_{d_C}}$, which is defined as

$$V_{d_C T} = -3\frac{k_B T_{\text{dev}}}{q_e}\ln\{t_N\} + \mathbf{V_{d_C}}\, t_N + (1 - t_N)\mathbf{V_{g_C}}, \tag{4.14}$$

here $\mathbf{V_{g_C}}$ is the bandgap voltage of the collector. Also, in relation (4.12), the temperature scaling $V_{er T}$ for the reverse Early voltage ($\mathbf{V_{er}}$) is given by

$$V_{er T} = \mathbf{V_{er}}\, t_N^{\mathbf{A_{Q_{B0}}}}\left(\frac{\mathbf{V_{d_E}}}{V_{d_E T}}\right)^{-\mathbf{p_E}}, \tag{4.15}$$

where $V_{d_E T}$ is the temperature scaling of the emitter-base diffusion voltage $\mathbf{V_{d_E}}$, which is defined as

$$V_{d_E T} = -3\frac{k_B T_{\text{dev}}}{q_e}\ln\{t_N\} + \mathbf{V_{d_E}}\, t_N + (1 - t_N)\mathbf{V_{g_B}}. \tag{4.16}$$

The quantities V_{t_E} and V_{t_C} in (4.12) are bias dependent junction voltages, which are defined as

$$V_{t_E} = \frac{Q_{t_E}}{(1-\mathbf{XC_{j_E}})C_{j_E T}}, \tag{4.17}$$

and

$$V_{t_C} = \frac{Q_{t_C}}{\mathbf{XC_{j_C}}\, C_{j_C T}}, \tag{4.18}$$

where Q_{t_E} and Q_{t_C} are the base-emitter and base-collector junction depletion charges, respectively. While $C_{j_E T}$ and $C_{j_C T}$, are the corresponding (temperature scaled) zero-bias junction capacitances [8]; which are given by

$$C_{j_E T} = \mathbf{C_{j_E}}\left(\frac{\mathbf{V_{d_E}}}{V_{d_E T}}\right)^{\mathbf{p_E}}, \tag{4.19}$$

and

$$C_{j_C T} = \mathbf{C_{j_C}}\left[(1-\mathbf{X_P})\left(\frac{\mathbf{V_{d_C}}}{V_{d_C T}}\right)^{\mathbf{p_C}} + \mathbf{X_P}\right], \tag{4.20}$$

where $\mathbf{p_E}$ and $\mathbf{p_C}$ are respective junction grading coefficients [35]. We note that for SiGe transistors, q_0^I in (4.12) is redefined in order to carter for the effect of Ge grading on the Early effect, thus q_0^I is replaced by

$$q_{0_{Ge}}^I = \frac{\exp\left(\left[\frac{V_{t_E}}{V_{er T}}+1\right]\frac{q_e dE_{gT}}{k_B T_{\text{dev}}}\right) - \exp\left(\frac{-V_{t_C}}{V_{ef T}}\frac{q_e dE_{gT}}{k_B T_{\text{dev}}}\right)}{\exp\left(\frac{q_e dE_{gT}}{k_B T_{\text{dev}}}\right) - 1}, \tag{4.21}$$

where

$$dE_{gT} = \mathbf{dE_g}\, t_N^{\mathbf{A_{Q_{B0}}}} \tag{4.22}$$

is the temperature scaling for the $\mathbf{dE_g}$, the difference in the bandgap between the neutral edges of the base at zero bias. In the present Mextram model [8], the base current is defined in terms of the main current I_N (4.3)and the current gain β. This implies that the

corresponding temperature scaling rules for the current gain are also employed; these are defined as:

$$\beta_{fT} = \beta_{\mathbf{f}} t_N^{\mathbf{A_E}-\mathbf{A_B}-\mathbf{A_{QB0}}} \exp\left[\frac{-\mathbf{dV_{g\beta f}}}{V_{\Delta T}}\right], \tag{4.23}$$

for the forward current gain $\beta_{\mathbf{f}}$, and

$$\beta_{riT} = \beta_{\mathbf{ri}} \exp\left[\frac{-\mathbf{dV_{g\beta r}}}{V_{\Delta T}}\right], \tag{4.24}$$

for the reverse current gain $\beta_{\mathbf{ri}}$. The model parameters $\mathbf{dV_{g\beta f}}$ and $\mathbf{dV_{g\beta r}}$, are the bandgap voltage difference of forward current gain and reverse current gain, respectively.

Physical mechanisms interfering in the weak avalanche regime

From the above expressions, we see that in order to determine the weak avalanche current I_{avl}, which is defined by

$$I_{\mathrm{avl}} = (M_n - 1)\, I_N(V_{B_2E_1}, V^*_{B_2C_2}), \tag{4.25}$$

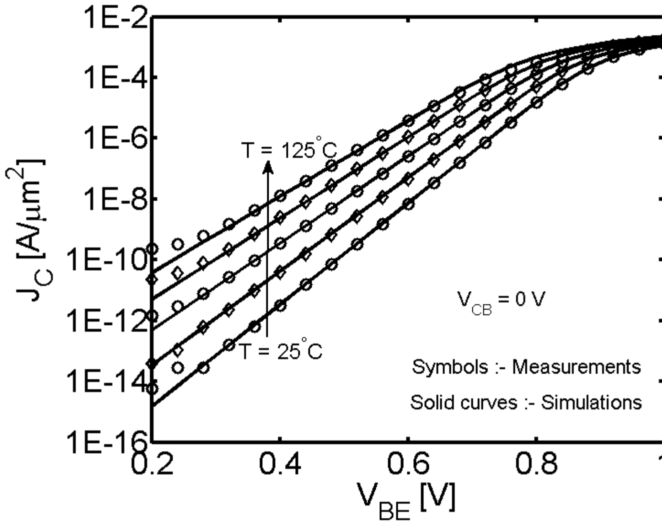
where $I_N(V_{B_2E_1}, V^*_{B_2C_2})$ is given by expression (4.3), other DC-measurements (as a function of both bias and temperature) need to be taken in order to determine first the input current I_N. As pointed out in the previous section, these measurements are essential in order determine the relevant input model parameters, and thus address all relevant physical effects that interfere with the temperature dependence of avalanche.

In Chapter 2 we presented a method for extracting $\mathbf{R}_{\mathrm{TH}}$ using the measured characteristics corresponding to the weak collector-base breakdown regime [5, 29], and it was visible that self-heating may as well be significantly interfering in the weak avalanche regime. Therefore, $\mathbf{R}_{\mathrm{TH}}$ needs to be extracted (accurately) first, before our extended version of the Mextram model (with the new non-local avalanche compact model) is employed. The Mextram simulation results as shown in this chapter, were performed on basis of values for the thermal resistance $\mathbf{R}_{\mathrm{TH}}$ as extracted by means of the method presented in Chapter 2 [5, 29]. We note that under the *weak avalanche* regime, as addressed by Mextram, the main transistor current I_N is approximately equal to the collector current I_C, in the forward normal operation of the transistor. The relevant model parameters, some of which are presented in the above expressions, were extracted from the DC-(temperature) measurements of the base-emitter and base-collector depletion capacitances, forward and reverse Gummel characteristics, forward and reverse Early characteristics, and output characteristics, as are summarized in Table. 4.1. A detailed definition of these Mextram model parameters can be found in e.g., [1, 8, 36]. The extraction process for each of these parameters [35] is left out of this thesis, but an example will be presented later in this section. In this table (in the second column), T_{ref} means that the corresponding Mextram parameters (in fourth column), were extracted at the chosen (fixed) reference temperature, while T means that those indicated parameters were extracted over all ambient temperatures. In the third column, we indicate the bias or current regime, as well as the employed sequence of the ambient temperatures $[T = (T_1, \cdots, T_5) = (25\,^\circ\mathrm{C}, \cdots, 125\,^\circ\mathrm{C})]$ over which such measurements were taken. The

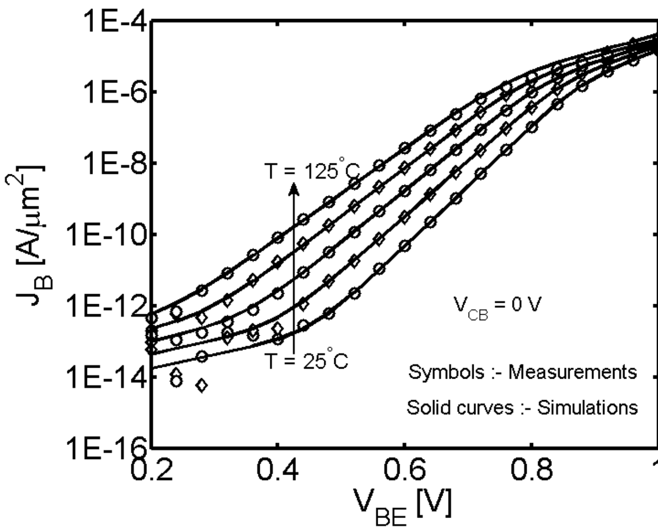
Table 4.1: Mextram model parameters which are relevant for our new non-local avalanche model verification. The first column represents the type of measurements taken, the second column the measured data over the regime described by the third column. The extracted model parameters are represented in the fourth column.

Characteristic	Measured data	Measurement regime	Extracted parameters
Base-emitter capacitance	C_{BE} [V_{BE}, T]	$V_{CB} = 0$ V, $T = T_1, \cdots, T_5$	C_{j_E}, V_{d_E}, p_E
Collector-base capacitance	C_{CB} [V_{CB}, T]	$V_{BE} = 0$ V, $T = T_1, \cdots, T_5$	C_{j_C}, V_{d_C}, p_C, X_P
Forward Gummel	I_C [V_{BE}, T_{ref}]	$V_{CB} = 0$ V, $T = T_{\text{ref}}$	I_s, V_{er}
	I_B [V_{BE}, T_{ref}]	$V_{CB} = 0$ V, $T = T_{\text{ref}}$	β_f, I_{Bf}, m_{Lf}
Reverse Gummel	I_B [V_{CB}, T_{ref}]	$V_{EB} = 0$ V, $T = T_{\text{ref}}$	β_{ri}, I_{Br}, V_{Lr}
	I_S [V_{CB}, T_{ref}]	$V_{EB} = 0$ V, $T = T_{\text{ref}}$	I_{Ss}
Forward Early	I_B [V_{CB}, T_{ref}]	$V_{BE} = 0.65$ V, $T = T_{\text{ref}}$	W_{avl}, V_{avl}
	I_C [V_{CB}, T_{ref}]	$V_{BE} = 0.65$ V, $T = T_{\text{ref}}$	V_{ef}
Reverse Early	I_E [V_{EB}, T_{ref}]	$V_{BC} = 0.65$ V, $T = T_{\text{ref}}$	V_{er}
Forward Gummel	I_C [V_{BE}, T]	$V_{CB} = 0$ V, $T = T_1, \cdots, T_5$	V_{g_B}
	I_B [V_{BE}, T]	$V_{CB} = 0$ V, $T = T_1, \cdots, T_5$	$dV_{g\beta f}$, V_{gj}
Reverse Gummel	I_B [V_{CB}, T]	$V_{EB} = 0$ V, $T = T_1, \cdots, T_5$	$dV_{g\beta r}$, V_{g_C}
	I_S [V_{CB}, T]	$V_{EB} = 0$ V, $T = T_1, \cdots, T_5$	V_{g_S}
Forward Early	I_C [V_{CB}, T]	$V_{BE} = 0$ V, $T = T_1, \cdots, T_5$	A_{QBO}
Output	I_S [V_{CE}, T_{ref}]	fixed I_B, $T = T_{\text{ref}}$	R_{Cc}, R_{Cblx}, R_{Cbli}
	I_c [V_{CE}, T_{ref}]	fixed I_B, $T = T_{\text{ref}}$	I_k

parameters corresponding to the measured Gummel plots, were extracted from either the low bias non-ideal region or the medium bias ideal region. In Fig. 4.4, we present an example of the measured and Mextram model simulated results of the collector and base current densities as a function of base-emitter voltage (V_{BE}) and temperature. From the results in Fig. 4.4, good fits between the measured and Mextram simulated results for the collector and base current densities, are achieved over all bias regimes and applied ambient temperatures. From these results, the relevant temperature dependent model parameters, as indicated in Table. 4.1, corresponding to the forward Gummel characteristics were extracted. Good model fits between the measured and Mextram simulated results were also achieved for the other relevant electrical characteristics (not included in this thesis); the relevant model parameters as indicated in Table. 4.1, were extracted accordingly, over bias and temperature.



(a)



(b)

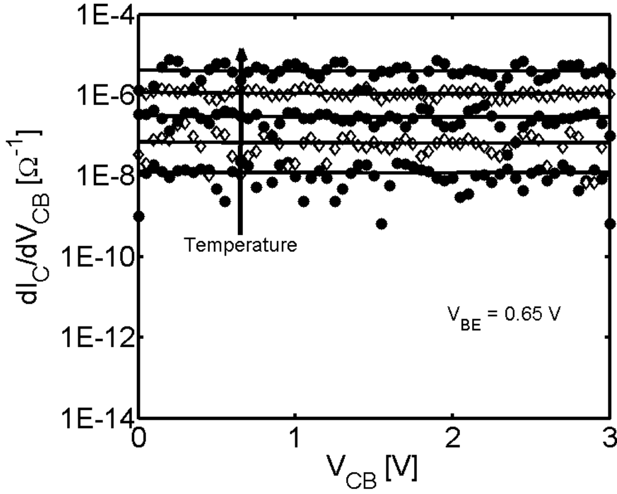
Figure 4.4: Measured (symbols) and Mextram model simulated (solid curves) collector current density J_C (a) and base current density J_B (b) as a function of base-emitter voltage (V_{BE}) and ambient temperature (T = 25, 50, 75, 100, and 125 °C), corresponding to the QUBiC4+ (NPN) SiGe-BTJ, with emitter area $A_E = 0.50 \times 20.7$ μm^2. A good agreement between the measured and simulated results can be observed over all bias regimes and ambient temperatures. From these forward Gummel plots, the corresponding Mextram model parameters as indicated in Table. 4.1, are extracted.

A specific interfering physical effect that we would like to point out, is the temperature dependence of the collector current through the Early effect. This involves the temperature dependent modulation of the depletion layers, and hence of the width of the neutral base. This leads to a modulation of the base Gummel number (due to variation of the integral limits) [25, 34], and hence the main transistor current I_N [33]. Since the weak avalanche current $I_{avl} = (M_n - 1) I_N$ results from impact-ionization induced by electrons that take part in I_N, then a temperature dependent modulation of I_{avl} is implied. In Mextram, the temperature dependence of the neutral base charge (at zero applied bias) Q_{B0} is represented by a model parameter $A_{Q_{B0}}$ (also presented in Table. 4.1). As can be seen from the above relations, $A_{Q_{B0}}$ affects a number of quantities, i.e., it appears in the temperature scaling rules for: saturation current (4.4), Early voltages (4.13) and (4.15), Ge induced bandgap difference (4.22), and forward current gain (4.23). In other words it directly affects the transistor main current I_N, and thus the avalanche current. This effect is also demonstrated in Fig. 4.7. In Fig. 4.5, we demonstrate how $A_{Q_{B0}}$ can be extracted from the measurements (symbols) and Mextram model simulations (curves) of the slope (dI_C/dV_{CB}) as a function of temperature, over the non-avalanche region (see Fig. 2.2), i.e., *independently of the avalanche effects*. The solid curves/lines in Fig. 4.5(a) shows the results corresponding to the optimized parameter value for $A_{Q_{B0}}$ using the measurements over temperature (i.e., $T = 25, 50, 75, 100$, and 125°C) and the Mextram model simulations. In Fig. 4.5(b), we show an example of the results corresponding to the optimized value, i.e., $A_{Q_{B0}} = 1.043$ (solid curve) for a fixed $V_{CB} = 1.0$V (taken from Fig. 4.5(a)) and variable temperature. In addition, the sensitivity to $A_{Q_{B0}}$ of the slope (dI_C/dV_{CB}) is demonstrated by cases corresponding to $A_{Q_{B0}} = 0$ and $A_{Q_{B0}} = 2$ (dashed curves in Fig. 4.5(b)). From these results, we can observe that the slope decreases as $A_{Q_{B0}}$ increases most especially for high ambient temperatures; as expected, for low temperatures, $A_{Q_{B0}}$ has a marginal effect on the slope (dI_C/dV_{CB}).

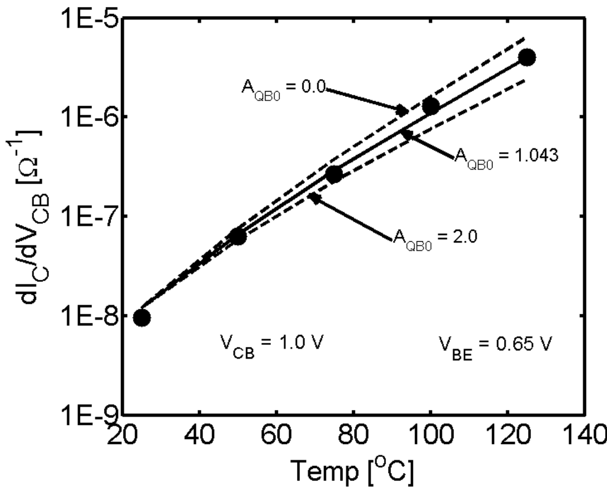
Based on the assumption that we have taken all other significant physical effects into account then, the remaining observed *differences* between the measured and simulated (using the Mextram *local* avalanche model [4, 8, 36]) avalanche characteristics, can be attributed to non-local avalanche effects. These effects are addressed by our new non-local avalanche compact model, both as a function of bias and temperature. The corresponding measurement and simulation results for the avalanche characteristics (both as a function of bias and temperature) will be presented in the next section, for the employed pure Si bipolar junction transistor.

