INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.

**ProQuest Information and Learning
300 North Zeeb Road, Ann Arbor, MI 48106-1346 USA
800-521-0600**

UMI[®]

Modeling of Thermal Coupling in Multifinger Bipolar Transistors

by

Dritan Celo

**A thesis submitted to the Faculty of Graduate Studies and Research
in partial fulfillment of the requirements for the degree of**

Master of Applied Science

**Department of Electronics
Carleton University
Ottawa, Ontario, Canada**

August 2001

© Dritan Celo

**National Library
of Canada**

**Acquisitions and
Bibliographic Services**

**395 Wellington Street
Ottawa ON K1A 0N4
Canada**

**Bibliothèque nationale
du Canada**

**Acquisitions et
services bibliographiques**

**395, rue Wellington
Ottawa ON K1A 0N4
Canada**

Your file Votre référence

Our file Notre référence

The author has granted a non-exclusive licence allowing the National Library of Canada to reproduce, loan, distribute or sell copies of this thesis in microform, paper or electronic formats.

L'auteur a accordé une licence non exclusive permettant à la Bibliothèque nationale du Canada de reproduire, prêter, distribuer ou vendre des copies de cette thèse sous la forme de microfiche/film, de reproduction sur papier ou sur format électronique.

The author retains ownership of the copyright in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author's permission.

L'auteur conserve la propriété du droit d'auteur qui protège cette thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

0-612-66887-8

Canada

The undersigned recommend to
the Faculty of Graduate Studies and Research
acceptance of this thesis

“Modeling of Thermal Coupling in Multifinger Bipolar Transistors”

Submitted by Dritan Celo, B. Eng.
in partial fulfilment of the requirements for the degree of

Master of Applied Science

A handwritten signature in black ink, reading "David J. Walkey", written over a horizontal line.

Thesis Supervisor

David J. Walkey, Ph. D., P. Eng., Associate Professor

A handwritten signature in black ink, written over a horizontal line. The signature is stylized and appears to be "P. J. P.". Below the line, the text "Chair, Department of Electronics" is printed.

Chair, Department of Electronics

Carleton University

August, 2001

Abstract

An extensive thermal analysis for high-power multiemitter bipolar transistor structures is performed, and, based on the analysis results, two simplified models are created to predict maximum temperature for devices with/without metal interfinger connection. Predicted temperatures obtained from these models are compared extensively with temperatures from numerical/analytical simulators. Short calculation time (approximately 10-15 times shorter than numerical one) makes simplified models very useful in preliminary stages of thermal design process where accurate numerical models are prohibited.

The model created for HBTs without interfinger metal was based on a normalized approach, which shows that coupling between sources in the normal multi-finger configuration can be handled in a normalized way by expressing the inter-finger spacing as a proportion of finger length. Normalization coefficient $\alpha=0.62$ was used, and the normalized result for a device with emitter length 20 μm is quite close to 10 and 40 μm , therefore, this case can be used as the predictor. The error in temperature prediction is less than 2%.

The model for HBTs with interfinger metal takes into account the effect of heat distributed through the metal. The model is applied to devices with a variety of geometries (finger length 10-100 μm , number of fingers 1-23, and spacing 4-100 μm), and materials (GaAs, InP, Si for the substrate, and Au, Al for metal structure). It was found that metal layer reduces maximum temperature by about 13%, while increasing metal thickness more than 3 μm has no practical effect on temperature reduction. Compared with numerical results, calculation time is reduced by about 15 times, maintaining two percent accuracy.

Acknowledgments

My gratitude is owed, first of all, to my thesis supervisor, David Walkey. Thanks for your encouragement, expertise, and endless patience, qualities which have made this work a truly fulfilling experience.

I would like to acknowledge the support of Nortel Networks for providing funding for this research, and the Faculty of Graduate Studies and Research for financial support through awarding an Ontario Graduate Scholarship in Science and Technology (OGSST). This scholarship was contributed in parts from Cadabra and Ontario Ministry of Education and Training.

I would like to thank Professor Tom Smy for his helpful assistance during my thesis research and provision of the numerical simulator Atar which was extensively used throughout this work.

I would also like to thank the staff and faculty in the Department of Electronics at Carleton University, my fellow research group members Gerardo Romo and Lei Zhang for their continuous help and participation in private discussions, as well as Helen Triantafelou who helped to edit this material.

Of course, this work would not have been possible without the support of my wife, my family, and friends. Thank you all.

Table of Contents

Abstract	iii
Acknowledgments	iv
List of Tables	viii
List of Figures	ix
List of Abbreviations and Symbols	xii
Chapter 1 Introduction	1
1.1 Thermal Design Considerations for Multi-Finger HBT	1
1.2 Thesis Objectives	2
1.3 Thesis Outline	3
Chapter 2 HBT-Overview and Thermal Properties	4
2.1 Heterojunction Bipolar Transistor	4
2.1.1 Basic Structure and Concept	5
2.1.2 HBT Structure and Fabrication Technology	7
2.2 Thermal Effects in a HBT Structure	9
2.2.1 Modes of Heat Transfer in Semiconductor Devices	9
2.2.2 The Heat Conduction Equation-General	10
2.2.3 HBT Thermal Performance	13
Chapter 3 Simulation of Thermal Effects	16
3.1 Discussion of Existing Approaches	16
3.2 T _{source}	18
3.2.1 The Thermal Model	19
3.2.2 Solution of Heat Conduction Equation	20

3.3 Numerical Simulator Atar	24
Chapter 4 Normalized Approach for Modeling of Thermal Coupling	29
4.1 Introduction	29
4.2 Nature of Thermal Coupling	30
4.2.1 Modeling Assumption	31
4.2.2 Temperature Distribution in a Device, Lateral vs. Longitudinal	33
4.2.3 Thermal Coupling Along Long and Short Edges	35
4.2.4 Lateral Temperature Distribution	38
4.3 Normalized Approach of Lateral Coupling	42
4.4 Normalized Approach to Prediction of Maximum T_{rise}	52
4.4.1 Geometry Dependence of Maximum T_{rise}	52
4.4.2 Effect of Normalization Case	54
4.5 Verification Using Numerical Simulation	55
4.6 Design Scenarios where ‘Optimal’ Use of Area is Required	59
4.6.1 Constant Footprints	59
4.6.2 Design Scenarios	61
4.6.3 Finger Geometry Variations at Constant Area	63
4.6.4 Use of the Normalized Approach for Optimization	68
4.7 Width Representation	69
4.8 Summary	70
Chapter 5 The Effect of Interfinger Metal on Temperature Distribution	72
5.1 Heat Removal Techniques in Power HBTs	72
5.2 Temperature Distribution from Numerical Simulations	75
5.2.1 Theoretical Background	75
5.2.2 Structure Description-General	80