4.4 Verification results for Si bipolar junction transistor

After addressing all the independent physical mechanisms that may interfere with the weak avalanche characteristics and their temperature dependence, we are now in position to address the influence of non-local avalanche effects on the measured avalanche characteristics of the QUBiC4+ (NPN) Si industrial bipolar transistor. Employing the experimental variant of the Mextram model (i.e., Mextram model [8] extended with our new non-local avalanche model [37]), we confront model simulations for our new non-local avalanche model with measured data taken on this Si-BJT. The corresponding model verification results, will be presented in this section. Note that unless stated otherwise,



(a)



(b)

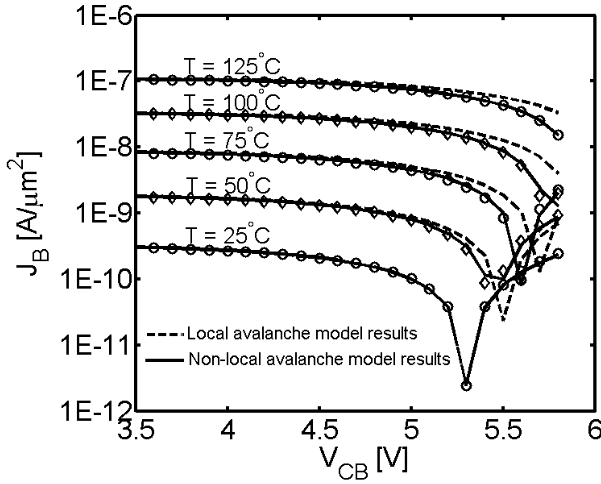
Figure 4.5: Independent extraction of the Mextram's model parameter $A_{Q_{B0}}$ for the temperature dependence of the neutral base charge, from the forward Early measurements of I_C as a function of ambient temperature (i.e., $T = 25, 50, 75, 100$, and $125\,^\circ$C), over the non-avalanche region. The symbols represent measured values of the slope (dI_C/dV_{CB}) and the curves are the corresponding Mextram model simulations. Figure (a) shows the slope as a function of collector-base voltage (V_{CB}) and ambient temperature; from this plot, $A_{Q_{B0}}$ is optimized over all temperatures using the measurements and Mextram model simulations. Figure (b) shows the results corresponding to the optimized parameter value for $A_{Q_{B0}}$ (solid curve) taken from (a) at V_{CB} = 1.0 V. The sensitivity to $A_{Q_{B0}}$ of the slope is also demonstrated by showing cases for $A_{Q_{B0}}$ = 0 and $A_{Q_{B0}}$ = 2.0 (dashed curves in (b)). In all these results, the base-emitter voltage is fixed to $V_{BE} = 0.65$ V.

the measurement results will be represented by *open symbols*, while the simulation results corresponding to our new non-local avalanche model will be represented by *solid curves*. For comparison purposes, best fitting results obtained with the standard *local* avalanche compact model of the Mextram model [4] will also be presented (by *dashed curves*) on the same plots.

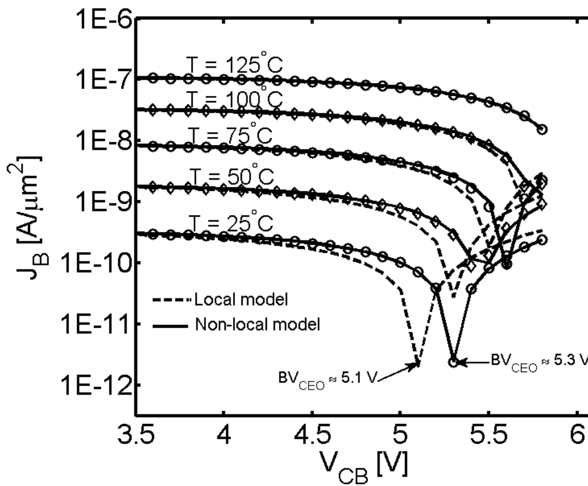
For the model parameter extraction process, first the avalanche model parameters, i.e., V_{avl}, W_{avl}, and λ_e are simultaneously optimized from the base current I_B measurements as a function of the collector-base voltage V_{CB} (over the weak avalanche region) at a chosen fixed target ambient temperature (T_{tgt}). Next, the temperature coefficient of the relaxation length non-local avalanche parameter A_{λ_e} is optimized from the I_B measurements as a function of V_{CB} (over the weak avalanche region) and over all employed ambient temperatures (i.e., $T = 25, 50, 75, 100$, and $125\,°\mathrm{C}$), including the initially chosen target temperature. In order to fully account the effect of self-heating on the final extracted model parameters, this employed parameter extraction process is repeated a number of times (iterated) until stable model parameters are achieved. We note that this effect of self-heating on the simulated characteristics can as well be taken into account by simultaneously optimizing all model parameters over all ambient temperatures, provided the employed software supports this strategy. In practice, this approach is faster, but has a number of limitations in terms of the capabilities of the employed software to carry out such global parameter optimization over multiple input data sets.

In all the simulation results for the non-local avalanche model, $T = 25\,°\mathrm{C}$ is taken as the target temperature, and the avalanche model parameters (i.e., V_{avl}, W_{avl}, and λ_e) are optimized first at this temperature. To enable comparison with the semiconductor device physics (e.g., [9, 10, 31, 38]), for simulation results of the local avalanche compact model, two target temperatures are considered, i.e., $T = 25\,°\mathrm{C}$ and $T = 125\,°\mathrm{C}$. In both cases, the electric field model parameters (i.e., V_{avl} and W_{avl}) are optimized first at each target temperature. This means that we will have two sets of results, i.e., one for each case of the target temperature for the local avalanche model (i.e., top figure corresponding to $T_{\mathrm{tgt}} = 25\,°\mathrm{C}$, while bottom figure to $T_{\mathrm{tgt}} = 125\,°\mathrm{C}$), in combination with the results for the non-local avalanche model.

The measurement (symbols) and simulation (curves) results for the avalanche characteristics as a function of collector-base voltage (V_{CB}) for a sequence of ambient temperatures, are presented in Fig. 4.6. According to these results, deviations can be observed between the best fitting results for the local compact model simulations (dashed curves) and the measurements (symbols) as a function of temperature and V_{CB}, over the weak avalanche region (i.e., $V_{CB} > 4.0$ V). The dashed curves in Fig. 4.6(a), correspond to local avalanche model results where the model parameters V_{avl} and W_{avl}, were optimized first at target temperature $T = 25\,°\mathrm{C}$, while those in Fig. 4.6(b), at target temperature $T = 125\,°\mathrm{C}$. The major observation from Fig. 4.6(a), is the limited fitting capability of the standard local avalanche model [8] for the measured avalanche characteristics, as a function of temperature. Here, the local model seems to underestimate avalanche (mostly at higher temperatures), and thus overestimates the base current (over the weak avalanche regime). Since the avalanche parameters were optimized first at $T = 25\,°\mathrm{C}$, where non-local effects are more significant, it implies that the electric field was underestimated as the corresponding model parameters were significantly comprised at the



(a)



(b)

Figure 4.6: Measured (symbols) and simulated (curves) avalanche characteristics. Dashed curves represent a local avalanche compact model simulation results, while solid curves represent a non-local avalanche compact model simulation results. Shown is the base current density (J_B) as a function of collector-base voltage (V_{CB}), for a sequence of ambient temperatures (i.e., $T = 25, 50, 75, 100$, and $125\,^{\circ}$C), as observed for a QUBiC4+ (NPN) Si-BJT, for fixed base-emitter voltage $V_{BE} = 0.65$V. In (a), the temperature independent local avalanche model parameters (i.e., V_{avl} and W_{avl}) are optimized first at target temperature $T = 25\,^{\circ}$C, while in (b) they are optimized first at $T = 125\,^{\circ}$C. From (a) and (b), we can observe that our non-local avalanche model yields excellent fits with the measurements, in contrast to the local compact model simulations (over the weak avalanche region, i.e., $V_{CB} > 4.0$ V). In addition, it can be observed from (b) that the local avalanche model overestimates the avalanche current and thus underestimates the breakdown voltage BV_{CEO}.

cost of achieving good model fits (dashed curves). This underestimation of the electric field, implies that the local avalanche is now modeled by an effective electric field that is lower than the real electric field. This same effective electric field is used to model avalanche process at high temperatures, where the physical avalanche will behave more like local avalanche. Thus, though the local compact model is still a local model, now it is operated by an effective electric field (that was optimized for operation at $T = 25\,^{\circ}$C) that underestimates the true field. As a result, the local compact avalanche model, then underestimates the avalanche current, which results into an overestimate of the base current (over the weak avalanche region) at higher temperatures (i.e, for $T > 25\,^{\circ}$C). This is indeed what we observe by the dashed curves in Fig. 4.6(a). On the other hand, when the local model parameters V_{avl} and W_{avl} are optimized first at a higher target temperature, i.e., $T = 125\,^{\circ}$C, as in Fig. 4.6(b), the physical avalanche process approximates best the local avalanche process. Due to the high temperatures, the crystal lattice vibrations are more intense and this will lead to a stronger randomizing impact on the mobile electron cloud, thus a short relaxation time and hence a shorter relaxation length (λ_e). This can be viewed as: the incident electrons travel a much shorter distance within the high field region before they gain the threshold energy to cause impact-ionization. Since λ_e is much shorter, the incident electrons can be assumed to gain the threshold energy instantaneously once they enter the high field collector-base depletion region, thus physical collector-base breakdown process approximates the local avalanche process. When the model is optimized against the highest temperature, the deviations between the local avalanche model simulations and the measured characteristics become apparent where they naturally belong, according to semiconductor device physics. Indeed, at lower temperatures, the energy relaxation length is longest, and non-local avalanche effects are the most significant. Due to non-local avalanche effects, avalanche current is weaker than predicted by the local avalanche model, hence the measured current at the base terminal is larger than predicted. This is indeed what we observe in Fig. 4.6(b) by the dashed curves. Here we see that due to the high electric fields induced by the local model (at lower temperatures), avalanche current is overestimated and this results into an underestimate of the collector-emitter breakdown voltage with an open base BV_{CEO} [9, 10, 38]. We note that, this will be the same presentation format for local avalanche model simulation results in Section 4.6 for SiGe HBTs, thus the same explanation will hold. It is noteworthy to remark that as demonstrated by Fig. 4.6, the present Si transistor technology (even non-RF technology) has evolved to the extent that *local* avalanche (even though appropriate temperature scaling of Chynoweth's law is applied [4]), can no longer capture the full family of avalanche characteristics over temperature.

Excellent fits between the measurement results (open-symbols) and the non-local avalanche model simulation results, as represented by the solid curves in Fig. 4.6, as a function of collector-base voltage and ambient temperature, were achieved. These non-local avalanche model results correspond to the extracted parameter value for the relaxation length $\lambda_e = 64.75$ nm, and the subsequently optimized value for the temperature coefficient $A_{\lambda_e} = -0.89$.

With reference to the adopted model verification ambition (as outlined in the Section 4.1), we note that our non-local avalanche model yields good fitting capabilities of the measured avalanche characteristics as a function of both bias and temperature. Ver-

ification of the physical content of our non-local avalanche model, however, requires us to test the uniqueness, accuracy, and precision of the extracted non-local avalanche model parameters. This will be addressed in detail in the next section on error analysis.

Example: Effect of neglecting the interfering physical mechanisms

We carried out a demonstration of the effect of neglecting the influence of the interfering physical mechanisms in the weak avalanche regime on the simulated results of the extended version of Mextram with non-local avalanche model. As an example, we compared the effect of model parameter $A_{Q_{BO}}$ on the simulation results; here we took values presented earlier in Fig. 4.5, i.e., $A_{Q_{BO}}$ = 1.043 (the extracted model parameter value), $A_{Q_{BO}}$ = 0.0, and $A_{Q_{BO}}$ = 2.0, and carried out the corresponding model simulations at T = 75 °C. The results are presented in Fig. 4.7, where the electric field model parameters for the local model (dashed curves) i.e., V_{avl} and W_{avl} are optimized at T = 25 °C. As demon-

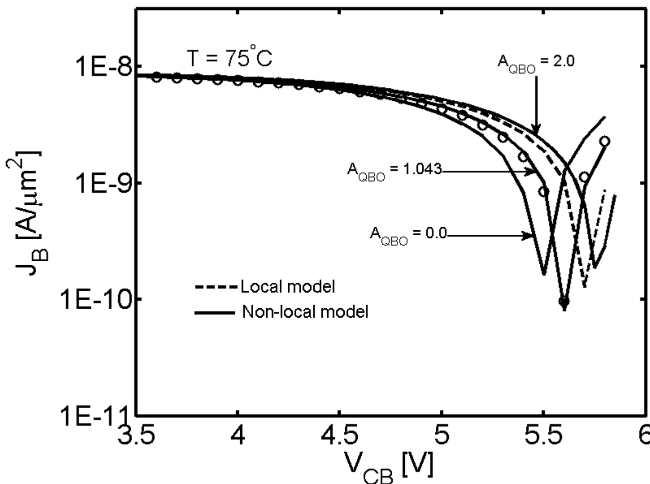


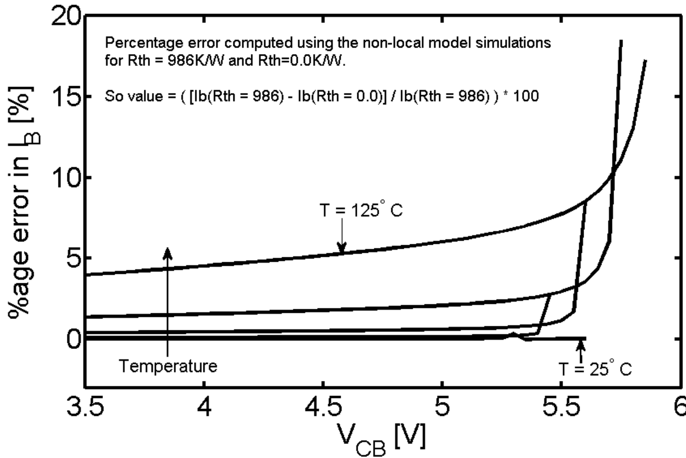
Figure 4.7: Measured (symbols) and simulated (curves) avalanche characteristics at different values of $A_{Q_{B0}}$ presented earlier in Fig. 4.5. Dashed curves represent a local avalanche compact model simulation results, while solid curves represent a non-local avalanche compact model simulation results. Shown is the base current density (J_B) as a function of collector-base voltage (V_{CB}), as observed for a QUBiC4+ (NPN) Si-BJT, for fixed base-emitter voltage V_{BE} = 0.65V at T = 75 °C. Large deviations between the non-local avalanche model simulations and the measured data (symbols) can be observed for $A_{Q_{B0}}$ = 0.0 and $A_{Q_{B0}}$ = 2.0; in contrast a good model fit is achieved when the true extracted model parameter $A_{Q_{B0}}$ = 1.043 is used.

strated earlier in Fig. 4.5, where $A_{Q_{B0}}$ = 0.0 and $A_{Q_{B0}}$ = 2.0 yielded an underestimate and overestimate of the slope, respectively, the same behavior can be observed in Fig. 4.7 where large deviations between non-local avalanche model simulations (solid curves) and measured data are observed (over the weak avalanche region, i.e., V_{CB} > 4.0 V) for $A_{Q_{B0}}$ = 0.0 and $A_{Q_{B0}}$ = 2.0. This is in contrast to extracted model parameter $A_{Q_{B0}}$ = 1.043, which yield good model fits with the measured data; this is the same value used for the simulation results (solid curves) in Fig. 4.6. We recall that the $A_{Q_{B0}}$ model parameter corresponds to the temperature dependence of the collector current through the Early

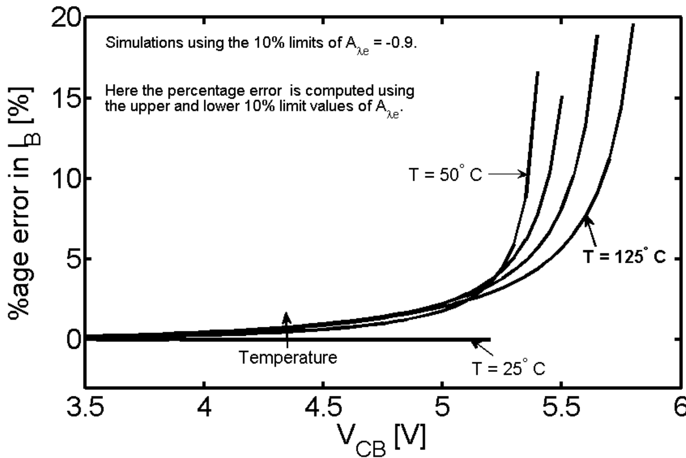
effect i.e., how Early effect interferes in the weak avalanche regime. Here, we can see a direct influence of this physical effect on the simulated results of the non-local avalanche compact model over the weak avalanche regime. From Fig. 4.7, parameter $\mathrm{A_{Q_{B0}}}$ appears to have a rather similar impact on the avalanche characteristics as the non-local model parameter A_{λ_e}. Hence, optimizing $\mathrm{A_{Q_{B0}}}$ and A_{λ_e} together would likely introduce large error margins in the estimated parameter values of both, to the point that the parameters values are actually dependent (i.e., an error in one parameter can be compensated for by another parameter) and hence undetermined. In this sense, a good fit to the measured data may still be achieved, but the parameters cannot be interpreted any longer, since they become undetermined. This demands for extraction of $\mathrm{A_{Q_{B0}}}$ independently of weak avalanche regime, over which A_{λ_e} is extracted.

Another example of the relevant interfering physical effect in the weak collector-base breakdown regime is self-heating. This effect was addressed for the results in Fig. 4.6 using the extraction method for the thermal resistance R_{TH} presented in Chapter 2, together with the already built-in Mextram thermal network [8]. This R_{TH} is then used to determine the increase in junction temperature, due to self-heating caused by the dissipated power in the device. For this device the extracted parameter value of $R_{TH} = 986$ K/W is used. During the model simulations, the self-heating effect can be switched off by taking $R_{TH} = 0$ K/W. We carried out the non-local avalanche model simulations, with and without self-heating; the from these results, we computed the percentage error in the base current I_B, and the corresponding results are presented in Fig. 4.8(a). From these results, we can see that the percentage error in I_B increases with temperature, both over the non-avalanche region and the weak avalanche region (i.e., $V_{CB} > 4.0$ V). For lower temperatures (i.e, $T = 25\,°$C and $T = 50\,°$C), the error is relatively small for both low and moderate V_{CB} values, which shows that for such temperatures, the effects of self-heating on the avalanche characteristics is relatively small. On the other hand, for relatively high temperatures (i.e., $T > 50\,°$C), the percentage error is relatively large and it increases significantly over the weak avalanche region. The level of significance of self-heating effect is based on the accuracy level of the non-local avalanche model parameters, i.e, λ_e and A_{λ_e} of 10% in relation to the published literature values for these quantities. This accuracy level for these parameters is discussed in detail in the next section. By considering the 10% accuracy level in the temperature coefficient $A_{\lambda_e} = -0.9$, we carried out non-local avalanche simulations for the lower (i.e, $A_{\lambda_e} = -0.99$) and upper (i.e, $A_{\lambda_e} = -0.81$) limits of this error interval; from these results, we computed the percentage error as a function of V_{CB} and temperature. The corresponding results are presented in Fig. 4.8(b). From the results in Fig. 4.8(b), we can observe that at ambient temperature $T = 25\,°$C, the computed error is very small ($\approx$ 0), but for $T > 25\,°$C, the error increases with temperature, with a sharp increase observed over the weak avalanche region. The maximum error is within the range of 15 to 20%, the same error margin is observed in Fig. 4.8(a); thus the significance of self-heating effect in relation to the targeted accuracy levels of λ_e and A_{λ_e}.

In summary, the results in Fig. 4.8 shows that even for low base-emitter voltage, i.e., $V_{BE} = 0.65$ V, self-heating effect is still significant in relation to the targeted accuracy level of 10% for the non-local avalanche model parameters λ_e and A_{λ_e}. Therefore, in



(a)



(b)

Figure 4.8: Percentage error in the simulated avalanche characteristics using the non-local avalanche model. The aim of these simulation results is to demonstrate the influence of self-heating effect on the avalanche characteristics in the weak avalanche regime. In Fig 4.8(a), the error is computed using the cases with and without self-heating, which is achieved by setting thermal resistance R_{TH} = 986 K/W and R_{TH} = 0 K/W, respectively. In Fig. 4.8(b), the percentage error is computed using the 10% error margin in the temperature coefficient $A_{\lambda_e} = -0.90$, i.e., the literature value. These results are a function of V_{CB} and temperature, for fixed base-emitter voltage $V_{BE} = 0.65$V.

order to achieve both good model fits to the measured weak avalanche characteristics and unbiased estimates for model parameters values which can be compared to their published literature values, the influence of self-heating on the measured characteristics in the weak avalanche regime has to be taken into account.

4.5 Error analysis of the non-local avalanche compact model

In the previous section, we demonstrated the physics based non-local avalanche compact model on a pure Si industrial bipolar transistor. Though this model showed good fitting capabilities with respect to the measured data as a function of both collector-base bias and ambient temperature, still we need to access the uniqueness, accuracy, and precision of the extracted non-local avalanche model parameter values. This will make our adopted model verification approach (mostly the physical nature of the extracted model parameters) complete, since it demonstrates how good our extracted non-local avalanche model parameter values correspond to the previously published literature values. For this analysis, we will employ the same temperature measurements for the avalanche characteristics taken on the pure Si QUBiC4+ (NPN) bipolar junction transistor, as that used in the previous section, though now with a different data treatment. Still ICCAP software [11], together with Levenberg-Marquardt optimization algorithm will be used in the model simulations. We will start by accessing the root mean square (RMS) error measure in the extracted model parameters. The magnitude of the effects of the numerical errors and stochastic errors in the measured data on the final extracted model parameter values will be discussed next, in the respective order.

4.5.1 Dependence of the RMS error measure on the model parameters

In general, the root mean square (RMS) measures the deviation between the measured data and the model simulations. For the employed simulation software (ICCAP [11]), the RMS error is a by-product of an optimization process for a given set of input model parameters, and the returned value (in case of model convergency) corresponds to an optimized set of model parameters that yields the minimum possible residual/deviations. We shall inspect the RMS error value, corresponding to the optimized set of model parameters, that minimizes the differences between the measured and simulated characteristics. Thus, the smaller the RMS error value, the better the model fits.

Using the experimental variant of Mextram model, we accessed the RMS error measure as a function of electric field Mextram model parameter $W_{\rm avl}$ which was stepped using a specified small step size. For each fixed input value for $W_{\rm avl}$, first the parameters $V_{\rm avl}$ and λ_e are simultaneously optimized from the base current (I_B) measurements as a function of the collector-base voltage (V_{CB}) at a chosen target temperature (i.e., T = 25°C for this case); the final extracted parameter values as well as the corresponding value for the RMS error are recorded. Next, the parameter A_{λ_e} is optimized from the I_B measurements as a function of V_{CB} and ambient temperature (i.e., T = 25, 50, 75, 100, and 125°C); the final extracted parameter value for A_{λ_e} together with the corresponding value for the RMS error [i.e., RMS error (A_{λ_e})] are recorded as well. In order to account for self-heating, this process is iterated until stable model parameters are achieved. This process is repeated for a sequence of fixed values for $W_{\rm avl}$. The final extracted results for the avalanche model parameter values (together with the corresponding values for the RMS errors) as a function of $W_{\rm avl}$, are presented in Fig. 4.9. From

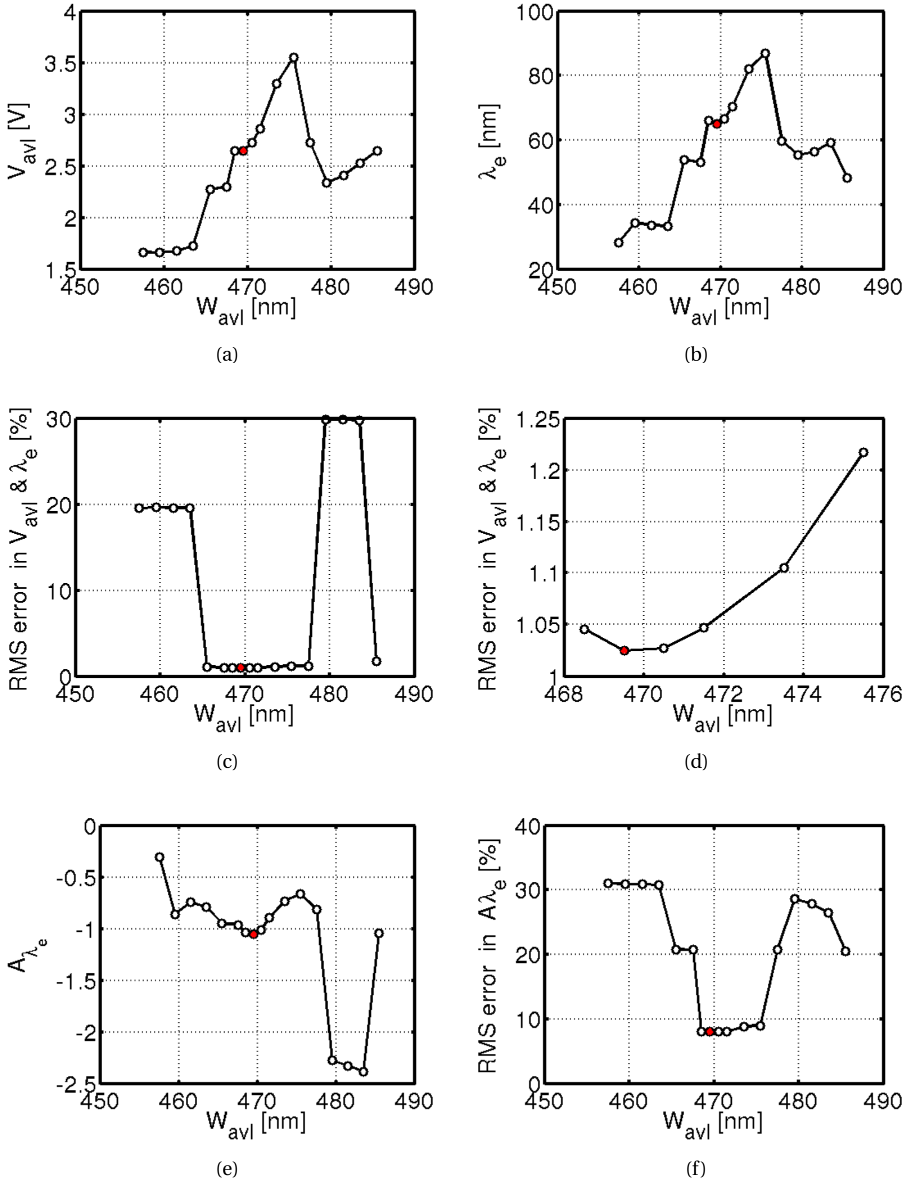


Figure 4.9: Dependence of the extracted avalanche parameter values for V_{avl} (a) and λ_e (b) on the electric field Mextram model parameter W_{avl}. The corresponding values for the root mean square (RMS) error in V_{avl} and λ_e (as a function of W_{avl}) are represented in figures (c) and (d); where (d) is the zoom-in of (c) around the range of values for W_{avl} where relatively small variations in the RMS error values are observed. The dependence of the extracted values for A_{λ_e} on W_{avl} and the corresponding values for the RMS error are presented in figures (e) and (f), respectively.

Figs. 4.9(a) and 4.9(b), we observe an irregular increasing trend in the extracted values for V_{avl} and λ_e as a function of W_{avl}. This is in contrast to what is expected from optimization of such a smooth model over a fixed data set. In fact for a well defined model, as the one we have, in a well defined optimization setting, the optimized V_{avl} and λ_e should be smooth functions of W_{avl}. In this numerical experiment, this observed behavior signifies the incapability of the employed software to accurately reproduce the expected smooth functional dependencies in the extracted values for V_{avl} and λ_e as a function of W_{avl}. Such irregularities can be attributed to the weakly imposed convergence criteria for the employed numerical algorithm. Such numerical artefact can in principle be decreased to any desired degree, but this would require extra computational resources, which in general may not be needed, since there would always be inaccuracies caused by the measurement stochastic errors. And as long as such stochastic errors dominate the the final results, the effect of numerical errors would be small enough to be considered negligible, hence this requires further quantification, which will be the subject of Subsection 4.5.2.

As we described earlier, at each value of W_{avl}, we get the RMS error value which measures the minimum deviation between the measured and simulated characteristics. With the minimized set of values for V_{avl} and λ_e plotted in Figs. 4.9(a) and 4.9(b), the corresponding RMS error values are plotted in Fig. 4.9(c) as a function of W_{avl}. We observe that for values of W_{avl} between 467.5 nm and 471.5 nm, the variations in the RMS error values are relatively small, i.e., in the range of 2.0%. This shows that if one were to plot RMS(W_{avl}, V_{avl}, λ_e) in a 3D parameter space, the RMS along the curve (W_{avl}, $V_{\mathrm{avl}}(W_{\mathrm{avl}})$, $\lambda_e(W_{\mathrm{avl}})$) would denote the height of the bottom of this curved surface, along W_{avl}. From Fig. 4.9(d), we observe a parabolic behavior of the RMS error, which implies that along the W_{avl} trajectory, the RMS error still has a well defined minimum. In short, these results show that if care is taken to optimize the model parameters with an initial condition sufficiently close to the minimum, the RMS error is nicely parabolic and the minimum of such is well defined. This indeed is the important conclusion from this experiment.

4.5.2 Influence of numerical accuracy of the employed simulation software on the final extracted model parameter values

In order to analyse the effects of numerical accuracy of the employed simulation software (i.e., ICCAP [11]) on the final extracted model parameter values, we carried out a simulation experiment, in which we perturbed W_{avl}, and for each value of W_{avl} we simultaneously optimized the model parameters W_{avl}, V_{avl}, λ_e, and A_{λ_e}, over all ambient temperatures (i.e., T = 25, 50, 75, 100, and 125°C). The corresponding extracted results (together with the RMS error values) as a function of the initial perturbed values for W_{avl} are presented in Fig. 4.10. The plots in Fig. 4.10 show a random variation of the extracted model parameter values as a function of the initial conditions for W_{avl}. Since everything was fixed, except the small perturbations in the initial conditions, this observed random variation in the extracted model parameters has to be a result of the numerical errors of the employed commercial simulation software. From the extracted model parameter values, we get $\lambda_e = 65.7 \pm 2.8$ nm and $A_{\lambda_e} = -0.92 \pm 0.09$, where we quote the mean values and two standard deviations [39].

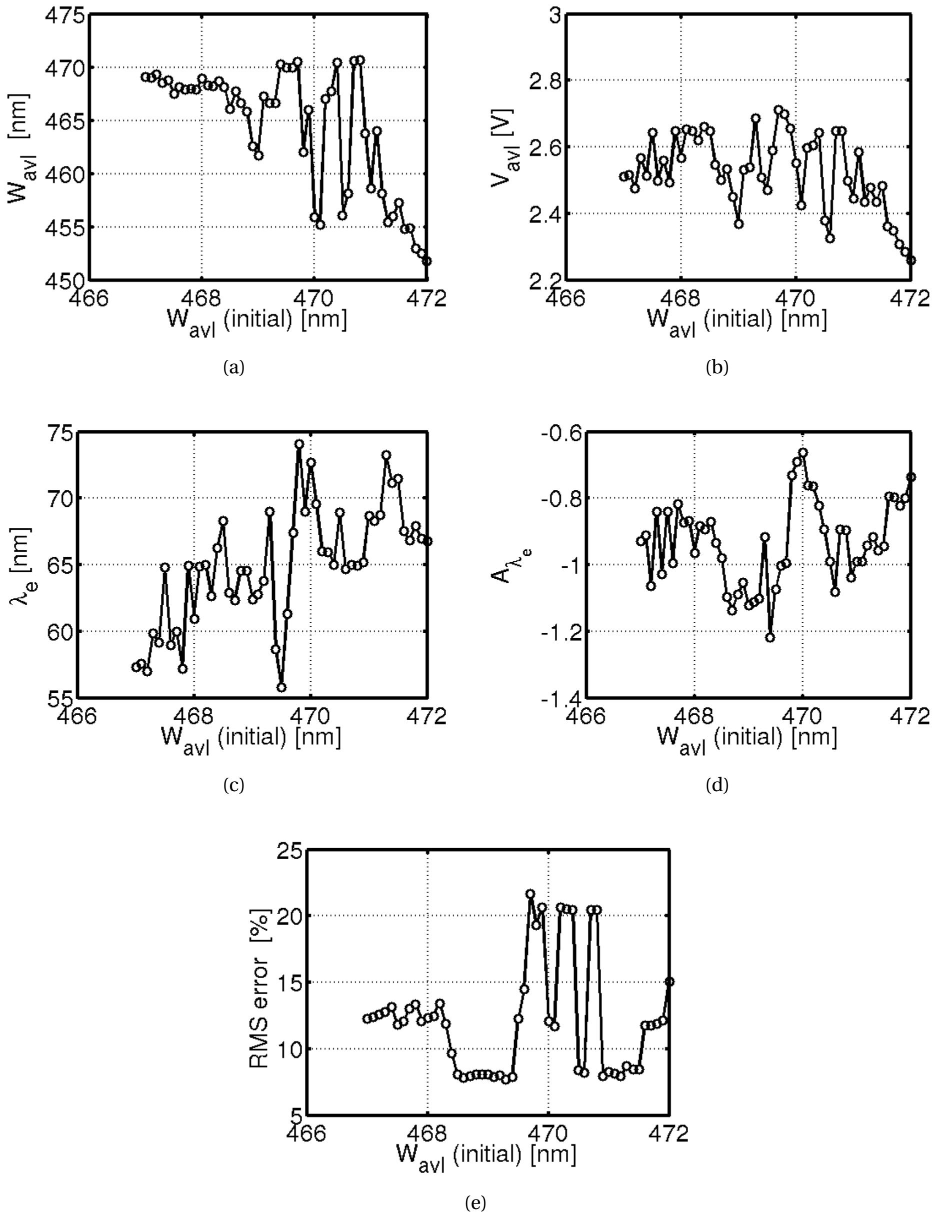


Figure 4.10: Dependence of the extracted avalanche parameter values [i.e., W_{avl} (a), V_{avl} (b), λ_e (c), and A_{λ_e} (d)] on the initial conditions for W_{avl}; and the corresponding values for the RMS error (e). A random variation of the extracted model parameter values can be observed as a function of the perturbed initial conditions.

4.5.3 Analysis of the effect of stochastic measurement errors on the final extracted model parameter values

We study the effect of the stochastic errors in the measured data on the final extracted avalanche model parameter values. Ideally, this would start with repeating all measurements as those described earlier on the same device a number of times, which would result into independent input measurement data sets. Carrying out such repetitions is beyond current resources though, thus for our case, we instead employed a numerical experiment that mimics such repeated measurements; we optimized the model parameters against different subsets of our measured data, and then observed the effect of such on the distribution of the final extracted parameter values. For total number of data points $p = 118$, we optimized against every n^{th} subset, where $n = 1, \cdots, 20$; with $n = 1$, corresponding to all points (whole data set), $n = 2$, corresponding to all even points, and so on. In other words, for each value of n, the corresponding subset contains every n^{th} point of the original data set.

In this simulation experiment, the model parameters, i.e., V_{avl}, W_{avl}, λ_e, and A_{λ_e} are simultaneously optimized over all ambient temperatures (i.e., T = 25, 50, 75, 100, and 125°C) for the different data subsets. The corresponding results are presented in Fig. 4.11, where the final extracted parameter values for λ_e and A_{λ_e} are plotted as a function of $(1/\sqrt{p})$, where p is number of data points in the various employed data subsets. From these plots, we observed a random variation in the final extracted parameters val-

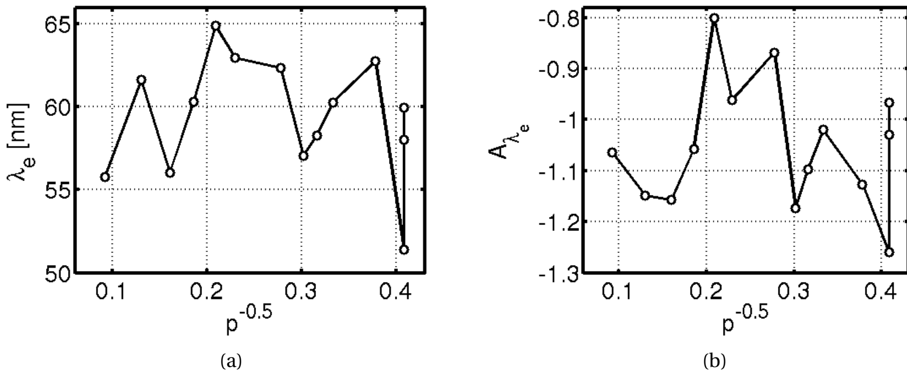


Figure 4.11: Dependence of the extracted non-local avalanche parameter values for λ_e (a) A_{λ_e} (b) on the inverse of the square root of the number of data points of the various employed subsets (i.e., $1/\sqrt{p}$). In this plot, we access the effect of the stochastic errors in the measured data on the final extracted model parameter values.

ues for λ_e and A_{λ_e} as a function of $(1/\sqrt{p})$. These can be attributed to the variations in the magnitudes of the stochastic errors in the input measurement data subsets. These extracted model parameter values yields $\lambda_e = 62.1 \pm 7.8$ nm and $A_{\lambda_e} = -1.08 \pm 0.30$; where we quote mean values and two standard deviations [39, 40] of the final parameter values optimized independently against different input measurement data points. We see that these error intervals contain values for silicon as reported in published semiconductor

device literature: $\lambda_e = 65.0$ nm [41] and $A_{\lambda_e} = -0.90$ [42, 43]. In addition, these intervals are wider than those observed in the case of the numerical errors; which implies a dominant influence of the stochastic errors in the measured data on the final extracted model parameter values.

We also verified the physical nature of the extracted electric field model parameters $\mathrm{V}_{\mathrm{avl}}$ and $\mathrm{W}_{\mathrm{avl}}$, by employing Mextram's physical interpretation of these avalanche parameters. According to Mextram model [1, 4, 8], these model parameters are defined so as to obey the relation

$$N_{\mathrm{epi}} = 2\frac{\varepsilon \mathrm{V}_{\mathrm{avl}}}{q_e {\mathrm{W}_{\mathrm{avl}}}^2}, \tag{4.26}$$

where ε is the permittivity and N_{epi} is the collector epilayer doping concentration. We found that the value of N_{epi} calculated from (4.26) using the extracted parameter values for $\mathrm{V}_{\mathrm{avl}}$ and $\mathrm{W}_{\mathrm{avl}}$ corresponding to all data points, equals that as independently estimated from capacitance-voltage [C(V)] characteristics [44] to within about 20%. This error margin can be attributed to the uncertainty in our knowledge about the effective collector-base junction area of the devices under consideration. These results shows that our new non-local avalanche compact model, also yields extracted electric field parameters that are in agreement with semiconductor device physics, which supports further that our non-local avalanche model captures presupposed underlying semiconductor device physics.

4

4.5.4 Summary

Based on an experimental variant of the Mextram model (Mextram model [1] extended with our new non-local avalanche model), we studied the error margins in the extracted avalanche model parameters and the effects of the various error sources on the final extracted model parameter values. This analysis was carried out using (temperature) measured data taken on a QUBiC4+ Si-BJT device, presented earlier in Section 4.3.