5.2.3 Heat Flux and Isothermal Contours	82
5.2.4 The Effect of Materials and Geometry in Maximum T_{rise}	91
5.3 Structure Representation in Circuit Simulation	98
5.3.1 Lumped Model	98
5.3.2 Model Development	104
5.3.3 Input Data	105
5.3.4 Device Model	105
5.3.5 Metal Structure Model	119
5.3.6 Extraction of the Thermal Network	111
5.3.7 Circuit Simulation	115
5.3.8 Model Specifications and Simulation Results	116
5.4 Model Simplification	117
5.5 Maximum Temperature Analysis, Approximated Model vs. Atar	122
5.5.1 Interfinger Material Dependence of Metal Coupling Effect	122
5.5.2 Substrate Material Dependence of Metal Coupling Effect	123
5.5.3 Geometry Dependence of Metal Coupling Effect	124
5.6 Summary	129
Chapter 6 Conclusions.....	130
6.1 Thesis Summary	130
6.2 Future Research	132
References	133

List of Tables

Table 1. The magnitude of error evaluated in different locations	48
Table 2. Maximum T_{rise} comparison for structures with three different finger length ..	54
Table 3. Example of trade off between footprint area and T_{rise}	62
Table 4. Example of trade off between active area and T_{rise}	63
Table 5. Maximum finger temperature predicted with lumped model	101
Table 6. Simulated maximum T_{rise} . Approximated model vs. Atar	117
Table 7. Simulated maximum T_{rise} , comparison between two models	120
Table 8. Circuit size-approximated model vs. Atar	121
Table 9. Maximum temperatures for different metal structure material	122
Table 10. Maximum T_{rise} for three different substrate materials	123

List of Figures

Figure 2.1 Schematic of GaAs mesa HBT structure-cross section.	6
Figure 2.2 A simplified representation of energy band diagram for the HBT	7
Figure 2.3 Cross-section view of a self-aligned HBT.	8
Figure 2.4 Two different HBTs structure	9
Figure 2.5 One dimensional heat conduction through a flat plate	11
Figure 2.6 A unit-volume element in a semiconductor.	12
Figure 2.7 Schematic diagram of 3-finger emitter structure.	15
Figure 3.1 Heat source simulating transistor power dissipation	20
Figure 3.2 Illustration of the method of images-cross section	22
Figure 3.3 Contribution of each source in the device temperature profile	23
Figure 3.4 Heat flows between blocks represented by thermal resistors	27
Figure 3.5 A 3D model of GaAs HBT with metal structure and temperature contours .	28
Figure 4.1 Schematic for thermal simulation structure in semi infinite substrate	31
Figure 4.2 Temperature profile for single finger device	34
Figure 4.3 Temperature profile for single finger device	36
Figure 4.4 Temperature profile for two different directions	38
Figure 4.5 Temperature profiles for three finger device with different L	39
Figure 4.6 Normalized temperature as function of distance in lateral direction.	41
Figure 4.7 Normalized temperature as function of normalized distance	43
Figure 4.8 Magnitude of error for different values of $\alpha=0.58$ to 0.72	45
Figure 4.9 Normalized temperature as function of normalized distance for $\alpha=0.62$	46
Figure 4.10 Normalized temperature as function of normalized distance for $\alpha=0.5$	47
Figure 4.11 Magnitude of prediction error vs. normalized length	48
Figure 4.12 Coupling effect and the error associated with predicted T_{rise}	49
Figure 4.13 Accumulated error for 23 finger structure.	51
Figure 4.14 Normalized center T_{rise} for $L=20\text{ }\mu\text{m}$	53
Figure 4.15 Normalized center T_{rise} utilizing the normalized approach.	55

Figure 4.16 Power HBT model for numerical simulation	57
Figure 4.17 Simulation results-numerical vs. analytical.	58
Figure 4.18 Example for equal area arrays with two and three fingers	60
Figure 4.19 Equal area curves superimposed with normalized T_{rise} curves	61
Figure 4.20 Spacing vs. number of fingers for two different total L	65
Figure 4.21 Array factor vs. number of fingers for constant temperatures.	66
Figure 4.22 Normalized spacing vs. N for three different constant T_{rise}	67
Figure 4.23 Finger width as function of number of fingers	69
Figure 5.1 Simplified schematic of “Flip-Chip” approach	73
Figure 5.2 Three finger emitter structure with metal interconnection	74
Figure 5.3 Isothermal surfaces and heat flow paths in a solid body	77
Figure 5.4 Network for numerical analysis of 2-D heat conduction.	79
Figure 5.5 Power HBT with metal interconnection numerical model	81
Figure 5.6 Heat flux and isothermal contours x-y plane for $z=0\text{ }\mu\text{m}$	83
Figure 5.7 Heat flux and isothermal contours y-z plane for $z=-2\text{ to }2\text{ }\mu\text{m}$	84
Figure 5.8 Temperature profiles along emitter finger for $z=0:1\text{ }\mu\text{m}$	86
Figure 5.9 Heat flux and isothermal contours y-z plane for $z=-0.05:2\text{ }\mu\text{m}$	87
Figure 5.10 Heat flux and isothermal contours y-z plane for $z=0:2\text{ }\mu\text{m}$	88
Figure 5.11 Heat flux and isothermal contours x-y plane for $z=1.1\text{ }\mu\text{m}$	89
Figure 5.12 Isothermal contours x-y plane for $z=1.9\text{ }\mu\text{m}$	90
Figure 5.13 The effect of metal contact on temperature reduction.	92
Figure 5.14 Structure model, device with metal contacts only.	94
Figure 5.15 Normalized temperature vs. contact thickness	95
Figure 5.16 The effect of metal interconnection layer on maximum T_{rise}	96
Figure 5.17 Normalized temperature vs. metal thickness for different spacings	97
Figure 5.18 Lumped model representation.	99
Figure 5.19 Lumped model for metal structure	100
Figure 5.20 Thermal network extracted from lumped model	101
Figure 5.21 Temperature profiles across fingers, structure with and without metal	102

Figure 5.22 Temperature profiles along fingers, structure with and without metal	103
Figure 5.23 Model development flow chart	104
Figure 5.24 Schematic of single emitter structure with metal	106
Figure 5.25 Temperature profile along emitter finger.	107
Figure 5.26 Elementary sections in a heat source.	109
Figure 5.27 Temperature profile approximation for 17 sections/finger	110
Figure 5.28 Simplified model for metal structure.	111
Figure 5.29 Device model and thermal network extracted	112
Figure 5.30 Example of thermal resistance calculation	113
Figure 5.31 Electrical network modeling the whole structure	114
Figure 5.32 Nodal structure of middle finger	118
Figure 5.33 Temperature profiles along emitter for different sublayers in metal	119
Figure 5.34 Result consistency as function of number of sections per finger	121
Figure 5.35 Temperature reduction vs. L for constant and variable spacing	125
Figure 5.36 Temperature difference center-edge vs. L	126
Figure 5.37 Temperature reduction vs. N for device with metal structure	127
Figure 5.38 Temperature reduction vs. S for device with metal structure	128