By considering a sequence of values for the electric field Mextram model parameter W_{avl} and then optimizing other model parameters, i.e., V_{avl}, λ_e, and A_{λ_e}, at each stepped value of W_{avl}, we accessed the quality of the root mean square (RMS) error measure. Here, we employed the argument that, the plot of RMS(W_{avl}, V_{avl}, λ_e) in a 3D parameter space would give a curved surface. Then, the plot of the RMS error re-optimized along the curve (W_{avl}, $V_{\mathrm{avl}}(W_{\mathrm{avl}})$, $\lambda_e(W_{\mathrm{avl}})$), gives the height of the bottom of this curved surface. We showed that the RMS error is parabolic and the minimum is well defined. Irregularities were observed in the plots for V_{avl} and λ_e as a function of W_{avl}; these are attributed to numerical artefact caused by the employed simulation software and the convergence criteria of the optimization algorithm. Such numerically related effects can be minimized to any desired degree (once one has access to the original code of the commercial software), but this demands more computational resources, and it is not needed since there is no need to be numerically more accurate than the error caused by the stochastic component in the measured data. Indeed, by a simulation experiment, which involved taking small perturbations of the initial input conditions of the software, using W_{avl} model parameter, we showed that the effect of the numerical errors on the

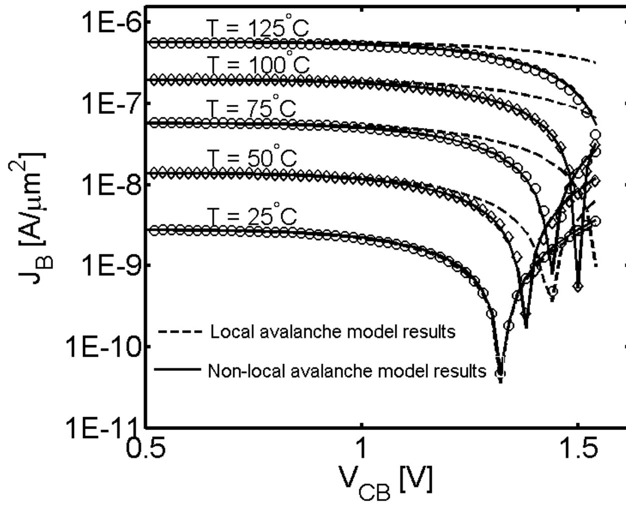
final model parameters to be relatively small, compared to the effect of stochastic errors in the measured data on the extracted model parameters. The latter was quantified by optimizing the model parameters against different subsets of the measured data. The resulting error intervals for the extracted non-local avalanche model parameter values are $\lambda_e = 62.1 \pm 7.8$ nm and $A_{\lambda_e} = -1.08 \pm 0.30$, where we quote mean values and two standard deviations [39, 40]. As these intervals contained the independently published literature values, i.e, $\lambda_e = 65.0$ nm [41] and $A_{\lambda_e} = -0.9$ [42, 43, 45], we observe that our new non-local avalanche model yields values for model parameters λ_e and A_{λ_e}, which are in agreement with the published semiconductor device physics values. This we present as a strong support for our claim that our model actually captures the presupposed underlying device physics.

4.6 Verification results for SiGe HBTs

In the previous model verifications, a pure Si-BJT was employed, but to test the application of our extended version of Mextram (with non-local avalanche compact model) to advanced bipolar technologies, we extended the model simulations to DC-measurements taken on modern industrial SiGe-HBTs. Here, the NPN devices used are the *high speed* QUBiC4X SiGe-HBT [46], the *high voltage* SiGe-HBT, and the *very high voltage* SiGe-HBT. Due to the presence of Ge in the neutral base of these devices, enhanced extra effects are induced on the Early effect [47] and thus on the collector current, through the base Gummel number [38, 44, 48, 49]. The physical nature of these effects is not well captured by the standard Mextram compact model, as an empirical relation is used to estimate such effects [8]. For this reason, our major emphasis in this section will be to verify our non-local avalanche model on the fitting capabilities against the measured avalanche characteristics, as a function of both bias and temperature. Of course we will also take a keen observation of the corresponding values for the non-local avalanche model parameters, i.e., λ_e and A_{λ_e}.

4.6.1 Results corresponding to a high speed QUBiC4X SiGe-HBT

The employed high speed QUBiC4X (NPN) SiGe-HBT device has the (effective) emitter area $A_E = 0.30 \times 20.7\ \mu m^2$, $BV_{CEO} \approx 1.3$ V, and $f_T/f_{\max}(V_{CE} = 1.5V) = 120/150$ GHz at T = 25 °C [46]. The experimental variant of the Mextram model, extended with our non-local avalanche model was applied on the DC-temperature measurements taken on this device (with terminal currents measured at the base, emitter, collector, and substrate terminals). For comparison purposes, the standard Mextram's local avalanche model [1] was also applied on the same measured data (at different reference temperatures as employed in Fig. 4.6). The corresponding results for the measured and simulated avalanche characteristics as a function of the collector-base voltage V_{CB} and ambient temperature are presented in Fig. 4.12. These results in Fig. 4.12 are analogous to those in Fig. 4.6, though now corresponding to a modern high speed QUBiC4X SiGe-HBT. The difference between the best achievable results based on the *local* avalanche model (dashed curves)



(a)

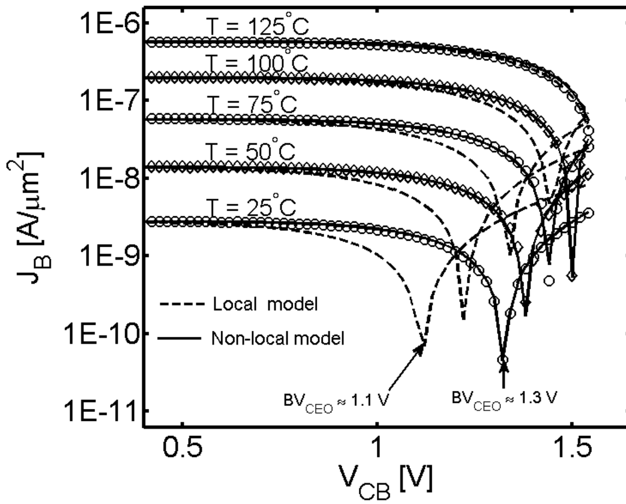


Figure 4.12: Measured (symbols) and simulated (curves) avalanche characteristics. Dashed curves represent a local avalanche compact model simulation results, while the solid curves represent a non-local avalanche compact model simulation results. Shown is the base current density (J_B) as a function of the collector-base voltage (V_{CB}), for a sequence of ambient temperatures (i.e., $T = 25, 50, 75, 100$, and $125\,°$C), as observed for a modern high speed QUBiC4X (NPN) SiGe-HBT, for fixed base-emitter voltage $V_{BE} = 0.65$V. In figure (a), the temperature independent local avalanche model parameters (i.e., V_{avl} and W_{avl}) were optimized first at $T_{\text{tgt}} = 25\,°$C, while in (b) they were optimized first at $T_{\text{tgt}} = 125\,°$C. From these results, we observe that the non-local avalanche model yields excellent fits with the measured data, in contrast to the local avalanche compact model (over the weak avalanche region, i.e., $V_{CB} > 0.8$ V).

and results from our *new non-local* avalanche model (solid curves) are now much *more distinct*. With our new non-local avalanche model yielding good fits with the measured data (symbols), in contrast to the local avalanche model.

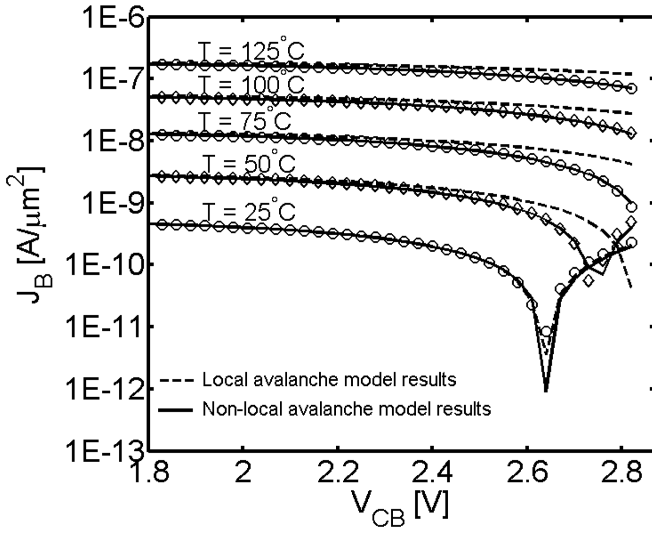
The key result as observed from Fig. 4.12 lies in the fact that Mextram 504 model extended with our non-local avalanche model can adequately reproduce the family of measured characteristics, also in the case of advanced SiGe HBTs. More results corresponding to SiGe bipolar devices will be presented in the following subsections. Our hypothesis that the difference between the dashed and solid curves is due to non-local avalanche effects is strongly supported by the extracted values for the non-local avalanche model parameters. Indeed for the relaxation length we found $\lambda_e = 64.31\,\text{nm}$, which is comparable to $\lambda_e = 65.0\,\text{nm}$ published earlier in [41], and for the temperature coefficient we found $A_{\lambda_e} = -1.034$, which is also comparable to the value $A_{\lambda_e} = -0.90$ reported in [42, 43, 45].

As observed and explained earlier in Section 4.4, from Fig. 4.12(a), the local avalanche model results (which were first optimized at target temperature $T = 25\,°\text{C}$), seem to underestimate avalanche (mostly at high temperatures), and thus to overestimate the base current (over the weak avalanche region); also for this high speed device. In contrast, when the local avalanche model parameters are optimized first at a higher target temperature, i.e., $T = 125\,°\text{C}$, as presented in Fig. 4.12(b), the local model overestimates avalanche, and thus underestimates BV_{CEO}; which is in agreement with semiconductor device physics [9, 50]. Similar conclusions as those drawn from the results in Fig. 4.6, can also be drawn from the simulation results in Fig. 4.12 in relation to our adopted non-local avalanche model verification strategy (mostly the fitting capabilities). That is to say, our non-local avalanche model simulations achieved good fits with the measured avalanche characteristics, as a function of both V_{CB} and temperature. Also, the corresponding (extracted) non-local avalanche model parameters are related to the independently published literature values.

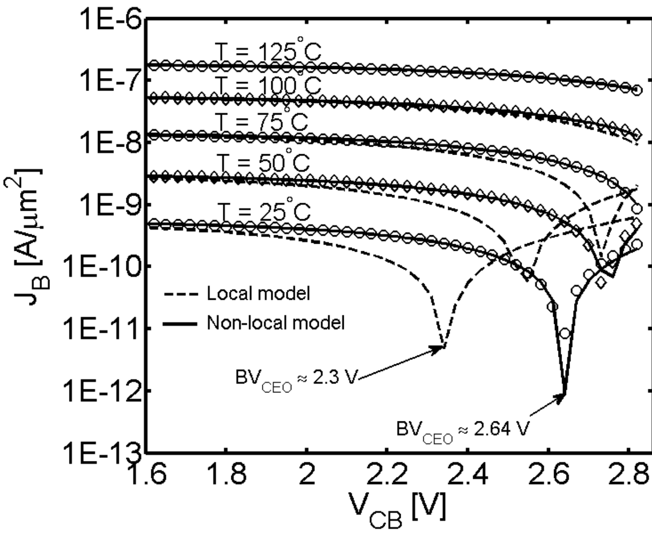
It was reported in [51] that in SiGe HBTs, excessive neutral base recombination significantly increases the total base current of the device, which in the end leads to a decrease in the device current gain. This would be due to the fact that in such SiGe devices, the base is highly doped, which increases the chances for Auger-recombination in the neutral base [52]. But from our analysis using the experimental variant of the Mextram model, this effect was negligible. That is to say, excellent model fits (solid curves in Fig. 4.12) were achieved for negligible neutral base recombination, correspond to the setting of the Mextram model parameter $X_{REC} = 0$ [1, 8] during the model simulations. This does not come as a surprise since such modern high speed QUBiC4X SiGe devices have a highly doped narrow base whose width is much smaller than the carrier recombination length [51]; thus neutral base recombination is highly minimized in such devices.

4.6.2 Results corresponding to a high voltage SiGe-HBT

In contrast to the model verification results presented in previous subsection, where a high-speed SiGe-HBT was used, in this subsection a *high voltage* SiGe-HBT device is employed. This NPN device has the drawn emitter size $A_E \approx 0.35 \times 5.2\,\mu\text{m}^2$ and $f_T/f_{\max} \approx 40/70\,\text{GHZ}$ at $\text{T} = 25\,°\text{C}$. In addition, this device does not have a collector sub-



(a)



(b)

Figure 4.13: Measured (symbols) and simulated (curves) avalanche characteristics (fully analogous to Fig. 4.12), here for high voltage (NPN) SiGe-HBT device, for a fixed base-emitter voltage $V_{BE} = 0.65$ V. In figure (a), the temperature independent local avalanche model parameters (i.e., V_{avl} and W_{avl}) were optimized first at $T_{\rm tgt} = 25°$C, while in (b) they were optimized first at $T_{\rm tgt} = 125°$C. Excellent fits with the measurements are achieved for simulations results corresponding to our new non-local avalanche model (solid curves), in contrast to those corresponding to the local avalanche compact model (dashed curves).

strate, meaning that it is more likely to experience high self-heating during the measurement process for the relevant electrical characteristics. Thus, accurate extraction of the thermal resistance (R_{TH}) is necessary in order to achieve meaningful model verification results for the avalanche characteristics. The results for the measured and simulated avalanche characteristics are presented in Fig. 4.13; and they are analogous to those presented in the previous subsection, though now corresponding to a high voltage SiGe-HBT device. From Fig. 4.13, deviations between the best achievable simulation results based on the local avalanche compact model (dashed lines) and the measured results (symbols) are clearly observed over the weak avalanche region (i.e., $V_{CB} > 1.5$V). On the other hand, excellent fits between the measured results and simulated results using our new non-local avalanche model (solid curves) were achieved for extracted value of the relaxation length $\lambda_e = 63.75$ nm and that of its temperature coefficient $A_{\lambda_e} = -0.81$. These extracted non-local avalanche model parameter values correspond pretty well with $\lambda_e = 65.0$ nm and $A_{\lambda_e} = -0.90$, as published in, e.g., [41] and [43, 45], respectively. Thus similar conclusions as those drawn from the results in Fig. 4.12, can also be drawn from the simulation results in Fig. 4.13, in relation to our adopted non-local avalanche model verification strategy.

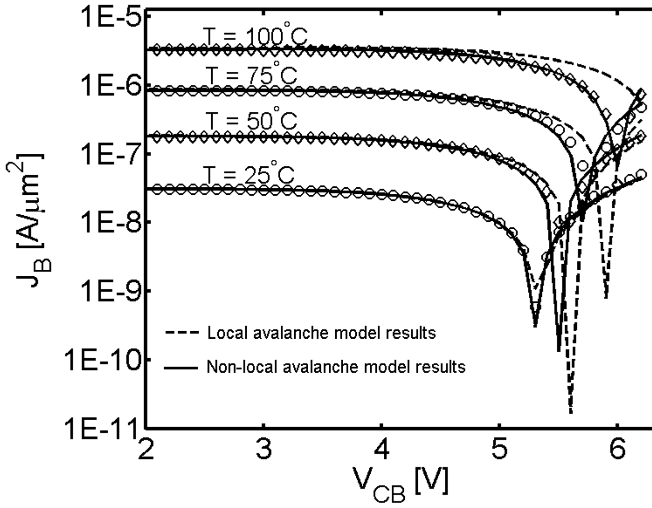
4

The same local avalanche model behavior as that observed in Fig. 4.12, can also be observed here, when the local model parameters are optimized first at target temperature $T = 25\,°$C (Fig. 4.13(a)), and at $T = 125\,°$C (Fig. 4.13(b)). From the simulation results in Fig. 4.13(b), we can observe that the standard local avalanche model [36] overestimates the avalanche current, which leads to underestimation of the collector-emitter breakdown voltage with an open base BV_{CEO}, this is in agreement with semiconductor device physics, e.g., [9, 53].

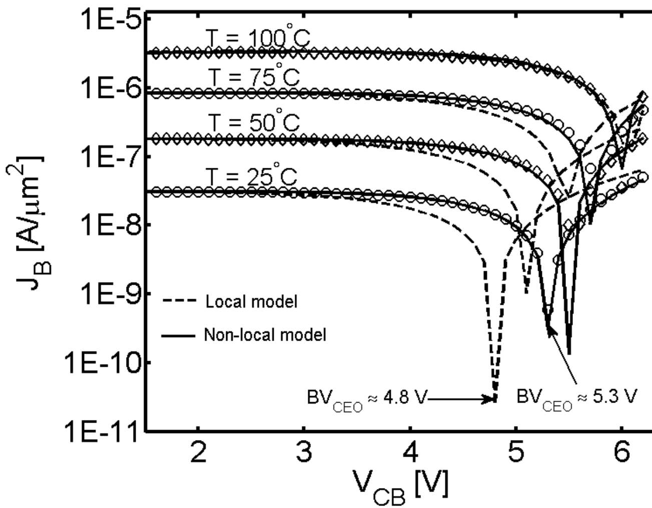
The major difference in relation to the results in Fig. 4.12, is that now smaller deviations between the measurement (symbols) and the local model simulation (dashed curves) results can be observed. That is to say, the non-local avalanche effects are less pronounced in this high voltage SiGe-HBT than in the high-speed QUBiC4X SiGe-HBT. This will be discussed further in Section 4.7. Model verification results corresponding to a *very high* voltage device technology will be presented in the next subsection.

4.6.3 Results corresponding to a very high voltage SiGe-HBT

After observing from the previous model verification results (both for Si and SiGe industrial bipolar transistors) that the extracted parameter values for the relaxation length (λ_e) and its temperature coefficient (A_{λ_e}) agree pretty well with $\lambda_e = 65.0$ nm and $A_{\lambda_e} = -0.90$, as published in [41] and [42, 43, 45], respectively, we decided to apply these values for the simulations of our non-local avalanche model, corresponding to a *very high* voltage (NPN) SiGe-HBT device. This device has got the emitter area $A_E \approx 0.4 \times 3.2\ \mu m^2$ and $f_T / f_{max}(V_{CE} = 3V) \approx 19/60$ GHz at T = 25 °C. In this case, the non-local avalanche model parameters (λ_e and A_{λ_e}) are taken as material coefficients. Since this device does not have a collector-substrate, the effect of device temperature (through self-heating) on the measured electrical characteristics is significant. This was observed by a sharp increase of the measured base current I_B for high ambient temperatures (i.e., T > 100 °C), even for the non-avalanche regime (i.e., very low V_{CB}). For this reason, we opted to ignore



(a)



(b)

Figure 4.14: Measured (symbols) and simulated (curves) avalanche characteristics as a function of collector-base bias V_{CB} and ambient temperature (i.e., $T = 25, 50, 75, 100$, and $125\,^\circ$C), here for very high voltage (NPN) SiGe-HBT, for a fixed base-emitter voltage $V_{BE} = 0.65$V. In figure (a), the temperature independent local avalanche model parameters (i.e., V_{avl} and W_{avl}) were optimized first at $T_{\text{tgt}} = 25\,^\circ$C, while in (b) they were optimized first at $T_{\text{tgt}} = 125\,^\circ$C. The non-local avalanche model simulation results (solid curves) correspond to standard parameter values published in semiconductor device literature, which yield good model fits with the measured data, in contrast to the local model simulation results (dashed curves).

the results corresponding to T = 125 °C in this model verification process. As in the previous subsections, we carried out the local and non-local avalanche model simulations, against the measured avalanche characteristics, and the corresponding results are presented in Fig. 4.14.

From the results for the avalanche characteristics as presented in Fig. 4.14, deviations between the best achievable results based on the local avalanche compact model (dashed curves) and the measured results (symbols) can be clearly observed over the weak avalanche region (i.e., $V_{CB} > 2.5$V). On the other hand, good fits between the measured and simulated results using our non-local avalanche compact model (solid curves) were achieved for the employed value of the relaxation length $\lambda_e = 65.0$ nm and its temperature coefficient $A_{\lambda_e} = -0.90$.

4

Note that for this device, the extracted non-local avalanche model parameter values are: $\lambda_e = 68.26$ nm and $A_{\lambda_e} = -0.78$, and they deviate a bit from the previous employed values from semiconductor device literature. Hence indeed the parameter values are poorly determined in this case.

4.6.4 Summary

In this section, we presented the verification results for the extended version of Mextram model (Mextram with our new non-local avalanche compact model), using measured data taken on present day modern industrial SiGe HBTs; with an emphasis on testing the model fitting capabilities against the measured data. Here, the employed RF SiGe HBTs are a *high-speed* QUBiC4X device [46], a *high voltage* device, and a *very high voltage* device. The extended version of Mextram model, was optimized against the measured avalanche characteristics. For comparison purposes, the standard Mextram local avalanche model [4, 8] was also optimized against the same measured data.

From the measurement and model simulation results, deviations were observed between the measured data (symbols) and the simulated results of the local avalanche model (dashed curves) as a function of temperature, for all employed devices. In contrast, good fits were achieved between the measured data and the simulated results of our new non-local avalanche model (solid curves), as function of both bias and temperature. In addition to the achieved good model fits with the measured data, the extracted values for the non-local avalanche model parameters (i.e., λ_e and A_{λ_e}) are related to the published semiconductor device physics literature values (i.e., $\lambda_e = 65.0$ nm [41] and $A_{\lambda_e} = -0.90$ [42, 43]). For the case of the *very high* voltage SiGe-HBT, these literature values for the non-local avalanche model parameters (now taken as material coefficients) were employed in the non-local avalanche model simulations, good model fits to the measured avalanche characteristics were achieved. This confirms further that λ_e and A_{λ_e} are indeed material coefficients, thus published literature values suffice.

A comparison of the magnitude of the observed deviations between the measured data (symbols) and the local avalanche model simulation results (dashed curves), showed that non-local effects are more significant in the high-speed SiGe-HBT device than in the high voltage SiGe-HBT devices or pure Si-BJT; this will be discussed in detail in the next section.

4.7 Magnitude of non-local avalanche effect in different device types

By comparing the results in Fig. 4.6, for a pure QUBiC4+ Si-BJT with those in Fig. 4.12, which correspond to a *high speed* QUBiC4X SiGe-HBT (over the weak avalanche regime), we observe (from the dashed curves) that non-local avalanche effects are more significant in the *high speed* SiGe-HBT device. Also, if we compare the results for the *high speed* QUBiC4X SiGe-HBT device with those corresponding to the *high voltage* SiGe-HBT device (i.e., in Fig. 4.13) and *very high voltage* SiGe-HBT device (i.e., in Fig. 4.14), we again see that non-local avalanche effects are more pronounced in the *high speed* QUBiC4X SiGe-HBT device. This observed variation in the magnitude of the non-local avalanche effects in relation to bipolar device technologies will be discussed in detail next, and we will focus on SiGe-HBT devices.

From bipolar transistor device technology [23, 34, 52, 54–57], the doping concentration of the collector epilayer (N_{epi}) in relation to the base doping concentration, plays a significant role in the classification of bipolar transistors as *high speed* or *(very) high voltage* devices. In this, high speed bipolar devices have a highly doped (thin) epilayer, while high voltage devices have a low doped (thick) epilayer. Here we assume that the transistor operates in forward normal operation [44] (V_{BE} is taken moderate to minimize high current effects), with the emitter more highly doped than the base and the collector epilayer. The epilayer doping concentration directly affects the magnitude of the electric field $E(x)$ across the base-collector depletion region. This dependence of the electric field on the epilayer doping (for uniform N_{epi}) is described by Poisson's equation [52] (for a mono-junction due to the high doped (SiGe) base) in the presence of charges

$$\frac{dE(x)}{dx} = \frac{q_e}{\epsilon} N_{\text{epi}} \left(1 - \frac{I_{\text{epi}}}{I_{\text{hc}}}\right), \tag{4.27}$$

where $I_{\text{hc}} = q_e A_{em} N_{\text{epi}} v_{\text{sat}}$ is the hot-carrier collector current. This relation shows that the gradient of $E(x)$ is directly proportional to N_{epi}, which means that high-speed bipolar devices such as QUBiC4X (with high N_{epi}) have higher magnitudes of the field gradient $dE(x)/dx$ in comparison to the high voltage devices (with low N_{epi}). Hence, high speed devices have a more narrow sharply-peaked electric field distribution (whose maximum value is located within the collector-base depletion region) in comparison to the high voltage devices [42], with a wider electric field distribution. In addition, since high speed bipolar devices have a high doped epilayer in comparison to the high voltage devices, at a given value of the collector-base junction bias, the epilayer of such high speed devices will be less depleted than that of high voltage devices [7]. Hence, in high voltage devices, the distance which the charge carriers (electrons) can travel so as to gain enough/steady energy to cause impact ionization is relatively small in comparison to the available collector-base junction depletion width. This means that carriers travel a considerable distance within the depletion region with the new steady kinetic energy capable of causing impact-ionization. Therefore, the non-local avalanche effects in such devices, behaves to a large extent as local avalanche effects. This is in contrast to high-speed devices with narrow depletion width, since the energy relaxation length λ_e is comparable to the avail-

able collector-base junction depletion width [9]. This implies that before impact-ionization starts, a considerable portion of the depletion width has been covered by the incident charge carriers [42], thus resulting into 'partial' impact-ionization. This makes non-local avalanche effects to be more significant in such high-speed (highly scaled) bipolar devices. Though the magnitude of non-local effect differ with device technology, our experimental assessment has shown it to be significant in all of them, and the new non-local avalanche compact model can sufficiently address such effect in different device technologies, both as a function of bias and temperature.

4.8 Summary

In this chapter, we presented an extensive experimental assessment of the significant and relevant physical mechanisms that are present in the weak collector-base breakdown regime of present day Si/SiGe industrial bipolar transistors, and do affect the terminal characteristics (i.e., lead to discrepancies between the measured and simulated family of characteristics as a function of bias and temperature). For this, we adopted a more ambitious model verification procedure of the extended version of Mextram, in a way that not only do we demonstrate the adequate model fitting capabilities in relation to the measured avalanche characteristics, but also show that our model actually captures the presupposed underlaying semiconductor device physics. By employing Mextram's extensive capabilities and dedicated parameter extraction process, we were able to take into account all other relevant physical effects that are significant in the weak avalanche regime, and affect the observed terminal characteristics. These were taken into account by considering the fact that the measured characteristics over temperature in the weak avalanche regime are also significantly affected by other temperature effects, through device self-heating; thus the need for adequate extraction of the thermal resistance R_{TH} [5] for each employed device. Using measured data in the same weak breakdown regime, a good estimated for R_{TH} was achieved by employing the method discussed in Chapter 2, this was used to address self-heating effect. One of these physical mechanisms we exhaustively discussed (and demonstrated), is the temperature dependence of the collector current through the Early effect, which was addressed via an independent extraction of the temperature model parameter corresponding to the zero-bias neutral base charge [8]. We also demonstrated the significance of self-heating effect on the measured characteristics in the weak collector-base breakdown regime, even for lower V_{BE} bias conditions, and thus the need to take such physical effect into account.

We showed that once all relevant and significant physical mechanisms in the weak collector-base breakdown regime are adequately taken into account, then the Mextram model extended with our new non-local avalanche compact model can accurately reproduce Si-BJT and SiGe-HBT weak avalanche characteristics as a function of both bias and temperature. The extended model also yields values for model parameters λ_e and A_{λ_e}, which are in agreements with the published semiconductor device physics values e.g., [41, 42]. Hence, a strong support for our claim that our model actually captures the presupposed underlying device physics. This confirms both the adequacy and validity

of the physical basis of our model formulation. We note that in most available literature e.g. [9, 10], avalanche measured data is normally represented in form of plots for multiplication factor $(M_n - 1)$, tacitly based on the assumption that avalanche indeed is the only significant mechanism in the regime it dominates. As we shown in this chapter, this is not always true as other significant physical effects are present in the weak avalanche regime, and these must be taken into account in present day industrial bipolar devices if one is to achieve good model fits to the measured data and physical model parameters. Though most of the relevant and significant physical mechanisms have been addressed in literature, they cannot be used individually to address the earlier observed deviations between the measured and simulated family of characteristics in the weak avalanche regime. As we demonstrated in this chapter, they must all (collectively) be taken into account if one is to achieve accurate compact model simulation results as a function of both bias and temperature.

References

[1] R. van der Toorn, J. C. J. Paasschens, W. J. Kloosterman, and H. C. de Graaff, *Compact Modeling: Principles, techniques, and applications,* (Springer-Verlag, 2010) Chap. 7.

[2] W. J. Kloosterman and H. C. de Graaff, *Avalanche multiplication in a compact bipolar transistor model for circuit simulation,* in *Proc. BCTM* (1988) pp. 103–106.

[3] W. J. Kloosterman and H. C. de Graaff, *Avalanche multiplication in a compact bipolar transistor model for circuit simulation,* IEEE Trans. Electron Devices **36**, 1376 (1989).

[4] W. J. Kloosterman, J. C. J. Paasschens, and R. J. Havens, *A comprehensive bipolar avalanche multiplication compact model for circuit simulation,* in *Proc. BCTM* (2000) pp. 172–175.

[5] R. Setekera, R. van der Toorn, and W. Kloosterman, *Local extraction of base and thermal resistance of bipolar transistors,* in *Proc. Bipolar Circuits and Technology Meeting* (IEEE, 2013) pp. 21–24.

[6] J. Alvin *et al.*, *Neutral base recombination and its influence on the temperature dependence of Early voltage and current gain-early voltage product in uhv/cvd SiGe heterojunction bipolar transistors,* IEEE Trans. Electron Devices **44** (1997).

[7] S. M. Sze, *Physics of Semiconductor Devices*, 2nd ed. (John Wiley and Sons, INC, 1981).

[8] J. C. J. Paasschens, W. J. Kloosterman, and R. van der Toorn, *Model Derivation of Mextram 504, The physics behind the model,* Tech. Rep. Technical note TN-2004/01010 (Philips Research Eindhoven, 2005).

[9] G. B. Hong and J. G. Fossum, *Implementation of nonlocal model for implact-ionization current in bipolar circuit simulation and application to SiGe HBT design optimization,* IEEE Trans. Electron Devices **42**, 1166 (1995).

[10] M. M. Mahim and J. S. Yuan, *Modelling of the avalanche current including the non-local effect,* Int Journal Electronics **85**, 400 (1998).

[11] *IC-CAP User's Guide*, Agilent Technologies (2008).

[12] P. E. Gill and W. Murray, *Algorithms for the solution of the nonlinear least-squares problem,* SIAM J. Numerical Analysis **15**, 977 (1978).

[13] P. E. Gill, W. Murray, and M. H. Wright, *Practical optimization,* (London: Academic Press, 1981) Chap. 4.

[14] D. M. Bates and D. G. Watts, *Nonlinear RegSression and its Applications* (New York: Wiley, 1988).

[15] J. C. J. Paasschens, S. Harmsma, and R. van der Toorn, *Dependence of thermal resistance on ambient and actual temperature,* in *Proc. Bipolar/BiCMOS Circuits and Technology Meeting* (2004) pp. 96–99.

[16] T. Vanhoucke, H. M. J. Boots, and W. D. V. Noort, *Revised method for extraction of the thermal resistance applied to bulk and SOI SiGe HBTs,* IEEE Electron. Device. Lett Electron Devices **25**, 150 (2004).

[17] R. Setekera, L. F. Tiemeijer, W. J. Kloosterman, and R. van der Toorn, *Analysis of the local extraction method of base and thermal resistance of bipolar transistors,* in *Proc. Bipolar Circuits and Technology Meeting* (IEEE, 2014) pp. 215–218.

[18] T. H. Ning and D. D. Tang, *Method for determining the emitter and base resistance of bipolar transistors,* IEEE Trans. Electron Devices **31**, 409 (1984).

[19] C. T. Kirk, *A theory of transistor cut-off frequency (f_t) fall-off at high current densities,* IEEE Trans. Electron Devices .

[20] P. Ashburn, *SiGe Heterojunction Bipolar Transistors* (John Wiley and Sons Ltd, 2003).

[21] I. E. Getreu, *Modeling the Bipolar Transistor:* CAD of Electronic Circuits, vol.1 (Elsevier Scientific Publishing Company, 1978).

[22] M. Shur, *Introduction to Electronic Devices*, edited by S. M. Elliot (John Wiley and Sons, Inc, 1996).

[23] P. Ashburn, *Design and Realization of Bipolar Transistors*, edited by D. V. Morgan and H. R. Grubin (John Wiley and Sons, Inc, 1988).

[24] W. D. van Noort and R. Dekker, *Thermal resistance of (H)BTs on bulk, SOI and glass,* in *Proc. IEEE* (2003) pp. 129–132.

[25] P. A. H. Hart, ed., *Bipolar and Bipolar-MOS Integration* (Elsevier, 1994).

[26] N. Rinaldi and V. d'Alessandro, *Theory of electrothermal behavior of bipolar transistors: Part III- Impact-ionization,* IEEE Trans. Electron Devices **53**, 1683 (2006).

[27] L. L. Spina, V. d'Alessandro, S. Russo, N. Rinaldi, and L. K. Nanver, *Influence of concurrent electrothermal and avalanche effects on the safe operating area of multifinger bipolar transistors,* IEEE Trans. Electron Devices **56**, 483 (2009).

[28] G. Verzellesi, R. Turetta, P. Pavan, A. Collini, A. Chantre, A. Marty, and C. Canali, *Extraction of dc base parasitic resistance of bipolar transistors based on impact-ionization-induced base current reversal,* IEEE Electron Device Letters **14**, 431 (1993).

[29] R. Setekera, L. F. Tiemeijer, and R. van der Toorn, *Verification of the simultaneous local extraction method of base and thermal resistance of bipolar transistors,* WASET International Journal of Electrical, Computer, Energetic, Electronic and Communication Engineering **08**, 1456 (2014).

[30] J.-S. Rieh, D. Greenberg, B. Jagannathan, and S. Subbanna, *Measurement and modeling of thermal resistance of high speed SiGe heterojunction bipolar transistors,* in *Proc. Topical Meeting on Silicon Monolithic Integrated Circuits in RF Systems* (2001) pp. 110–113.

[31] M. Schröter and A. Chakravorty, *Compact Hierarchical Bipolar Transistor Modeling with HICUM* (World Scientific Publishing Co. Pte. Ltd, 2010).

[32] C. McAndrew *et al.*, *VBIC95 the vertical bipolar inter-company model,* IEEE J. Solid-State Circuits **31**, 476 (1996).

[33] H. K. N. Gummel, *A charge control relation for bipolar transistors,* Bell Sys. Techn.J **31**, 115 (1970).

[34] H. C. de Graaff and F. M. Klaassen, *Compact transistor modelling for circuit design* (Springer-Verlag, 1990).

[35] J. C. J. Paasschens, W. J. Kloosterman, and R. J. Havens, *Parameter extraction for the Bipolar Transistor Model Mextram Level 504,* Unclassified report NL- UR 2001/801 (Philips Research Eindhoven, 2001).

[36] M. Reisch, *High-Frequency Bipolar Transistors* (Springer-Verlag, 2003).

[37] R. Setekera and R. van der Toorn, *Compact model for non-local avalanche effect in advanced bipolar transistors: an assessment of the relaxation length and its temperature dependence,* Solid-State Electronics (2015), submitted.

[38] J. A. Yuan, *SiGe, GaAs, and InP Heterojunction Bipolar Transistors,* edited by K. Chang (John Wiley and Sons, Inc, 1999).

[39] N. Pfenning, *Elementary statistics: Looking at the big picture,* (Cengage Learning, 2010) Chap. 111.

[40] M. L. Martin, *Business statistics: for MBA,* (JMC PRESS, INC, 1999) Chap. 55.

[41] J. W. Slotboom, G. Streutker, M. J. v. Dort, P. H. Woerlee, A. Pruijmboom, and D. J. Gravesteijn, *Non-local impact ionization in silicon devices,* in *Proc. IEDM* (1991) pp. 127–130.

[42] M. R. van den Berg, L. K. Nanver, and J. W. Slotboom, *Temperature dependence of avalanche multiplication in spiked electric fields,* in *Proc. IEDM* (2000) pp. 79–82.

[43] C. Jacoboni, C. Canali, G. Ottaviani, and A. A. Quaranta, *A review of the some charge transport properties of silicon,* IEEE Trans. Electron Devices **20**, 77 (1977).

[44] S. M. Sze and K. N. Kwok, *Physics of Semiconductor Devices,* 3rd ed. (Wiley, 2006).

[45] C. Canali, G. Majni, R. Minder, and G. Ottaviani, *Electron and hole drift velocity measurements in silicon and their empirical relation to electric field and temperature,* IEEE Trans. Electron Devices **22**, 1045 (1975).

[46] P. Deixler, A. Rodriguez, W. De Boer, H. Sun, R. Colclaser, D. Bower, N. Bell, A. Yao, R. Brock, Y. Bouttement, G. A. M. Hurkx, L. Tiemeijer, J. Paasschens, H. Huizing, D. Hartskeerl, P. Agrarwal, P. H. C. Magnee, E. Aksen, and J. Slotboom, *QUBiC4X: An* f_T/f_{max} *= 130/140 GHz SiGe: C-BiCMOS manufacturing technology with elite passives for emerging microwave applications,* in *Bipolar/BiCMOS Circuits and Technology, 2004. Proceedings of the 2004 Meeting* (2004) pp. 233–236.

[47] R. J. E. Hueting, *Charge Carrier Transport in Silicon-Germanium Heterojunction Bipolar Transistors,* Phd. thesis, Delft University of Technology (1997).

[48] M. Rudolph, *Introduction to Modeling HBTs* (Artech House, Inc, 2006).

[49] P. Roblin and H. Rohdin, High-Speed heterostructure Devices, *From device concepts to circuit modelling* (Cambridge University Press, 2002).

[50] P. Palestri, L. Selmi, G. A. M. Hurkx, J. W. Slotboom, and E. Sangiorgi, *Energy dependent electron and hole impact ionization in Si bipolar transisitors,* in *Proc. IEDM* (1998) pp. 885–888.