List of Abbreviations and Symbols

HBT	heterojunction bipolar transistor
MMIC	monolithic microwave integrated circuit
Si BJT	silicon bipolar junction transistor
A/D and D/A	analog to digital, digital to analog conversion
PA	power amplifier
VCO	voltage controlled oscillator
MBE	molecular beam epitaxy
MOCVD	metal-organic chemical vapor deposition
FET	field effect transistor
PHS	plated heat sink
T	temperature
T_{rise}	temperature rise
P	power dissipated in base-collector junction
R_{th}	thermal resistance
F_a	footprint area
N	number of fingers in HBT
L	finger length
W	finger width
S	spacing between fingers
U_s	unit of spacing
U_a	unit of area
k_{th}	thermal conductivity
ρ	material density
c	thermal capacity of material
q	rate of heat flow
α	thermal diffusivity

Chapter 1

Introduction

1.1 Thermal Design Considerations for Multi-Finger HBT

The demands of the electronic industry have dictated that electronic assemblies operate at increased speeds in a smaller package with higher reliability. These requirements create extreme challenges in thermal management for the designer engineer [1]. More power generates more heat, while smaller packages increase the difficulty of removing it from sensitive devices, which are adversely affected by increased temperature.

The term thermal management encompasses the technology of the generation and control of heat in electronic circuits. Heat is an unavoidable by-product of every electronic device and circuit, and is usually detrimental to performance and reliability. Heat may be generated by the devices themselves or may be present from other sources, internal or external. Today's tendency of placing more functions in a smaller package has resulted in higher heat densities, mandating that thermal management be given a high priority in the design cycle, in order to maintain system performance and reliability. Even modest reductions in peak temperatures, especially in critical regions, can increase the lifetime of these devices significantly. However, speed or power handling capabilities could be increased by running at higher current densities with equivalent lifetimes.

Although sophisticated and highly accurate thermal simulators exist, they are often impractical for use in early design stages, where detailed information about the system to be designed is not yet available. Instead, simplified estimation methods have a clear

advantage due to immediate availability of the results, and due to straightforward relations between the structure of the system and its thermal behavior. Based on such consideration, and the knowledge of the important role that HBTs have played in high speed digital and microwave analog circuit application in recent years, an extensive thermal analysis for the HBT is presented, with the intention of creating simplified models which are accurate enough to represent the physical behavior in the device, without being computationally expensive.

1.2 Thesis Objectives

The first objective is to develop a first order model to predict the thermal coupling effect in a multifinger bipolar transistor. This model is based on generalized simulation results for a given structure, which can be used to predict the thermal performance for a ‘family’ of devices, without repeated simulation. Part of the first objective is also to provide some guidance in the design process by showing the trade-off between active area and device operation temperatures.

The second objective is to develop a simplified model for predicting the effect of interfinger metal on maximum T_{rise} in a multifinger bipolar transistor as an alternative to numerical simulator. The model development is aimed at capturing the fundamental nature of heat flow in the structure, and quickly estimating the device temperature for a wide variety of structures with different geometries and technological parameters. Accuracy and short simulation time are the most essential requirements for this new simplified model.

1.3 Thesis Outline

Chapter 2 provides some background information related to thermal analysis performed throughout the work, including an overview of HBT and its thermal properties, and several basic concepts of heat transfer mechanism that take place in a HBT device. Chapter 3 presents the tools for thermal effect simulation: it begins with a brief discussion of existing approaches and continues with a description of two thermal solvers developed at Carleton University previously to this work: an analytical and a numerical solver utilized to simulate the thermal aspects of HBTs. Chapter 4 introduces a new approach for the modeling of thermal coupling effects in multifinger HBT. Several examples are illustrated to present the environment in which this new approach can be used to predict maximum device temperature. Predicted temperature is compared with the result obtained from analytical and/or numerical simulation. Chapter 5 introduces a new first order model for multifinger HBT with an interfinger metal structure. In the initial stage, a lumped model is created and the results are analysed. Non-physical results obtained from the lumped model lead to a more sophisticated model, which can better represent thermal properties in the device. The maximum device temperature simulated with the new model and simulation time required are compared with numerical results. In addition, some devices with a large emitter array size are simulated using the new model. Chapter 6, finally, outlines the conclusions drawn from this thesis, while indicating the contributions of this research to the bipolar device thermal modeling and design, and explores some related future research areas.

Chapter 2

HBT-Overview and Thermal Properties

2.1 Heterojunction Bipolar Transistor

The principle of HBT was first suggested by Shockley in 1951 [2], followed by Kramer's [3] proposed theory and advantages of what was called a wide-gap emitter in 1958. In 1969, Feucht [4] fabricated the first HBT which remained in an experimental stage until after the 1980's. At the present, however, HBT technology has reached a mature stage as many foundries have fabricated promising HBTs and their application-related circuits.

Compared to its predecessor, HBT offers higher frequency of operation and is often known as the high frequency version of Si BJT [5], [6]. Like its Si bipolar counterpart, it is characterized by high power density and linear operation. It is also distinguished for its high speed and monolithic microwave integrated circuit (MMIC) compatibility of GaAs FETs. A combination of such strengths is the main attribute of a high performance microwave device that can be fabricated using relaxed geometries [7]. In addition, HBTs offer higher transconductance, efficiency, and gain. Extensive research and development efforts have been undertaken globally over the last 15 years to develop, design, and fabricate HBTs for diverse applications with higher power densities and frequency of operation. For modern microwave HBTs, a multiple emitter finger structure has frequently been used, in which several HBT emitters, each with its own HBT operation, are placed parallel to each other with proper spacing in between [8]. Such a structure can reduce the HBT sig-

nal propagation and delay time, as well as allow a large emitter area without extreme lengths. Over the years, many HBT digital and analog circuits have been demonstrated [5],[6], including digital signal processing, mixed signal, and microwave power amplification [9], where higher output power and efficiency are in demand. In particular, A/D and D/A converters, phase detectors, sampling and hold circuits, and voltage comparators have been developed for GBits/s digital applications. For microwave analog circuits, high performance PA, wide band amplifiers, low noise oscillators, and VCO's using HBT have been achieved and produced.

2.1.1 Basic Structure and Concept

The HBT is a bipolar transistor with a wide band-gap emitter. To achieve this, the emitter of an HBT is made of material different from that used for the base and collector. The dissimilar band-gap energies at the emitter-base junction form the heterojunction and, hence, the name heterojunction bipolar transistor. A detailed analysis of the operation concept and different approaches to realizing the HBT structure can be found in [10]. This section provides a brief description of the simplest GaAs mesa HBT structure (single hetero structure, emitter-up) consisting of a wide-bandgap AlGaAs emitter (n-type), a heavily doped GaAs base (p-type), a lightly doped GaAs collector (n-type), and a heavily doped subcollector (n-type). These layers are grown on a semi-insulating GaAs substrate by either molecular beam epitaxy (MBE) or metal-organic chemical vapor deposition (MOCVD). The schematic of such a device is shown in Fig. 2.1. The energy band diagram of the HBT under forward-active operation (base-emitter voltage $V_{BE} > 0$ and collector-

base voltage $V_{CB} > 0$) is given in Fig. 2.2. The superior performance of HBT results directly from the valence-band discontinuity ΔE_V at the hetero interface, which stems from the proper choice of the heterojunction pair with emitter bandgap E_{GE} greater than base bandgap E_{GB} . ΔE_V increases the valence-band barrier height in the emitter-base heterojunction and, thus, reduces the injection of holes from the base to the emitter, which is a dominant component of the base current in the silicon homojunction BJT. This allows the base to be heavily doped while maintaining a large current gain. The HBT can then have a very thin base without being concerned about punch-through in the base region. Thus the benefit of having ΔE_V is twofold. First, the high base doping concentration decreases the base series resistance, which, in turn, reduces the ac, dc, and transient emitter crowding [13]; second, the very thin base reduces the base transit time and enhances the cut-off frequency [10].