[51] Z. A. Shafi, C. J. Gibbings, P. Ashburn, I. R. C. Post, C. G. Tuppen, and D. J. Godfrey, *The importance of neutral base recombination in compromising the gain of Si/SiGe heterojunction bipolar transistors,* IEEE Trans. Electron Devices **38**, 1973 (1991).

[52] G. A. M. Hurkx, *Bipolar and bipolar-MOS integration,* (Elsevier, Amsterdam, 1994) Chap. 3.

[53] J. S. Yuan and J. J. Liou, *Semiconductor Device Physics and Simulation* (Plenum Press, New York, 1998).

[54] H. C. de Graaf, W. J. Kloosterman, and T. N. Jansen, *Compact bipolar transistor model for CACD with accurate description of collector behavior purposes,* in *Proc. 18th Conf. Solid St. Dev. and Material, Tokyo, Japan.*

4

[55] M. F. Chang, ed., *Current Trends in Heterojunction Bipolar Transistors* (World Scientific Publishing Co. Pte.Ltd, 1996).

[56] H. C. de Graaf, *Electrical Behavior of Lightly Doped Collectors in Bipolar Transistors,* Phd. thesis, Eindhoven University of Technology (1975).

[57] J. C. Paasschens, W. J. Kloosterman, R. J. Havens, and H. C. de Graaf, *Improved compact modeling of output conductance and cut-off frequency of bipolar transistor,* in *Proc. BCTM* (2000) pp. 62–65.

4

5

Conclusions and Recommendations

In this thesis, we carried out a study and analysis of the significant and relevant physical mechanisms in present day Si/SiGe bipolar transistors that cause the deviations between the measured and simulated family of characteristics in the weak collector-base breakdown regime. We developed a physics-based extended version of the Mextram model that takes non-local avalanche effects and their temperature dependencies into account; which is aimed at resolving the model fitting issues in weak breakdown regime, over bias and temperature. The relevant physical effects virtually depend on temperature, and device self-heating takes the central role here; this is addressed using an improved method developed in this thesis that accurately estimates the device thermal resistance from the same measured characteristics in the weak collector-base breakdown regime. Other relevant physical mechanisms that interfere in the weak breakdown regime, were adequately addressed through application of the physical formulation and corresponding temperature scaling rules of the Mextram compact model. We demonstrated that the weak breakdown characteristics are not just the classical local/non-local avalanche characteristics, but there is more to it through these interfering physical effects. With these effects adequately addressed, we showed that indeed non-local avalanche is responsible for the deviations between the measured and simulated characteristics over the weak avalanche regime; when carefully taken into account, good fits are achieved for different present day industrial Si and SiGe bipolar transistors. Also the extended model formulation preserves the original carefully chosen presupposed underlying semiconductor device physics. Conclusions from each thesis chapter will be presented in this chapter. Specific recommendations for potential future work will be summarized as well in the second section of this chapter.

5.1 Conclusions

Extraction method of the base and thermal resistance of bipolar transistors

The resistance R_B and thermal resistance R_{TH}, play a crucial role both in characterization and for compact modeling of present day bipolar transistors. Therefore, extraction methods are needed to accurately determine their values for a wide range of bipolar technologies. We presented in Chapter 2 a method to extract both the base resistance (R_B) and the thermal resistance (R_{TH}) of Si and SiGe bipolar transistors, from the weak collector-base breakdown characteristics which are used also in the other chapters of the thesis. Our approach extends the category of methods [1, 2] that utilizes the fixed emitter current conditions to distinguish between emitter and base resistance; which are known [3, 4] to pose a fundamental difficulty in DC-methods that explore the normal forward bias regime. The key idea is to measure the differentials in the external base-emitter voltage while forcing a constant emitter current, which ensures that the voltage drop across the emitter resistance is constant, hence zero for constant R_E.

Since methods in this category [1, 2] employees the weak collector-base breakdown characteristics to vary the base current independently of the emitter current, then the differentials in internal base-emitter voltage with respect to the collector-base voltage, would strongly be affected by Early effect and self-heating. These two effects were consistently accounted for by our extraction method so as to arrive to accurate extraction of both base and thermal resistance. The method includes an experimental identification of its regime of applicability, i.e., an experimental check on the validity of the presupposed underlying assumptions; with three suggested approaches that can be employed. Since violation of the employed assumptions has observable consequences, the input data for the extracted method must be selected from this regime of applicability.

The extraction method was demonstrated on DC-measured data taken on a modern RF SiGe-HBT, and the corresponding R_B and R_{TH} values were extracted. From these results, the presupposed underlying assumptions were verified. We compared extracted results for R_B and R_{TH} from our method, with corresponding results from other independent DC- and AC-methods. Among these methods, is the DC-method proposed by Vanhoucke and Hurkx [2], which estimates both R_B and R_{TH} by considering the change of the internal base-emitter voltage (V_{BEi}) due to self-heating, while it ignores the influence of collector-base Early effect on V_{BEi}. For a fixed constant emitter current, this method yields the same extracted R_B value as the method presented in this work (since they both take self-heating effect into account), while the method in [2] yields the extracted value of R_{TH} that is at least 10% larger than that from our extraction method. This observed difference in R_{TH} value can directly be related to the ignored influence of Early effect (in the method proposed in [2]), which leads to an underestimation of the collector current and thus the device dissipated power [5–8], this results into an overestimation of the device thermal resistance. These results were also confirmed by the self-consistency check process, where Mextram model [9] simulation results were used as input data for our extraction method. Results from Mextram's operating point information (OP-info), which takes into account self-heating effect on the model parameters,

are treated as the known exact values of R_B and R_{TH}, thus compared with the extracted results from our method and those from the method proposed in [2]. The deviations between the value of R_{TH} from the OP-info and that from the method we presented in this work is ≈ 0.9%, while that with the method proposed in [2] is ≈ 16.0%. This shows that our proposed extraction method yield more accurate extracted results of R_{TH} in comparison to the method proposed in [2]; which is the key result from the improvements provided by the method in this work. A deviation of less than 2.0% was observed between the OP-info value of R_B and that extracted from our proposed method (which is the same as that for the method proposed in [2]).

The extraction method for R_B and R_{TH} presented in this work is local in bias- and temperature space: measured data involved for a single parameter extraction includes only small variations in emitter-base bias conditions and temperature. With respect to collector-base biasing, only a traverse of the *weak* avalanche regime is involved. This enables parameter extraction as a function of bias and temperature. The method is also local in device space: it is directly applicable to individual bipolar transistors; no additional (DC-) test structures are needed. This makes the method cost efficient e.g., with respect to parameter extraction as a function of geometry. We note that, extracted results for the thermal resistance R_{TH} from this method were used in Chapter 4 of this thesis to address self-heating effects on the measured electrical characteristics that significantly interfere in the weak avalanche regime of present day Si/SiGe bipolar transistors.

Compact model for non-local avalanche effect in bipolar transistors

Following a literature survey to find out the actual relevant and significant physical mechanisms in present days Si BJT's and SiGe HBT's that cause the deviations between the measured family of breakdown characteristics and their simulated counterparts, we developed a physics based compact model for non-local weak avalanche in bipolar transistors. The physical background of this model is the approximate energy balance equation [10] and Chynoweth's empirical law [11] for impact-ionization. The resulting expression for the impact-ionization as a function of electron temperature, turns out to be a sharply peaked function about the maximum positions, for any realistic regime of parameters as valid for conditions in the collector region of bipolar transistors. To enable integration of such sharply peaked functions, we developed a method which is based on the Taylor series expansion of the integral in terms of the width of the peak. We employed this method to derive an approximate expression for the integral of the impact-ionization rate (as a function of the electron temperature) over the width of the collector-epilayer, i.e., the approximate expression for the multiplication factor $(M_n - 1)$; whose multiplication with the collector current, yielded the non-local avalanche current I_{avl}. The resulting formulation is explicit in nature and in terms of elementary functions, so it can be readily implemented in existing complete compact bipolar transistor models, such as Mextram [12], whenever the model provides an explicit relation for the electric field. We took into account the temperature dependence of the new introduced physical parameter, i.e., the relaxation length λ_e, and this provide possible means to compare the model parameters to the earlier published literature values. The introduced new parameter is

the temperature coefficient A_{λ_e}. Thus, only two new non-local avalanche model parameters were introduced, and these correspond to independent material coefficients, so that their values can be compared to values independently published in semiconductor device literature.

For purposes of implementation in a complete compact transistor model, like Mextram [9] and to enable practical applications, the approximate expression for the multiplication factor $(M_n - 1)$ was modified further, in order to take into account cases where it becomes subject to evaluation for parameter and bias values outside its range of physical validity, for example during automated parameter optimization procedures. The resulting expression is globally robust, in a way that it preserves the accuracy of the original formulation for the low and intermediate collector currents, but suppresses the avalanche current in the high current limit. For the case of Mextram 504 [12] bipolar transistor model, the existing formulation for the local electric field is used to define the non-local model input field value at the maximum position. The multiplication factor $(M_n - 1)$ corresponds to the Mextram model variable for the generation factor G_{EM} [12], while other quantities correspond to existing Mextram model parameters or model constants [13]; in fact, as we have demonstrated, their values can be directly obtained from independent published literature. Therefore, this new physics-based non-local avalanche compact model can easily be implemented in the standard Mextram compact model (yielding an extended version of Mextram model), to enable accurate simulation of non-local avalanche effects and their temperature dependencies in present day industrial bipolar transistors. This extended version of Mextram, can possibly be used to resolve the misfits (demonstrated in the introduction chapter of the thesis) between the measured and simulated family of characteristics in the weak collector-base breakdown regime. Note that this extended version of Mextram is kept as closely as possible to the carefully chosen semiconductor device physics, such that the values for new introduced model parameters can easily be related to the available device physics.

Assessment of the physical mechanisms in the weak collector-base breakdown regime

Using the derived physics based extended version of the Mextram compact model, in Chapter 4 we carried out an experimental assessment of the relevant and significant physical mechanisms in the weak collector-base breakdown regime of present day Si BJTs and SiGe HBTs that cause the deviations between the measured and simulated characteristics. For this, we adopted a model verification ambition of not only verifying the fitting capability regarding the targeted characteristics, but also demonstrating that our model actually captures the presupposed underlaying semiconductor device physics. With this latter goal being achieved by showing that the extracted values for the two new model parameters (the relaxation length λ_e and its temperature coefficient A_{λ_e}) correspond to values that have been previously published from independent sources; which generally requires estimation of the accuracy of the extracted values for these parameters.

Employing the extended version of the Mextram model together with the measured

data taken on present day industrial bipolar devices, we demonstrated that the breakdown characteristics in the weak avalanche regime are actually not just the classical local/non-local avalanche characteristics, but other physical mechanisms are also significant in this regime, and affect the terminal characteristics. These interfering physical mechanisms *must* be adequately addressed in order achieve good model fits to the measured characteristics and physical values of λ_e and A_{λ_e}. The presence of these effects limits the traditional approach of plotting the breakdown characteristics in terms of the multiplication factor $(M_n - 1)$, which is generally based on prejudice of only avalanche being the significant mechanism in this regime it dominates. The fact that these physical effects are coupled and virtually depend on temperature, (a) makes parameter extraction so challenging and (b) device self-heating becomes important due to scaling and device isolation technologies in modern RF-devices. Though self-heating is a simple concept, an experimental determination of the thermal resistance (R_{TH}) is a generally recognized challenge. By using the measured characteristics in the same weak breakdown regime together with the extraction method presented in Chapter 2, a good estimate for R_{TH} was determined. With this value for R_{TH}, the effect of self-heating on the measured electrical characteristics that significantly interfere in the weak avalanche regime was addressed. Among other significant effects that were taken into account is the temperature dependence of the collector current through the Early effect. This was addressed independently of the avalanche effects by determining the temperature coefficient for the zero bias neutral base charge in the non-avalanche regime. Generally these physical effects were distinguishable, and thus accounted for independently of the weak avalanche effects.

With these interfering physical effects/mechanisms adequately taken into account, we demonstrated that non-local avalanche is indeed responsible for the deviations between the measured and simulated family of characteristics in the weak collector-base breakdown regime, for present day bipolar devices. This was achieved by showing that the extended version of the Mextram model (with our new non-local avalanche compact model) can accurately reproduce Si-BJT and SiGe-HBT weak collector-base characteristics as a function of both bias and temperature, for all employed industrial bipolar transistors (corresponding to different transistor technologies). The physical content of our extended model (i.e., the model captures the presupposed underlying physics) was verified using measured data taken on a pure Si-BJT, as this choice avoids SiGe-specific interfering effects. This required an estimation of the accuracy of the extracted values for the new model parameters λ_e and A_{λ_e}. Our model produced results for values of λ_e and A_{λ_e}, that correspond to within approximately 10% to independently published literature values for these parameters; which confirms both the adequacy and validity of the physical basis of the model formulation.

Therefore, in the weak collector-base breakdown regime, a number of coupled physical mechanisms other than avalanche are present and they play a key role in the deviations between the measured and simulated family of characteristics in this regime. Avalanche alone as portrayed in literature is not sufficient to address such deviations. The family of these coupled physical effects must all be adequately addressed in order to enable accurate simulation of the measured characteristics; this also enables the extended model to preserve its carefully chosen underlying device physics. In practical

applications of our model formulation for non-local avalanche, no parameter extraction should be needed to find the added non-local avalanche parameters; known material coefficients from independently published semiconductor device literature should suffice. Moreover, once our new non-local avalanche model would be available in any complete transistor model (for our case the extended version of Mextram), the fact that one can rely on its physical basis and known parameter values may assist in extraction or verification of model parameters representing other effects (e.g., the temperature dependence of the Early effect), that interfere in the weak collector-base breakdown regime.

5.2 Recommendations for future work

Generalization of the R_B and R_{TH} extraction method

In chapter 2, we presented an extraction method for both the base resistance R_B and thermal resistance R_{TH} of bipolar junction transistor from the DC-measurements. One of the key assumptions applied in the derivation of this extraction method is that the device Early voltage V_A is much greater than the applied base-collector voltage (V_{CB}). This assumptions limits the application of this extraction method to bipolar devices with *very low Early voltages*. Such devices in this category may include *ultra* high-speed SiGe-HBT devices, with relatively high epilayer dopping in relation to the base dopping concentration [5, 14]. Due to the high doping across the base-collector junction, the Early voltage of such devices will tremendously reduce [7, 15, 16]. Our proposed extraction method, need to be modified in order to enable its applications on such devices with low Early voltages. Attempts to achieve this was taken [17], where we relaxed the assumption of $V_A \gg V_{CB}$. In this we reformulated relation (2.20) as

$$Y = [R_B + \alpha_T R_{TH}]\,(V_{CB} + V_A^{\mathrm{eff}})\,, \tag{5.1}$$

where

$$Y = -\left[\frac{dV_{BE}}{dI_C} + \frac{V_T}{I_C}\left(\frac{V_A + V_{CB}}{V_A^{\mathrm{eff}}} - 1\right)\right]. \tag{5.2}$$

Here we see that Y contains the Early voltage (V_A) and the thermal voltage (V_T). Using (5.1), we define an improved extraction procedure for R_B and R_{TH} in a way similar to that described at the end of Section 2.2. First, we determine the extraction region from the DC-measurements taken at a sequence of constant emitter currents (I_E) as described in Section 2.3, and the input data is taken from this region of validity. In this extraction region, the intercept on the vertical axis of the plot of Y as a function of ($V_{CB} + V_A^{\mathrm{eff}}$) gives the extracted R_B and the corresponding slope ($S_{\mathrm{TOT_{new}}}$) is equivalent to $\alpha_T R_{TH}$; where $\alpha_T = -dV_{BE}/dT$ is determined as described in Section 2.3.
In order to determine Y (assuming small local temperature variations), the Early voltage V_A need to be determined first. This means that another extraction method for V_A is needed. There are various methods published in literature that can be used to extract the bias dependent V_A such as those published in [1, 18]. Note that, in this generalized

extraction method, the accuracy of the extracted V_A for a given bipolar transistor will affect that of the extracted R_B and R_{TH}.

To study the limitations of this method, we need to use measurement data taken on real bipolar devices with very low Early voltage. Similar experimental verification steps as those employed in Chapter 2 can still be employed here as well, but requires the measured data taken on the real low V_A device. These devices were not accessible during the time of this extraction method development, thus we were unable to carry out a demonstration of this generalized method on real experimental data. This can be done in the future once the necessary measurement data becomes available. We believe that this will provide a key step in analyzing the application limits for this proposed extraction method, as well as a room for further improvements of this method.

Verification of the non-local avalanche model in high current regime

A physics base compact model for non-local avalanche effects in present day Si/SiGe bipolar transistors was derived in Chapter 3 of this thesis. As this compact expression is globally robust, it can also be applied on non-local avalanche model simulations corresponding to flat-field current regimes. It was demonstrated extensively in [19, 20] that through the small-signal admittance parameters (Y-parameters), the effect of avalanche on the various characteristics in the high current regime can be studied. Using the real part of the Y_{12} admittance parameter, i.e., Re(Y_{12}) the effect of avalanche on different characteristics such as the unit power gain G_U can be analyzed. From the small signal circuit, Re$(Y_{12}) = -g_\pi$, the conductance in the base-collector junction due to avalanche. Note that the imaginary part of Y_{12}, gives the collector-base depletion capacitance. Since Re$(Y_{12}) = -g_\pi$, it suggests that we in practice -Re(Y_{12}) might be an observable suitable to quantify and study avalanche in the small signal AC regime of transistor operation. Using a small signal equivalent circuit of a full compact model for bipolar transistors e.g., Mextram model [12, 21], in combination with the two-port network circuit representation in common-emitter configuration, we obtain

$$-Re(Y_{12}) = g_\pi + O(\omega^2), \tag{5.3}$$

where $\omega = 2\pi f$ is the angular frequency. Here $O(\omega^2)$ denotes the additional terms of at least second order in ω [21]; which represents effects from parasitic resistances and capacitances (in combination with other model quantities like transconductance). Note that the effects of parasitics on Re(Y_{12}) are of second order as a function of frequency, as opposed to, e.g., the first order contribution of the collector-base capacitances to the Im(Y_{12}) [19]. On basis of this observation, it can be expected that, if the measurements are to be taken at sufficiently low frequencies and sufficiently high output bias voltages, the real part of Y_{12} is going to be dominated by the avalanche conductance g_π [19]. But at sufficiently high frequencies, the $O(\omega^2)$ parasitic terms will dominate $-$Re(Y_{12}). In [20] an experimental verification of avalanche effect in the AC-regime was demonstrated using the Mextram's standard local avalanche model [9, 22, 23]. The qualitative behavior of the real part of Y_{12} admittance parameter as a function of frequency and bias conditions, as described earlier, was clearly observed in the measured data that was taken

on real industrial SiGe HBTs. As discussed previously we observed the impact of weak avalanche on Re(Y_{12}) under conditions of collector current at about the peak cut-off frequency f_T. In this regime, the interpretation of data is not much hampered by other high current effects (e.g., base push out and current spreading [24]); avalanche can be observed in Re(Y_{12}), analogous to current modulation effects in the intrinsic collector-base capacitance, which can be observed in Im(Y_{12}). From the simulation results, the effect of non-local avalanche on the various AC-characteristics was clearly observable, and our non-local avalanche model showed fitting capabilities to the relevant AC-characteristics. But since the major goal is to study the temperature dependence of avalanche current; this non-local avalanche model need to be verified on the relevant AC-characteristics measured at a range of temperatures. Since the original physics based non-local avalanche model i.e., (3.26) was comprised in order to achieve a counterpart that is relevant for complete standard compact models i.e., (3.30), the major goal would now be to verify the fitting capabilities of the non-local avalanche model against the measured AC-characteristics. Similar to the verifications in Chapter 4, AC-measurements taken in different industrial bipolar devices can be used; this will help to study the robustness of the model in terms of device technologies.