Figure 2.1 Schematic of GaAs mesa HBT structure-cross section.

The conduction band discontinuity ΔE_C (spike), on the other hand, is not as desirable as ΔE_V because the spike makes it necessary for the free carriers in the heterojunction to transport by means of thermionic and tunnelling mechanisms [10] instead of drift-diffu-

sion transport in homojunctions, thus impeding the emitter junction efficiency and decreasing the collector current. This problem can be alleviated to some extent by inserting a thin layer (graded layer) before the hetero-interface, or by inserting a thin, undoped GaAs layer (setback layer) after the heterointerface[10].

Figure 2.2 A simplified representation of energy band diagram for the HBT under forward active operation.

2.1.2 HBT Structure and Fabrication Technology

Although the idea of the HBT is as old as the BJT itself, an actual HBT was not demonstrated until the 1980's. Such lateness was mainly attributed to the technical difficulties in forming an acceptable heterojunction of III-V semiconductors of unequal band gap [4]. However, with recent advances in the material growth techniques such as MBE and MOCVD, high performance HBTs are made possible. These techniques allow the deposition of one material onto the other, layer by layer, nearly on an atomic scale. As a result, both the composition (hence the band structures and the inherent electric field) as well as the doping levels of the HBT can be precisely controlled. In practice, there are many ways of fabricating an HBT. Depending on transistor performance requirements,

different degrees of complexity on the fabrication process are used. The fabrication of HBT, using GaAs material, results in higher frequency of operation, because of higher electron mobility in GaAs than Si. Here a brief review of HBT fabrication methods is presented while a detailed coverage of this subject can be found in [14]. Three different processing technologies which are non-self aligned, self-aligned, and planar are addressed here. The first two technologies can be grouped under the mesa-etched technology. The structure for AlGaAs/GaAs non-self-aligned HBT is similar with the schematic shown in Fig. 2.1

It is often desirable to place the emitter and base contacts as close to each other as possible to reduce the base series resistance and extrinsic base surface recombination. The self-aligned technique shown in Fig. 2.3, is often used to achieve this goal. In such a technique, the base contact is produced in a way that its separation from the emitter is ensured by a controlled amount because of the emitter geometry. A separation of $0.1\ \mu\text{m}$ to $0.2\ \mu\text{m}$ between emitter and base contacts is achievable. Furthermore, using the self-aligned processing technique [14], higher f_t can be achieved. To date, HBT's with high switching speed and high cut-off frequency in the range of 100~200 GHz have been reported [5].

Figure 2.3 Cross-section view of a self-aligned HBT [14].

In both of these fabrication methods, the entire wafer surface was etched back every time a new layer was accessed in the transistor structure, thereby creating a nonplanar structure. Such a structure results in a mesa device structure and creates the need for dielectric crossovers to connect each device terminal to external circuits. A device structure can be made more planar by the use of ion implantation techniques. Ion implantation is intended to form conductive channels and convert conductive layers into semi-insulating layers. The quasi-planar structure shown in Fig. 2.4a, utilizes ion implantation only for isolation purposes. In the planar structure shown in Fig. 2.4b, device isolation is achieved by implantation as before, but the collector contact is brought to the base layer. Although this structure is more planar than the previous one, the wafer surface is still at the base level layer. A truly planar structure is not commonly applied to HBTs so as to maintain high emitter injection efficiency.

Figure 2.4 Two different HBTs structure [14]
a) Quasi-planar and b) fully planar.

2.2 Thermal Effects in a HBT Structure

2.2.1 Modes of Heat Transfer in Semiconductor Devices

The mechanism by which heat is transferred in a heat exchange or an energy con-

version system is quite complex. There are, however, three basic and distinct modes of heat transfer: conduction, convection, and radiation [15]. Because the heat transfer by conduction is the most important mode that takes place in a semiconductor, it is only this mode that is considered in this thesis. The other two modes are less important, even in the case when convectional heat transfer is used to cool the substrate externally, the heat transfer through the path between power dissipating devices and the semiconductor substrate is still governed by the conduction process. It has been pointed out that the convection from the top of the substrate does not affect the device itself [15]. The same conclusion is shown to be true for radiation mode as well, because the rate of energy transfer through thermal radiation is calculated to be negligible compared to the internal device power dissipation. It should be noted that throughout this work, heat conduction is analyzed in the absence of electrical carriers, so that the thermal-electric effects can be neglected.

2.2.2 The Heat Conduction Equation-General

Before analysing the thermal effects in a HBT device, it will be useful to briefly review the differential equation of heat conduction, which governs the device junction temperature as a function of power dissipation. Because heat transfer takes place every time there is a temperature gradient in a medium, a knowledge of values of temperature at all points of the medium is essential in heat transfer studies. The instantaneous value of temperature at all points of the medium of interest is called the *temperature distribution* or field. The temperature distribution can be unsteady if it varies with time, or steady if at a given point it never varies with time; it is a function of space coordinates only. In the fol-

lowing analysis, steady temperature distributions are considered.

Fourier's law is the basic law governing heat conduction based on the continuum concept. Consider a flat plate of material having a *surface area* A (m^2) and a thickness Δx , as illustrated in Fig. 2.5. One side of the plate is maintained at *temperature* T (K), and the other at $T+\Delta T$. The Fourier's law of heat conduction states that the rate of *heat flow* q (W) is directly proportional to the area A and the temperature difference ΔT , but inversely proportional to the thickness Δx . Mathematically it is expressed as:

$$q = -k_{th} A \frac{\Delta T}{\Delta x} = -k_{th} A \frac{\partial T}{\partial x} \quad (2.1)$$

A negative sign appears because the heat flow, from the hotter area to the colder area, is in - x direction. The constant of proportionality, k_{th} (W/mK) is called the *thermal conductivity*. According to [16], k_{th} is a property of the material. In a temperature range suitable for device operation, k_{th} decreases with increasing temperatures for all major semiconductors. It is expected that the k_{th} also decreases in a region where a high current level flows.

Figure 2.5 One dimensional heat conduction through a flat plate.

Following the introduction of some basic concepts, the temperature distribution of the substrate containing the device, once the power dissipation level is known, can be studied. For a unit-volume element in a semiconductor (Fig. 2.6), considering infinitesimal volume with $\Delta x \rightarrow 0$, $\Delta y \rightarrow 0$, and $\Delta z \rightarrow 0$, the overall rate of excess energy entering the volume is equal to

$$(\dot{q}_{1x} - \dot{q}_{2x}) + (\dot{q}_{1y} - \dot{q}_{2y}) + (\dot{q}_{1z} - \dot{q}_{2z}) = k_{th} \Delta x \Delta y \Delta z \left(\frac{\partial^2 T}{\partial x^2} + \frac{\partial^2 T}{\partial y^2} + \frac{\partial^2 T}{\partial z^2} \right) \quad (2.2)$$

Figure 2.6 A unit-volume element in a semiconductor.

The principle of conservation of energy requires that the amount of heat energy storage in a volume is equal to the heat energy entering the volume through the six surfaces, and the heat generated from source(s) inside the volume, which, in a semiconductor, can be a device that dissipates power. Mathematically, this principle can be expressed as

$$\Delta x \Delta y \Delta z \rho C \frac{dT}{dt} = k_{th} \Delta x \Delta y \Delta z (\nabla^2 T) + g \Delta x \Delta y \Delta z \quad (2.3)$$

where ρ is the *material density* (kg/m^3) and C is the *thermal capacity* of the material (J/kgK). By rewriting Eq. (2.3), the differential equation of heat conduction in general form

is obtained:

$$\nabla^2 T(x, y, z, t) + \frac{g(x, y, z)}{k_{th}} = \frac{1}{\alpha} \frac{\partial T(x, y, z, t)}{\partial t} \quad (2.4)$$

where α is *thermal diffusivity* (length²/time, which may be given the units m²/h or m²/s) and is defined as $\alpha = k_{th}/(\rho C)$. Differential equation (2.4) enables the calculation of the temperature distribution in a semiconductor containing the device, once the boundary conditions are known. This derivation assumes that the material is homogeneous, and that the thermal conductivity is independent of location and temperature.

Another aspect with interest in thermal device analysis is the first order, steady-state relationship between internal *power dissipation* P (W) in the region, and the *rise above ambient temperature* of a device T_{rise} (K), which can be expressed [17] using the linear formulation

$$T_{rise} = R_{th} P \quad (2.5)$$

where R_{th} (K/W) is the *thermal resistance*. The value of the thermal resistance depends on material properties such as the thermal conductivity, and the geometries of the heat generation and dissipation regions [18]. With a constant value for R_{th} , this relationship assumes a temperature independent thermal conductivity.

2.2.3 HBT Thermal Performance

Two thermal aspects are related to the electro-thermal performance of the HBT. The first one is the control of the average value of the junction temperature, which is directly related to the total dissipation power level. The second aspect is the temperature

distribution across the multi-finger power HBT. The first aspect has been intensively investigated by several authors [14], [19]-[21]. The second has only drawn limited attention until recently. Several authors [22], [23] have demonstrated that the current density distribution (or temperature profile) in a multi-finger HBT can become extremely non uniform when the HBT enters certain operation regions. These authors have also demonstrated that the phenomenon called the *collapse of current gain* [23] is caused by extreme non uniform temperature distribution across the fingers.

In a multi-emitter HBT, two mechanisms are involved in the creation of a non uniform temperature profile, known as *self-heating* and *thermal coupling effects*. For single finger devices, the power dissipated across the finger causes the junction temperature of the finger to rise above the ambient temperature. This phenomenon is called the device self-heating, where the junction T_{rise} of a finger is derived exclusively from the power dissipation of the same finger. Defining the power dissipated in the finger as $P = I_B V_{BE} + I_C V_{CE} \equiv I_C V_{CE}$, and using the self-heating thermal resistance R_{th} , the temperature T (K) in the device can be characterized by

$$T = T_A + R_{th} I_C V_{CE} \quad (2.6)$$

Where T_A is ambient temperature (K), I_C is device current (A), and V_{CE} is the bias voltage (V). In devices with more than one finger (Fig. 2.7), the junction temperature of one finger increases due to power dissipation both in itself and in the other fingers. According to [16], the junction temperature of one finger is a simple sum of the T_{rise} due to the power dissipation in this finger and in the others. To illustrate, in a three finger case, temperature

T for each finger can be found by the following equations:

$$\begin{aligned} T_1 &= T_A + R_{th11} V_{CE} I_{C1} + R_{th12} V_{CE} I_{C2} + R_{th13} V_{CE} I_{C3} \\ T_2 &= T_A + R_{th21} V_{CE} I_{C1} + R_{th22} V_{CE} I_{C2} + R_{th23} V_{CE} I_{C3} \\ T_3 &= T_A + R_{th31} V_{CE} I_{C1} + R_{th32} V_{CE} I_{C2} + R_{th33} V_{CE} I_{C3} \end{aligned} \quad (2.7)$$

where $R_{thj,k}$ represents the mutual coupling thermal resistance. I_{Cf} represents current density in the f_{th} finger and could have a different value for each finger due to the temperature variations, and the voltage V_{CE} which, generally, has the same value for all fingers. The amount of power dissipated in each finger differs from one finger to another because of the variation in current density.

Figure 2.7 Schematic diagram of 3-finger emitter structure.

Chapter 3

Simulation of Thermal Effects

3.1 Discussion of Existing Approaches

Regardless of the multifinger HBT's increasing importance and popularity in high-power microwave applications, efforts on theoretical modeling of such a device have been limited in the past because of the complexity of coupling the solution of both the semiconductor device equations and the heat equation [24]. This combined electro-thermal problem is a significant challenge. This is due in part to the complicated nature of negative feedback [10] of the thermal effect on the HBT current-voltage characteristic. The problem is further complicated by the thermal coupling between the neighbouring emitter fingers when a multifinger structure is considered [10]. However, this problem can be simplified and the device's thermal properties can be determined without solving the device equations, if the power generation region is known apriori [25]. By knowledge of the power generation region as a function of position, the heat equation can be solved independently for a calculation of the device temperature profile.

Since high power applications are the area where HBT's have some advantage over the heterojunction FET's, the thermal characterization and modeling of HBT's have been given considerable attention in the literature of the 1990's. Several published papers have dealt with the thermal analysis of multifinger structure HBT's [20],[26], etc. It has been shown that a two-dimensional analysis leads to reduced temperatures more strongly than the simple one-dimensional model where thermal-flow spreading is not included [27].

However, the temperature obtained from two-dimensional calculations are still higher than the experimental value since the transverse or longitudinal heat conduction is neglected. For design guidelines, a three-dimensional model where the heat flow is spread in a cone shape is more desirable. Also, temperature distributions along the depth direction of the substrate are required to analyse the thermal behavior for HBT structures [29]. In addition, for high-power multifinger HBT's, the thermal behavior is no longer correctly characterized if it is based on the assumption of uniform junction temperature distributions over the active region. In this case, a high power HBT can be regarded as an array constituted by $M \times N$ elementary transistors, and the distributions may be treated uniformly in each elementary transistor [20].