5

Using the known non-local avalanche model parameter values to extraction or verify model parameters for other physical effects

In Chapter 4, we verified that the non-local avalanche compact model presented in this work, sufficiently captures the presupposed underlying semiconductor device physics. This was achieved by demonstrating that the extracted values for the non-local avalanche model parameters λ_e and A_{λ_e} are in good agreement with the independently published literature values, i.e., $\lambda_e = 65$ nm [10] and $A_{\lambda_e} = -0.90$ [25–27]. Thus no parameter extraction for these non-local avalanche model parameters is needed, these known material coefficient values should suffice. This paves the way for employing these material coefficients together with the physical basis of the model in the extraction or verification of model parameters representing other physical effects that interfere in the weak avalanche regime. In this, once the new non-local avalanche is incorporated in a complete compact transistor model, then these two non-local avalanche model parameters can be fixed to the known material coefficients from the published literature, and then other model parameters representing the interfering physical effects, as discussed in Section 4.3 can be extracted. By incorporating the non-local avalanche compact model into a full compact bipolar transistor model, and then fixing λ_e and A_{λ_e} to the known material coefficients, it should be possible achieve more physical model parameters for such physical interfering effects, and thus enable verification of model on the presupposed underlying physics.

Extension of the non-local avalanche compact model

During our derivation of the non-local avalanche compact model, we followed the most

employed trend of compact models, i.e., the avalanche current being limited to weak avalanche regime. This weak avalanche formulation assumes that the generated carriers due to impact-ionization do not generate extra charge carriers, i.e., no avalanching. In practical applications, this assumption is not always valid and in some cases, the proposed weak avalanche based compact models fail to capture the targeted avalanche characteristics. In some cases, such characteristics can be captured at the expense of the physical nature of the relevant model parameters; which generally affects the overall compact model fitting capabilities. But due to the fact that the non-local avalanche model we presented in this work captures the presupposed underlying physics, it provides a foundation for further device-physics-based development of model in directions of strong avalanche regime. This can possibly be done by including the secondary ionization in the model for avalanche current, that is neglected in the weak avalanche model.

References

[1] G. Verzellesi, R. Turetta, P. Pavan, A. Collini, A. Chantre, A. Marty, and C. Canali, *Extraction of dc base parasitic resistance of bipolar transistors based on impact-ionization-induced base current reversal,* IEEE Electron Device Letters **14**, 431 (1993).

[2] T. Vanhoucke and G. A. M. Hurkx, *Simultaneous extraction of the base and thermal resistances of bipolar transistors,* IEEE Transactions on Electron Devices **52**, 1887 (2005).

[3] J. Steigerwald and P. Humphries, *TCAD assisted reflection on parameter extraction for compact modeling,* in *Proc. Bipolar Circuits and Technology Meeting* (2010) pp. 245–252.

[4] S. Maas, *Why I hate base resistance,* in *Microw. Mag* (2004) pp. 54–60.

[5] P. A. H. Hart, ed., *Bipolar and Bipolar-MOS Integration* (Elsevier, 1994).

[6] H. C. de Graaff and F. M. Klaassen, *Compact transistor modelling for circuit design* (Springer-Verlag, 1990).

[7] P. Ashburn, *SiGe Heterojunction Bipolar Transistors* (John Wiley and Sons Ltd, 2003).

[8] S. M. Sze and K. N. Kwok, *Physics of Semiconductor Devices*, 3rd ed. (Wiley, 2006).

[9] R. van der Toorn, J. C. J. Paasschens, W. J. Kloosterman, and H. C. de Graaff, *Compact Modeling: Principles, techniques, and applications,* (Springer-Verlag, 2010) Chap. 7.

[10] J. W. Slotboom, G. Streutker, M. J. v. Dort, P. H. Woerlee, A. Pruijmboom, and D. J. Gravesteijn, *Non-local impact ionization in silicon devices,* in *Proc. IEDM* (1991) pp. 127–130.

[11] A. G. Chynoweth, *Ionization rates for electrons and holes in silicon,* Phys. Rev **109**, 1537 (1958).

[12] J. C. J. Paasschens, W. J. Kloosterman, and R. van der Toorn, *Model Derivation of Mextram 504, The physics behind the model*, Tech. Rep. Technical note TN-2004/01010 (Philips Research Eindhoven, 2005).

[13] M. Reisch, *High-Frequency Bipolar Transistors* (Springer-Verlag, 2003).

[14] P. Roblin and H. Rohdin, High-Speed heterostructure Devices*, From device concepts to circuit modelling* (Cambridge University Press, 2002).

[15] S. M. Sze, *Physics of Semiconductor Devices*, 2nd ed. (John Wiley and Sons, INC, 1981).

[16] M. Rudolph, *Introduction to Modeling HBTs* (Artech House, Inc, 2006).

[17] R. Setekera, L. F. Tiemeijer, W. J. Kloosterman, and R. van der Toorn, *Analysis of the local extraction method of base and thermal resistance of bipolar transistors,* in *Proc. Bipolar Circuits and Technology Meeting* (IEEE, 2014) pp. 215–218.

[18] C. C. McAndrew and L. W. Nagel, *Early effect modeling in SPICE,* IEEE Journal of Solid-State Circuits **31**, 136 (1996).

[19] V. Milovanović and R. van der Toorn, *Rf small signal avalanche characterization and repercussions on bipolar transostor circuit design,* in *Proc. IEEE EUROCON* (2009) pp. 230–233.

[20] V. Milovanović, *Advanced Breakdown Modeling for Solid-State Circuit Design*, Phd. thesis, Delft University of Technology (2010).

[21] R. van der Toorn, J. J. Dohmen, and O. Hubert, *Distributon of the collector resistance of planar bipolar transistor: Impact on small signal characteristics and compact modeling,* in *Proc. BCTM* (2007) pp. 184–187.

[22] W. J. Kloosterman and H. C. de Graaff, *Avalanche multiplication in a compact bipolar transistor model for circuit simulation,* in *Proc. BCTM* (1988) pp. 103–106.

[23] W. J. Kloosterman, J. C. J. Paasschens, and R. J. Havens, *A comprehensive bipolar avalanche multiplication compact model for circuit simulation,* in *Proc. BCTM* (2000) pp. 172–175.

[24] J. C. Paasschens, W. J. Kloosterman, R. J. Havens, and H. C. de Graaf, *Improved compact modeling of output conductance and cut-off frequency of bipolar transistor,* in *Proc. BCTM* (2000) pp. 62–65.

[25] C. Canali, G. Majni, R. Minder, and G. Ottaviani, *Electron and hole drift velocity measurements in silicon and their empirical relation to electric field and temperature,* IEEE Trans. Electron Devices **22**, 1045 (1975).

[26] C. Jacoboni, C. Canali, G. Ottaviani, and A. A. Quaranta, *A review of the some charge transport properties of silicon,* IEEE Trans. Electron Devices **20**, 77 (1977).

[27] M. R. van den Berg, L. K. Nanver, and J. W. Slotboom, *Temperature dependence of avalanche multiplication in spiked electric fields,* in *Proc. IEDM* (2000) pp. 79–82.

A

Functions used for the non-local avalanche compact model

A.1 Integration of a Gaussian function

An illustrative demonstration of the practical application of expression (3.24) and of the key role of the coefficient γ_f would be the application of (3.24) to compute the integral of the Gaussian function, defined by

$$f_{A,\beta}(x) = A\exp\left(-\frac{x^2}{2\sigma^2}\right), \tag{A.1}$$

where A and σ (standard deviation) are parameters. A natural choice for the central point s in this case is the position of the maximum of this function, which is at $x = 0$. The corresponding values for $f_{A,\beta}(x)$ and its spatial derivatives at $s = 0$ are:

$$f_{A,\beta}(0) = A, \qquad f^{'}_{A,\beta}(0) = 0, \qquad f^{''}_{A,\beta}(0) = -A, \qquad f^{'''}_{A,\beta}(0) = 0, \qquad f^{''''}_{A,\beta}(0) = \frac{3A}{\sigma^2}\,. \tag{A.2}$$

Substitution of (A.2) into (3.24) and simplifying further, leads to an approximate expression

$$\int_{-\infty}^{\infty} f_\beta(x)\,dx \approx 2\sqrt{\frac{2}{3}}A\gamma_f\sigma\,. \tag{A.3}$$

The right hand side (RHS) of (A.3) in this case takes the form of the exact expression for the integral of the Gaussian function (A.1), for all values of parameters A and σ, when we choose $\gamma_f = \sqrt{3\pi}/2$. This value for γ_f is indeed a constant real value that does not depend on the parameters A and σ.

A

A.2 Smooth bounded functions

For the case of standard compact model formulations, smoothly bounded functions are usually employed in order to find a smooth transition between the two asymptotes of a given function. This is vital to enable convergence of the employed model (mostly near the point of discontinuity), during the compact model simulation process. Such functions have got the following properties

$$f(x) = \begin{cases} 0, & \text{for } x < 0, \\ x, & \text{for } x > 0, \end{cases} \tag{A.4}$$

and for compact modeling, a smooth transition between the two asymptotes is required. The commonly used functions are hyperbolic functions of the form [1]

$$f(x) = \frac{1}{2}\left(x + \sqrt{x^2 + 4\varepsilon^2}\right), \tag{A.5}$$

which is plotted in Fig. A.1; indeed this function describes a hyperbole. As can be seen

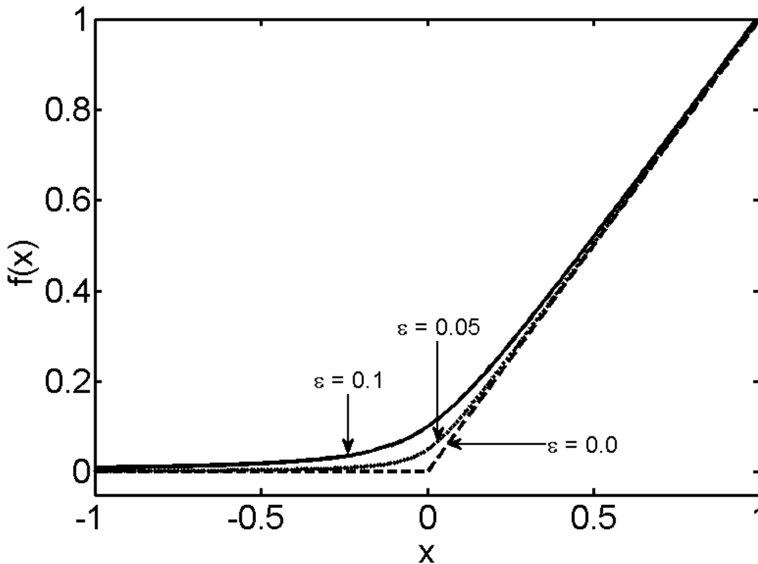


Figure A.1: The hyperbole function $f(x)$ as defined by relation (A.5) and the corresponding asymptotes for different values of ε.

from Fig. A.1, around $x = 0$, the function (A.5) does not converge very fast to the asymptotes. In this, other functions some of which are presented in [1], can be employed. For the work in this thesis (i.e., in Chapter 3 and Chapter 4), we used the smooth bounded functions of the form

$$f(x) = \begin{cases} x_0 + \varepsilon \ln\left[1 + \exp\left(\frac{x - x_0}{\varepsilon}\right)\right], & \text{for } x < x_0, \\ x + \varepsilon \ln\left[1 + \exp\left(\frac{-(x - x_0)}{\varepsilon}\right)\right], & \text{for } x > x_0, \end{cases} \tag{A.6}$$

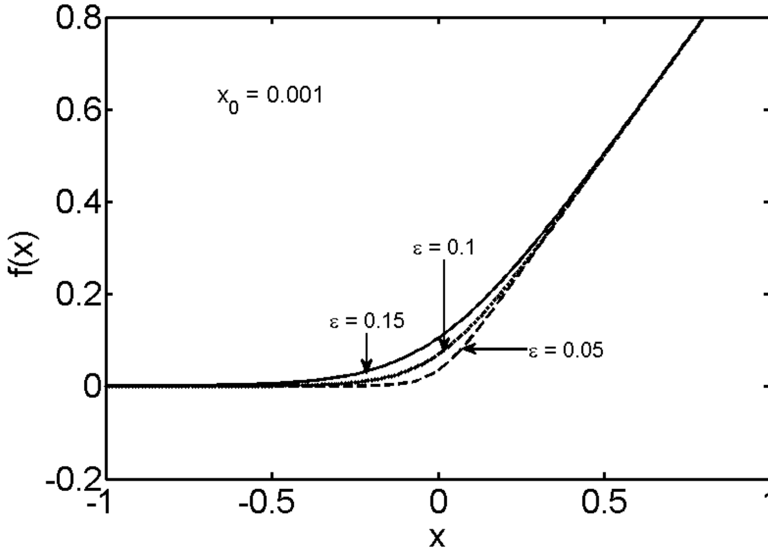


Figure A.2: The hyperbole function $f(x)$ as defined by relation (A.6) and the corresponding asymptotes for different values of ε.

which is plotted in Fig. A.2, for arbitrary values of ε and $x_0 = 0.001$. We can see that this function indeed describes a hyperbole and provides a smooth transition between the asymptotes. Such functions have been implemented in the standard Mextram compact model [1, 2].

For the work in Section 3.3, $x_0 = -0.5B_n(T)$, which is the lower limiting value from the square root function of equation (3.30). The constant $\varepsilon \approx V_{DC}/W_{avl}$, where V_{DC} is the collector-base diffusion voltage (a Mextram model parameter). The variable $x = E_{avl}$, where E_{avl} is defined by equation (3.31). Then from (A.6), the function $f(x) = \tilde{E}_{avl}$, the smoothly bounded function used in equation (3.30).

References

[1] J. C. J. Paasschens, W. J. Kloosterman, and R. van der Toorn, *Model Derivation of Mextram 504, The physics behind the model,* Tech. Rep. Technical note TN-2004/01010 (Philips Research Eindhoven, 2005).

[2] R. van der Toorn, J. C. J. Paasschens, W. J. Kloosterman, and H. C. de Graaff, *Compact Modeling: Principles, techniques, and applications,* (Springer-Verlag, 2010) Chap. 7.

Summary

With the current increasing demand for faster and more reliable communication and computing electronic devices such as faster wireless communication networks, circuit designers are forced to carry out device optimization in order to achieve the maximum possible performance. To enable full circuit optimization, designers depend on compact models for the different circuit elements; this means that the compact models for the individual elements need to be accurate and reliable in order to achieve the best possible performance under different bias and temperature conditions. For present day RF bipolar transistors, there is a tradeoff between device speed and breakdown voltage, i.e., faster devices have got a lower breakdown voltage; this imposes a great difficulty in devices such as power amplifiers, which require both high currents and voltages. For this, circuit designers are forced to explore device operation beyond certain breakdown voltages. Thus, the need for reliable compact models that can be used to accurately simulate the transistor characteristics in the breakdown regime under different operation conditions.

This thesis addresses the physical mechanisms that are relevant and significant to breakdown characteristics of present day industrial Si BJT's and SiGe HBT's. As is demonstrated in the thesis, the standard Mextram model version 504.10, for example, actually is not capable of accurately simulating characteristics in the weak collector-base breakdown regime as a function of both bias and temperature, even though all appropriate temperature scaling rules, well established as they have been in the past for previous generations of semiconductor technology, appear to have been applied. This thesis analyses and resolves this inadequacy of Mextram. With respect to the physics involved, the key questions are: what are the mechanisms in this regime responsible for these misfits and how can they be addressed in order to achieve accurate model simulations? Addressing these questions formed the goal of this thesis. Between the introduction to the thesis provided in Chapter 1, and the overview of the conclusions and potential future prospects presented in Chapter 5, the thesis progresses as follows.

In Chapter 2 an extraction method for base resistance (R_B) and thermal resistance (R_{TH}) of Si/SiGe bipolar transistors is presented. This method uses the measurements of the transistor characteristics in the *weak collector-base breakdown regime*, to accurately determine R_B and R_{TH}, by consistently accounting for the influence of self-heating and Early effect on the internal base-emitter voltage. Generally this method extends the category of the methods that utilizes the weak collector-base avalanche current to vary the base and collector currents independently of the emitter current, while eliminating voltage fluctuations across the emitter resistance by enforcing a fixed emitter current. The method was demonstrated on measured data taken on present day RF SiGe HBT's, and the corresponding results were compared with those from other earlier published extraction methods. By using the simulated data of the standard Mextram compact model as the input data of the method (instead of the measured data), we carried out a self-con-

sistency check of the method; from the corresponding results, our extraction method yielded more accurate values for R_{TH} than the earlier proposed methods. The R_{TH} from this method is used later in Chapter 4 to take into account the influence of self-heating on the electrical characteristics in this weak collector-base breakdown regime.

In Chapter 3, an extended review of the relevant and significant physical mechanisms that are responsible for the deviations between the measured family of weak breakdown characteristics and their simulated counterparts, as a function of both bias and temperature is carried out. Firstly, an extensive review of the physical mechanisms addressed by the Mextram compact model is undertaken. Secondly, a development of a physics-based compact model formulation of non-local avalanche effects in bipolar transistors and their temperature dependencies is carried out. The physical basis of this non-local avalanche model is the approximate energy balance equation and Chynoweth's empirical law for impact-ionization. The ionization coefficient (α) as a function of the electron temperature, turned out to be a sharped peaked function about the maximum position. The approximate expression for the multiplication factor ($M_n - 1$) is attained by Taylor series expansion of the integral of α over epilayer width, in terms of the relative width of the function peak. With only two new model parameters, the final expression for the weak avalanche current turned out to be explicit in nature and in terms of elementary function, so it could easily be implemented in existing full compact bipolar transistor models. A compact formulation suitable for industrial applications of this non-local avalanche model was also derived and it accurately approximates the original formulation for low and intermediate collector currents, but it suppress the avalanche current for high collector currents. Using the existing Mextram model electric field distribution, the Mextram model was extended with the new non-local avalanche model; the resulting extended version of Mextram is physical in nature with only two new introduced model parameters, i.e., the relaxation length and its temperature coefficient. This extended model version is used to assess the relevant and significant physical mechanisms in the weak collector-base breakdown regime.