In the available literature, there are many analytical, semi-numerical, and numerical methods aimed at solving this 3D thermal problem and finding the temperature distribution in a given structure. Nevertheless, an essential question is raised: which methods would be most effectively applied in a typical multiemitter HBT structure? According to [1], thermal behavior of an HBT can be characterized using experimental and/or computational simulations. The last one is the first chosen form for two reasons: it is faster to develop either an analytical or a numerical model based on finite elements or finite difference methods; and it allows parametric simulations. This means that key parameters influencing design are varied to measure their effects on junction temperature. A closer examination of computational methods [1] shows that the analytical methods are the first for forming the solution. Because of its analytical nature, computational simulation is typically referred as the first stage of the solution. The numerical solution is considered as a

second stage with a limited domain of applications. The experimental method is the highest level of solution and often follows the computational modeling, or used when the other two methods are not suitable for the specific problem addressed.

Based on this idea, computational simulations with both analytical and numerical methods are performed throughout this thesis with an intent to study the device thermal properties. A brief description of the tools used for this work is provided in the sections that follow.

3.2 Tsource

Tsource is a thermal solver which was developed at Carleton University prior to this work. It is based on an analytical model for a three-dimensional transient heat flow problem which takes into account the physical structure of the device and the actual region of power dissipation. At any point within the device, the model predicts the time dependent-temperature response to a power dissipation which is an arbitrary function of time. The thermal coupling between adjacent devices in an integrated circuit is also considered. In the following sections, the model [28] used in tsource is explained, starting with basic assumptions and the geometry involved in the thermal model of a bipolar transistor, followed by the solution of 3D heat equation which models a thermal point source. This point source problem is extended to dissipation due to a heat source with a given geometry and the steady state temperature at any point inside the substrate volume can be found. Finally, some limitations related with the analytical solver tsource are reported.

3.2.1 The Thermal Model

The assumptions made in the model are related to the device geometry and boundary conditions applied. The surface of the chip may be regarded as adiabatic, that is, no heat flows perpendicular to the boundary at the surface. Thus, the model neglects the possibility of conduction through the metal leads or convection from the surface. The approximation of the real medium which is considered is then a semi-infinite, homogeneous, with one adiabatic surface. Considering only one adiabatic surface and infinite medium in other directions in the substrate, the effect of all chip boundaries at the closest device surface is neglected. Homogeneity is a good approximation, because most of the region in which thermal spreading resistance occurs in a real transistor has sufficiently low impurity concentration (i.e. less than 10^{18} cm^{-3}) so that the thermal properties are nearly constant. Also, the diffused regions are usually thin compared to the dimensions of the region in which thermal spreading resistance occurs.

The source of power dissipation in a bipolar transistor is mainly in that part of the collector/base depletion region where most of the voltage drops. In this region, the carriers attain high velocities, and through phonon interactions give up their energy to the lattice. This region of power dissipation will be approximated by a solid rectangular parallelepiped submerged beneath the surface of the semi-infinite medium (Fig. 3.1). The volume in which the power is dissipated is shown as the region of dimensions $L \times W \times H$ located at a depth D below the surface (x - y plane). For a transistor, the distance L and W correspond to the length and width of the emitter stripe, whereas H is approximately the thickness of the depletion region, and D is the depth of the collector/base junction below the surface.

Figure 3.1 Heat source simulating transistor power dissipation in a semi-infinite medium [28].

3.2.2 Solution of Heat Conduction Equation

A starting point for constructing a thermal model is the continuous time dependent Green's function solution of the 3D heat equation [30], which models a thermal point source.

$$T = \frac{P}{8(\pi\alpha)^{3/2}} \int_0^t \exp[-r^2/(4\alpha(t-t'))] \frac{dt'}{(t-t')^{3/2}} \quad (3.1)$$

This equation gives the temperature T at any point (x,y,z) due to a point source located at (x',y',z') present from $t=0$ until a later time t , where r is the distance (m) from the point source location to the point (x,y,z) . Temperature independent thermal conductivity as well as dissipation into an infinite material are assumed. For application to thermal behavior of integrated devices, an expression of the effect of a rectangular source is required to model a base-collector depletion region. Following the method in [28], this point source model can be extended to dissipation due to a rectangular source by integrating the point source

location over the ranges $x_1 \leq x \leq x_2$, $y_1 \leq y \leq y_2$, $z_1 \leq z \leq z_2$ to yield

$$T = \frac{P}{8V\rho c(\pi\alpha)^{3/2}} \int_0^t \left(\left(\operatorname{erf}\left(\frac{x_2-x}{\sqrt{4\alpha(t-t')}}\right) + \operatorname{erf}\left(\frac{x_1-x}{\sqrt{4\alpha(t-t')}}\right) \right) \left(\operatorname{erf}\left(\frac{y_2-y}{\sqrt{4\alpha(t-t')}}\right) + \operatorname{erf}\left(\frac{y_1-y}{\sqrt{4\alpha(t-t')}}\right) \right) \left(\operatorname{erf}\left(\frac{z_2-z}{\sqrt{4\alpha(t-t')}}\right) + \operatorname{erf}\left(\frac{z_1-z}{\sqrt{4\alpha(t-t')}}\right) \right) \right) \frac{dt'}{(t-t')^{3/2}} \quad (3.2)$$

The above solution represents the temperature due to a power dissipation in an infinite medium. To obtain the solution for a semi-infinite medium with an adiabatic and conductive boundary conditions, the method of imaging can be used [15]. In this technique, the zero heat flow across an adiabatic boundary condition is modeled by placing a second “mirror” source in a position of symmetry relative to the position of the adiabatic boundary. This symmetry of position cancels the heat flow along the adiabatic plane, because heat sources are mirror images of each other and generate an equal amount of heat. A conductive boundary condition can be realized with the same technique but for a sink (source of negative strength) in place of the mirror source.

Figure 3.7 shows the method of images applied to the problem of a thermal source (a cross section is shown) within a semi infinite region with one adiabatic boundary condition at the substrate surface and one conductive boundary condition at the bottom of the substrate where the heat sink is located [31]. The source, indicated by a rectangular enclosing a “+”, is placed so that symmetry is preserved across the line defined by adiabatic boundary condition. The sink, indicated by a rectangular enclosing a “-”, is placed to ensure symmetry with sign reversal across the line defined by the conductive boundary

condition. The temperature inside the volume described by the boundary conditions due to the original source, is the sum of the effect of the original source and as many image sources as required to ensure convergence.

Figure 3.2 Illustration of the method of images-cross section.

The steady state temperature at any point inside the substrate volume can, therefore, be determined by first modeling the boundary conditions using a suitable arrangement of image sources and sinks; then by summing the effects of these image sources and sinks by repeated application of Eq. (3.2) with (x, y, z) as the point of interest, $(x_1, x_2, y_1, y_2, z_1, z_2)$ representing the translated location of the image source, and a numerical integration over time to a time where the temperature does no longer change significantly. Thermal resistance for the source is a bulk quantity. Since its value is function of the position, a point must be chosen at which to evaluate the temperature. Having the power dissipated and final value of the temperature increase at this point, corresponding R_{th} can be found using the relation

$$R_{th}(x, y, z) = \frac{T(x, y, z)}{P} \quad (3.3)$$

The solution for several heat sources in the medium, such as the case of a multi-emitter device, is found by using linear superposition: that is, since the differential equation is linear, the contribution from each source is added at all points in time and space. To demonstrate this, an example for a five-finger emitter structure is simulated and the temperature distribution for each finger and for the entire device is shown (Fig. 3.3).

Figure 3.3 Contribution of each source in the device temperature profile.

It is worthwhile to mention that although the model employed in *tsource* is fast and useful for a wide variety of transistor structures, some limitations exist and are related with assumptions made in the model. For example, the model works for planar or quasi planar devices, but mesa devices or multi-layer substrates can not be represented. If a high mesa technology is used, and the heat generation region is located near the top of the surface, the adiabatic boundary condition applied to the surface will confine the heat spread-

ing from heat source toward the heat sink, thereby reducing the effective spreading area in the substrate and resulting in non-accurate temperature representation. Another example of a device that can not be represented with this model is the multi-emitter power HBT with an interfinger metal structure. The assumption made in the model to neglect the possibility of conduction through the metal leads, makes the representation of this specific structure impossible.

3.3 Numerical Simulator Atar

Numerical simulators are based on finite element or finite difference solution engines, and consist of powerful software tools for producing the discretization of the 3D geometry needed for a solution. These tools generally have a sophisticated graphic user interface (GUI) as an aid in building the 3D model, specifying material parameters, and setting boundary conditions and heat generation regions. A significant drawback in the use of these tools is the time and expertise required to build a 3D model. At this point one of the advantages of Atar [25] becomes important.

Atar is a numerical simulator developed at Carleton University independently of this work. It offers a better solution, specifically in the 3D model creation. Using Atar, for a given technology description and layout information, a 3D model of the device will be generated automatically. Another advantage of Atar is that once the technology has been defined it is very straight forward to generate models with differing geometries or to investigate the effect of small modifications of the technology. Using this advantage, a large variety of device geometry and process variation can be generated in a short time and with

minimum effort.

The physics involved in solving for the conductive heat flow requires solving the partial differential equation Eq. (2.4) explained in section 2.2.2. To produce a numerical solution, an appropriately discretized 3D model must be built: boundary conditions must be applied, and a mathematical model should be extracted and then solved using a non-linear solution method. Due to the requirement of an automatic generation of the model from layout information, the model would be built from the chip surface down into the substrate. In this case, the model is constructed by first defining a surface layer (in geometry file) and then projecting this layer downwards into the substrate.

The model is created automatically with a high density of elements or nodes in the regions of large heat flow or temperature gradients. Such conditions require the creation of a large number of small elements near the heat generation regions in the simulation region. The heat generation regions in an integrated device are generally fairly close to the chip surface and well defined by the device layout.

The basic philosophy of the model building is to create a complex mesh geometry using rectangular blocks of varying sizes. The model will be created in such a way that a block and its neighbors form a relatively simple set of topologies whereby each block can have a maximum of two blocks on any vertical side and, a maximum of four blocks on the top or bottom of the block. From these block topologies, a network of thermal resistances and capacitances will be extracted, representing the thermal equation to be solved.

After the creation of the 3D model, a thermal network is extracted consisting of non-linear resistors and capacitors. Each block is visited and, using the topology of the

connections with the adjacent blocks, a network of resistors to the neighboring nodes is created. The value of each resistor is calculated using the following equation:

$$R_{th} = \frac{1}{k_{th}} \frac{L}{A} \quad (3.4)$$

where A is common cross-sectional area, L is the size of two blocks, and k_{th} is the local thermal conductivity.

The thermal capacitance of the material is used to calculate the value of a capacitor connected from the block node to the ground. As the material properties are, in general, temperature dependent, the resistor and the capacitor values are a function of the node temperature. In Fig. 3.6 the heat flows for a single block are shown, and each heat flow would be represented by a thermal resistor. This block has a single block above, four below, one on the front, back and left sides, and two on the right side. The heat flow associated with the heat capacity of the block is not shown.

A number of boundary conditions can also be placed on the model. Regions of the model can be designated to be at a fixed temperature, as would be used for a fixed back-side wafer temperature. A zero flow boundary condition is used for vertical sides of the simulation region, which implies a mirroring of all geometry on the simulation region at this surface. In addition, constant power generation can be specified for a particular region defined by the emitter geometry.

Figure 3.4 Heat flows between blocks represented by thermal resistors [25].

Once the thermal model has been extracted and the boundary conditions set, the network needs to be solved. The simulation tool allows for either a netlist to be outputted (allowing for solution to be found using an external circuit simulator such as spice), or an internal solver is used. Currently, two steady-state internal solvers are available: a direct one which solves the linear (or non-linear if temperature dependence of the k_{th} is to be considered) set of equations representing the thermal resistor network, and a second one based on a successive over relaxation (SOR) technique. This second solver is particularly useful for large problems, and it also handles the temperature dependence of the material properties. Once the solution has been found, the maximum temperature is obtained and a grey-scale representation of the temperature distribution can be displayed.

The primary use of the Atar tool is the quick building and thermal simulation of

integrated devices in a wide variety of integrated circuit (IC) processes. One of the technologies utilized in this work is a high speed GaAs mesa HBT with interfinger metal connection. To illustrate, a model for such a device, and temperature contours simulated with Atar are shown in Fig. 3.5. Once the model is created, geometrical parameters and material properties can be varied in order to create different scenarios and optimize the required one.

Figure 3.5 A 3D model of GaAs HBT with metal structure, and temperature contours.

Generally, a disadvantage of numerical methods is that computation time increases dramatically with model size because of its dependence on the number of blocks in the model. In order to reduce the model size and the computation time, the device symmetry and mirroring are used throughout this work. More specifically, these concepts are discussed in the following chapters (sections 4.5.1 and 5.2.2) where the numerical model for a specific device of interest is described. Due to this symmetry, the device model can be scaled down, and practical calculation can be performed for a smaller model.

Chapter 4

Normalized Approach for Modeling of Thermal Coupling

4.1 Introduction

Operation integrity and longevity of electronic components are directly affected by their operating temperature. Thus, the object in thermal design is to control the temperature sufficiently to meet the reliability requirements of the system. Achieving this goal requires the co-operation of many engineers on a design team, and the close attention of the thermal designer.