Chapter 4 is devoted to an experimental assessment of the family of weak collector-base breakdown characteristics by employing the physical extended version of the Mextram model (with included the non-local avalanche compact model) derived in Chapter 3. Using the measurement data taken on present day Si/SiGe industrial bipolar transistors, we demonstrate that these observed breakdown characteristics are actually not just classical local/non-local avalanche characteristics as portrayed in most published literature, but other physical mechanisms are significant as well in this breakdown regime. We showed that these physical effects can actually be distinguished, and thus taken into account independently of avalanche effects. Here, we took advantage of the well developed physical basis of the Mextram compact model in addressing these effects, together with our extraction method for R_{TH} developed in Chapter 2 to address self-heating effects on the measured electrical characteristics that significantly interfere in the weak collector-base breakdown regime. With these interfering physical effects well addressed, we demonstrated that the avalanche mechanism responsible for the deviations between the measured and simulated breakdown characteristics is non-local avalanche indeed, and with non-local effects adequately taken into account, accurate model fits to the measured characteristics taken on present day industrial Si/SiGe bipolar transis-

tors are achieved as a function of both bias and temperature. Also with these physical effects adequately taken into account, the developed model (Mextram extended with our new non-local avalanche compact model) actually captures the presupposed underlying semiconductor device physics. Here, our extracted values for the electron energy relaxation length and its temperature coefficient, in actual industrial bipolar devices under normal forward bias conditions, are in agreement with earlier independent published values for these material coefficients. Though the relevant and significant physical mechanisms in the weak collector-base breakdown regime, each separately, have already been published in semiconductor device literature, and while most of them have been adopted in standard compact models for bipolar transistors, these effects cannot be used individually to address the deviations between the measured and simulated family of characteristics in this regime. All of them must be jointly taken into account in order to achieve accurate simulation of the characteristics in this breakdown regime; this is well demonstrated and verified in Chapter 4.

Since the derived extended version of Mextram model is physical in nature, and the new model parameters are indeed material coefficients, with parameter values already available in published semiconductor device literature, the new compact model can be used to extract or verify model parameters representing other physical effects that interfere in the weak collector-base breakdown regime. Also the physical basis of the model can form a foundation for extension of the model to strong avalanche breakdown regime.

Samenvatting

Door de huidige, toenemende vraag naar snellere en meer betrouwbare elektronische apparaten voor communicatie en informatieverwerking, zoals snellere draadloze communicatie-netwerken, worden circuit ontwerpers gedwongen het uiterste uit elektronische componenten te halen. Voor volledige optimalisatie van circuits zijn ontwerpers afhankelijk van *compacte modellen* voor de verschillende circuit elementen. Opdat ontwerpers de beste prestaties kunnen realiseren onder uiteenlopende omstandigheden wat betreft temperaturen en elektrische randvoorwaarden dienen compacte modellen voor de individuele circuit elementen nauwkeurig en betrouwbaar te zijn. De huidige hoogfrequent (RF) bipolaire transistoren weerspiegelen typisch een compromis tussen snelheid en doorslagspanning; dat wil zeggen, snellere transistoren hebben een lagere doorslagspanning. Dit impliceert aanzienlijke moeilijkheden met betrekking tot bijvoorbeeld vermogensversterkers, in welke immers zowel grote stroomsterkten als hoge spanningen optreden. Door deze context zijn circuit ontwerpers gedwongen om componenten te laten werken bij spanningen die groter zijn dan sommige doorslagspanningen. Hieruit komt de behoefte voort aan betrouwbare compacte modellen die kunnen worden ingezet om de karakteristieken van transistoren te simuleren in het doorslagregime, en zulks onder uiteenlopende omstandigheden.

Deze dissertatie bestudeert de fysische mechanismen die relevant en significant zijn met betrekking tot de doorslagkarakteristieken van huidige industriële silicium (Si) bipolaire transistoren (BJT's) en silicium-germanium (SiGe) heterojunctie bipolaire transistoren (HBT's). Zoals wordt gedemonstreerd in de dissertatie is bijvoorbeeld het standaard Mextram model, versie 504.10, niet in staat om nauwkeurig karakteristieken te simuleren, als functie van elektrische randvoorwaarden en temperatuur, in het zwakke collector-basis doorslagregime. Dit ondanks het feit dat alle van toepassing zijnde temperatuurschalingsregels in Mextram lijken te worden toegepast, terwijl die regels in het verleden toch grondig zijn vastgesteld voor toenmalige generaties van halfgeleider technologie. Deze dissertatie analyseert deze tekortkoming van Mextram en presenteert er een oplossing voor. Met betrekking tot de betrokken fysica zijn de sleutelvragen dan: welke *zijn* de mechanismen die in dit regime verantwoordelijk zijn voor de waargenomen afwijkingen en hoe kunnen ze in acht worden genomen, opdat nauwkeurige model simulaties worden gerealiseerd? Tussen de inleiding verschaft in hoofdstuk 1, en het overzicht van conclusies en mogelijke toekomstige ontwikkelingen gepresenteerd in hoofdstuk 5, is de opbouw van deze dissertatie als volgt.

In hoofdstuk 2 wordt een extractie methode gepresenteerd voor de basisweerstand R_B en de thermische weerstand (R_{TH}) van Si/SiGe bipolaire transistoren. Deze methode benut gemeten karakteristieken van een transistor in het zwakke collector-basis doorslagregime om nauwkeurig R_B en R_{TH} te bepalen, op basis van een consistent in rekening brengen van de invloed van zelf-verwarming en van het Early effect. De methode breidt de categorie uit van methoden die de zwakke collector-basis doorslagstroom be-

nutten om de basis- en collector stromen onafhankelijk van de emitterstroom te variëren. Tegelijkertijd worden spanningsfluctuaties over de emitter weerstand voorkomen door een constante emitterstroom op te leggen. De methode wordt gedemonstreerd op basis van data die gemeten is aan huidige hoogfrequent SiGe HBT's en de resultaten worden vergeleken met corresponderende resultaten verkregen met eerder gepubliceerde extractiemethoden. Door gesimuleerde data van het standaard Mextram model te benutten als invoerdata voor de methode (in plaats van gemeten data) hebben we zelf-consistentie tests uitgevoerd; onze extractie methode bleek nauwkeurigere waarden voor R_{TH} op te leveren dan de eerder voorgestelde methoden. De waarde voor R_{TH} als verkregen met de methode wordt in hoofdstuk 4 benut om de invloed van zelfverwarming in rekening te brengen met betrekking tot de elektrische karakteristieken in het zwakke doorslagregime.

In hoofdstuk 3 wordt een uitgebreid overzicht gegeven van de relevante en significante fysische mechanismen die van invloed zijn op de gemeten familie van zwakke doorslag karakteristieken en hun gesimuleerde tegenhangers, als functie van zowel elektrische randvoorwaarden als temperatuur. Om te beginnen wordt een uitgebreid overzicht opgesteld van de fysische mechanismen die door het standaard Mextram model in rekening worden gebracht. Vervolgens wordt een fysisch gebaseerd compact model ontwikkeld voor niet-lokale lawine effecten in bipolaire transistoren, inclusief temperatuur afhankelijkheden daarvan. De fysische basis van dit niet-lokale lawine model is de benaderende energiebalansvergelijking en Chynoweth's empirische wet voor inslag ionisatie. De ionisatie coëfficiënt (α), als functie van de elektronentemperatuur, blijkt een functie te zijn die scherp is gepiekt rond de positie van zijn maximum. Een benaderende expressie voor de vermenigvuldigingsfactor ($M_n - 1$) wordt verkregen middels een Taylor reeks ontwikkeling, in termen van de relatieve breedte van de piek van de functie, van de integraal van α over de dikte van de epilaag. Het resultaat is een expressie voor de lawine stroom, die slechts twee nieuwe modelparameters introduceert. De expressie is expliciet in termen van elementaire functies en daardoor zeer geschikt voor implementatie in ieder bestaand compleet compact transistormodel. Een compact model formulering, geschikt voor industriële toepassing van dit niet-lokale lawine model, werd ook afgeleid. Dit model benadert de oorspronkelijke formulering nauwkeurig voor lage en middelhoge stroomsterkten, maar het onderdrukt de lawine stroom bij hoge stroomsterkten. Onder gebruikmaking van Mextram's bestaande model voor het relevante elektrische veld, werd het Mextram model uitgebreid met het nieuwe niet-lokale lawine model. De resulterende uitgebreide versie van Mextram is fysisch van aard en heeft slechts twee nieuwe model parameters, te weten de relaxatie lengte en de temperatuur coëfficiënt daarvan. Deze uitgebreide model versie is benut om relevante en significante fysische mechanismen in het zwakke doorslagregime nader te onderzoeken.

Hoofdstuk 4 is gewijd aan een experimenteel onderzoek van de familie van zwakke collector-basis doorslagkarakteristieken, dat gebruik maakt van de fysische, uitgebreide versie van het Mextram model zoals ontwikkeld in hoofdstuk 3 (dus inclusief het compacte model voor niet-lokale lawinevorming). Gebruikmakend van gemeten data, vergaard met betrekking tot moderne, industriële Si/SiGe bipolaire transistoren, hebben we gedemonstreerd dat waargenomen zwakke doorslag karakteristieken feitelijk geen zuivere, klassieke lokale- dan wel niet-lokale doorslagkarakteristieken zijn, zoals die worden

geportretteerd in het overgrote deel van de beschikbare literatuur; integendeel, we laten zien dat andere fysische effecten evenzeer significant zijn in dit doorslagregime. We hebben gedemonstreerd dat deze fysische effecten feitelijk kunnen worden onderscheiden, en dus in rekening kunnen worden gebracht, onafhankelijk van lawine effecten. Bij het bestuderen van deze effecten hebben we profijt getrokken van de goed ontwikkelde fysische basis van het Mextram model, alsook van onze extractie methode voor R_{TH}, zoals ontwikkeld in hoofdstuk 2; dit laatste om zelf-verwarming goed in rekening te kunnen brengen met betrekking tot interpretatie van de waargenomen elektrische karakteristieken in het zwakke doorslagregime. Onder adequate rekenschap van alle zich manifesterende fysische effecten hebben we gedemonstreerd dat het mechanisme, verantwoordelijk voor de resterende afwijkingen tussen gemeten en gesimuleerde doorslag karakteristieken, inderdaad niet-lokale lawinevorming is. Na adequate rekenschap van niet-lokale effecten, kunnen gemeten karakteristieken van moderne Si/SiGe bipolaire transistoren nauwkeurig worden gereproduceerd, als functie van zowel elektrische randvoorwaarden als temperatuur. Voorts blijkt, wanneer deze fysische effecten zorgvuldig in rekening worden gebracht, het ontwikkelde model (Mextram, uitgebreid met ons nieuwe compacte model voor niet-lokale lawinevorming) daadwerkelijk getrouw de vooronderstelde, onderliggende halfgeleidercomponentfysica te vangen. Onze geëxtraheerde waarden voor de elektron energie relaxatie lengte, en diens temperatuur coëfficiënt, met betrekking tot hedendaagse industriële bipolaire transistoren, onder normale voorwaartse bedrijfscondities, zijn in overeenstemming met eerdere, onafhankelijk gepubliceerde waarden voor deze materiaal coëfficiënten. Hoewel de relevante en significante fysische mechanismen in het zwakke doorslagregime elk afzonderlijk eerder zijn gepubliceerd in de literatuur der halfgeleidercomponentfysica, en hoewel de meeste van deze zelfs zijn opgenomen in standaard compacte modellen, is een verklaring van de afwijkingen tussen waargenomen en gesimuleerde karakteristieken in dit regime niet goed mogelijk op grond van enig van deze effecten afzonderlijk. Al de betrokken effecten moeten in hun onderlinge samenspel in rekening worden gebracht om tot een getrouwe simulatie van de karakteristieken in dit doorslagregime te komen. Dit wordt duidelijk gedemonstreerd en geverifieerd in hoofdstuk 4.

Omdat de afgeleide, aangevulde versie van Mextram fysisch van aard is, en omdat de nieuwe modelparameters inderdaad materiaalcoëfficienten zijn, kan het nieuwe model worden ingezet om modelparameters voor andere fysische effecten, en wel die die zich ook manifesteren in het zwakke collector-basis doorslagregime, te extraheren of te verifiëren. Voorts verschaft de fysische basis van het model een fundering voor verdere uitbreiding van het model, bijvoorbeeld in de richting van het sterke lawine doorslagregime.

List of Publications

- **R. Setekera** and R. van der Toorn, “Compact Model for Non-local Avalanche Effect in Advanced Bipolar Transistors: an Assessment of the Relaxation Length and its Temperature Dependence", *Solid-State Electronics*, submitted for publication (September 2015).
- **R. Setekera**, L. Tiemeijer and R. van der Toorn, “Verification of the Simultaneous Local Extraction Method of Base and Thermal Resistance of Bipolar Transistors", *World Academy of Science, Engineering and Technology: International Journal of Electrical, Computer, Energetic, Electronic and Communication Engineering*, vol. 08, no. 10, pp. 1456-1460, September. 2014.
- **R. Setekera**, L. Tiemeijer, W. Kloosterman and R. van der Toorn, “Analysis of the Local Extraction of the Base and Thermal Resistance of Bipolar Transistors", *In Proceedings of IEEE Bipolar/BiCMOS Circuits and Technology Meeting, BCTM 2014*, September 28 - October 01 2014, San Diego, California, USA, pp. 215-218.
- **R. Setekera**, R. van der Toorn and W. Kloosterman, “Local Extraction of the Base and Thermal Resistance of Bipolar Transistors", *In Proceedings of IEEE Bipolar/BiCMOS Circuits and Technology Meeting, BCTM 2013*, September 30 - October 03 2013, Bordeaux, France, pp. 21-24.

Acknowledgements

Doing PhD is like walking a long journey on a hilly narrow bounded path, so one can choose to get tired a long the way and then turn back or can decided to press on for the price ahead; but if one chooses to press further, he must stay focused in order to avoid the various destructions a long the way that may lead him to slip off that narrow path. I would like to thank God the creator of the universe who guided me through this challenging journey, He provided enough knowledge, wisdom, strength and grace I needed each day, and He never left me alone. Not forgetting the wonderful gift of life, without it all will be vanity. During this journey, many people supported me in different aspects, I will do my best to appreciate them here, but in order not to forget any individual, I will try to avoid mentioning names, and instead, I will give a general acknowledgement.

First and far most, I would like express my utmost gratitude to my supervisor, Dr. Ramses van der Toorn for a number of reasons. For giving me a chance to work within his group on such an interesting but challenging research topic. I am really so grateful that you took a bold stand to offer this physics related position to a mathematician, this cannot easily be found, and it was so priceless to me; in fact it is the one which kept me going even when challenges can along the way, my goal was to reward your trust by all means. Though I had the necessary mathematics, but starting this physics related PhD was challenging, it required building sufficient knowledge on a number of subjects such as compact modeling, Mextram model, ICCAP, and taking measurements in advanced labs, I am so grateful for the materials you provided and the long discussions we had, these helped me to build the required knowledge in the shortest time possible. At times we had to spend some long hours in front of a computer showing me how ICCAP works and how Mextram simulations and parameter extraction are done, not forgetting the measurement process, this was a key to understanding what my research was all about-thank you very much. You also introduced me to various experts in the field (like those at NXP), which helped me to build more knowledge through various discussions with them, as well as exposure to the application of my research work. During my research period, I found out that it was very difficult to come up with something new in this already saturated research field, with the fact that many companies have got their own local research groups working on related topics, made it more challenging and at times discouraging, but you were always available for open discussions, always rising some valuable questions, which helped me to get a new picture of problem at hand; eventually we came up with some interesting results that have been appreciated by the industrial community. Last but not least, I would like you for the very long discussions we had about different life and philosophical topics, these helped me to step out of the PhD world and learn about at other interesting things.

Special thanks to my PhD promotor Prof. Arnold Heemink, for accepting me in his Mathematical Physics group and for his meekness. In all the interactions we had, he made me feel more than welcome, as well as comfortable during the various meetings.

Not forgetting, he is a good listener, who gives you room to express your point during discussions without interruptions.

I would also like to acknowledge NXP Semiconductors for the financial support of this work, and for providing us with some of the bipolar devices (wafers) used for the work in this thesis. In addition, I would like to extend my gratitude to special NXP members, i.e., Willy Kloosterman, Dr. ir. Andries Scholten, and Dr. L. F. Tiemeijer for the various open discussions we had, their ideas helped me to develop the work presented in this thesis. Great thanks to Ir. Ralf Pijper, for availing the advanced electronic lab at NXP and for teaching me how to carry out such advanced measurements. Not forgetting Atef Akhnoukh and Peter Swart, for their wonderful help with DC and RF measurements. In the same regard, I would like to extend my appreciation to DIMES TU Delft, for availing me with the (DC-) measurement lab, thank you very much.

During my PhD journey, I came across a number of individuals at TU Delft, who contributed in various ways to this journey. I cannot mention each one of them, as the list can go on and on, but I would like to thanks fellow PhD students and staff-members of ECTM, DIAM, and GS for their great support, greetings, smiles, and encouragement; thank you very much. Special thanks to Prof. L. K. Nanver, first for accepting me in her research group as a new PhD student and for continuous support/advise even when we left her research group. I wish to thank my fellow PhD student A. Rajab, for great jokes we had in our office and staying around; you created such a wonderful working environment.

Heartfelt thanks to Wim Tiwon for the dedicated efforts to ensure that our computers are always working perfectly: he was always willing to have long discussions about the problem(s) at hand and to reply continuous emails in order to solve the problem in the shortest time possible. I extend my great thanks to Dorothee Engering and Evelyn Sharabi the wonderful secretaries of DIAM, and Marian Roozenburg, secretary of ECTM, for their assistances with all administrative work.

Honestly I would not be able to complete this PhD without the continuous support and prayers from my beloved family. Thank you very much, I am so proud of you. My Church family, you mean a world to me, thank you very much for your prayers and encouragements; may the Almighty God blessed you abundantly. I also wish to thank all my friends both from Uganda and around the world, for the advise, jokes, encouragements, discussions, corrections, and support in one way or the other; I am so grateful.

Finishing PhD is one think, and using it to make a difference is another think; my prayer is that: "*God You help me to use this PhD for the glory of Your Name. I know this is a new journey to walk, but I cannot walk it successfully without you Lord, I pray that your hold my hand and lead me, guide me through this path which is unknown to me, but well known to You Father; help me to walk according to Your will and Lord make me fruitful everywhere You will lead me, I know without You I can do nothing, Ameny*".

About the Author

Robert SETEKERA

Robert Setekera was born in Mukono Uganda on June 1, 1983. He attained is BSc (in Mathematics and Physics) in October 2006 from Makerere University Kampala, Uganda. In September 2009, he received his double MSc degrees in Industrial and Applied Mathematics with specialization in Computational Science and Engineering, from Eindhoven University of Technology, Eindhoven, The Netherlands and Technical University Kaiserslautern, Kaiserslautern, Germany.

After his MSc, he joined the Postmaster program (Mathematics for Industry) at the Department of Mathematics, Eindhoven University of Technology, Eindhoven, The Netherlands, where he received a Professional Doctorate in Engineering (PDEng) degree in September 2011. Here, he worked on a number real-life industrial problems and used various mathematics skills to solve such challenging and interesting industrial problems.

In October 2011, he joined the faculty of Electrical Engineering, Mathematics and Computer Science, Delft University of Technology, Delft, The Netherlands, where he worked towards the PhD degree, under Delft Institute of Applied Mathematics (DIAM).

His current research focus includes compact modeling, semiconductor device characterization and parameter extraction with respect SiGe HBTs especially mm-wave devices.