One of the problems facing a thermal designer with regard to the practical design process of a multi-emitter HBT, is the exploration of the trade-off between device geometry and maximum T_{rise} . However, not all geometrical parameters are subject to change in a thermal design process, some of them being set by technological constrain, while others by electrical constrains. Generally, depending on the given constrains, there are three geometrical parameters left that the designer can modify: finger length, number of fingers, and spacing between fingers. Each combination of these parameters leads to a different device temperature operation and to different area occupied by emitter finger array or *array footprint*. The number of combinations of these variables that the designer should model and simulate in order to reach a target T_{rise} with optimum area usage is large and often non-practical. This problem becomes more complicated since multiple finger structure is

employed and T_{rise} is affected by both the self-heating and thermal coupling among fingers.

In order to simplify this task, a new normalized approach to model thermal coupling in multi-emitter structures is proposed. Using this method, the model for a specific structure can be used to predict thermal effects for a large ‘family’ of devices with several similar characteristics. As a first order method, it can be part of a preliminary design stage and reduce the number of models required to be simulated in late stages. Some guidance in design is also provided.

In the following sections of this chapter, coupling effects in a multi-emitter structure are studied, the results are obtained and the normalized approach to model these effects is explained. Normalized approach is applied to predict the maximum T_{rise} for different examples and the error in prediction is calculated. Analytic predictions are verified using numerical simulator. Based on this approach, some scenarios are carried out in which the relation between emitter area and device temperature are investigated.

4.2 Nature of Thermal Coupling

For HBT’s with a single-emitter finger, the thermal effects arise from the fact that all the heat produced in the generation region will be dissipated to the ambient area, thus increasing the lattice temperature, i.e. the self-heating effect [32]. In a multi-emitter finger HBT, another additional mechanism, thermal-coupling effect, is present, causing the inner fingers to be hotter than the outer ones [33].

The nature of thermal coupling in a multifinger device is studied using analytical

solver tsource. Based on the specified model, device temperature distribution in lateral and longitudinal direction are found. Further analysis is emphasized on devices with alignment along long edges and different geometry.

4.2.1 Modeling Assumption

The cross section and top view of a multi-emitter structure which was simulated using tsource is shown in Fig. 4.1. The model used to simulate analytically is generally explained in section 3.2.1. Parameter specifications and assumptions represented here tend to better specify the multifinger device to be simulated.

Figure 4.1 Schematic for thermal simulation structure in semi infinite substrate
a) Cross section view and b) Top view.

A simulation region with infinite dimension in x and y directions, and 75 μm thickness was created with adiabatic top surface and an ideal 300 K boundary condition on the bottom. In other words, the distance chip boundaries are absent in the model. The approximation of the real medium considered is reasonable since, as shown in [28], the adjacent devices in an integrated circuit are essentially thermally isolated when the distance in

between is in the range of or higher than five times the largest dimension of the heat source. Another argument that also supports this assumption is that most high power devices are not typically close to the edge of a chip due to the need for supporting circuitry, pads etc. If instead of assuming an infinite substrate in lateral and longitudinal direction, the adiabatic surfaces had been involved, to enforce the symmetry across both boundaries an infinite series of mirror sources would have been required.

The thermal generation was assumed to be uniformly distributed in the emitter finger area, and power density was taken 2 mW per square micron of the emitter area. Such power density is arbitrary in determining thermal resistance or normalized temperatures, provided thermal conductivity is not a function of temperature. This power density represents typical operating conditions [34] that could correspond to a device with $50 \mu\text{m}^2$ emitter area, $I_C=20 \text{ mA}$ at $V_{CE}=5 \text{ V}$. This is a reasonable assumption given that the device is being operated at moderate current densities.

The thermal generation region was intended to produce a reasonable representation of heat production in a base-collector depletion region under the substrate surface. The depth D_g and thickness T_g of the thermal generation region were set via z extents of the rectangular source, 0.05 and 0.1 μm respectively as shown in Fig. 4.1a, corresponding roughly to the base-collector depletion region [35]. For a transistor, the distance W and L would correspond to the width and length of the emitter stripe and were varied from 0.5 to 1.5 μm and 10 to 100 μm respectively, depending on model requirements, assuming an aspect ratio of at least 10 and the usual alignment along the long edges. The smallest aspect ratio 10 was chosen as the limiting case of a long stripe-emitter geometry often

used in high-speed circuits. Spacing between fingers was specified either as a function of the emitter length or as a constant value depending on the simulation purpose. The minimum spacing of 4 μm was considered according to typical technological limits.

Three different substrate materials were considered: Si, InP, and GaAs, with different thermal conductivity. Si has higher thermal conductivity which creates better conditions for heat to flow through the substrate. Unlike Si, GaAs is considered a poor thermal conductor which leads to higher device temperatures. Dependence of thermal conductivity on device temperature was neglected; in other words, temperature-independent thermal conductivity was assumed, by being able to consider power independent coupling. In applications where T_{rise} has a significant role, this effect can be taken into account later. According to [36], thermal conductivities at 300 K for Si, InP, and GaAs are 1.5, 0.69, and 0.44 W/cmK respectively. Unless otherwise specified, GaAs is used in the following simulations.

4.2.2 Temperature Distribution in a Device, Lateral vs. Longitudinal

Using the analytical solver tsource, for a single finger device with a finger size at 0.5x20 μm , temperature distribution in both directions (first along x and after along y-direction) is calculated and plotted in Fig. 4.2. Both temperature and distance are normalized respectively to maximum T_{rise} , and finger width W , or finger length L , expressing the temperature at each point in space as a function of normalized distance from the center of the finger. As it is observed for both cases, the maximum value of T_{rise} is obtained in the center of the finger and then decreased with the distance from the center point. Specifi-

cally, when distance between fingers in a multifinger device or the distance between devices in a multi chip module is smaller than a particular distance, thermal coupling effect becomes an important factor deserving constant attention. This suggests that large spacing is required for a small coupling effect between fingers or devices. Unfortunately, in practical applications, thermally speaking, it is not always possible to have the desired distance due to other applied restrictions. These restrictions are related, for example, to parasitic capacitances, device frequency operation, limited chip area etc.

Figure 4.2 Temperature profile for single finger device
a) across the finger in x-direction and b) along the finger in y-direction.

By comparing two profiles, temperature drops relatively faster when plotted along y-direction. For example, T_{rise} reaches 10% of maximum value at a distance $x/W = 34$ when plotted in x-direction, or at $y/L = 1$ when plotted in y-direction. This suggests that the contribution of thermal coupling in the lateral direction in maximum T_{rise} is more significant than the longitudinal one. What is stated is based on a single emitter device. To expand the analysis for devices closer to practical applications, a multifinger structure is

simulated and once again the contribution of coupling in T_{rise} is investigated.

4.2.3 Thermal Coupling Along Long and Short Edges

There are different geometries that can be used to locate heat sources in an array. Since the rectangular arrays are widely used in practice, the discussion of this thesis is restricted to these arrays. There are also different ways of arranging the heat sources in a rectangular array, depending on emitter configuration and the number of emitters. In section 4.2.1, a long-stripe emitter geometry was assumed, which limits the number of arrangements practically in two: array with one row of fingers, where fingers are coupled along long edges, and array with two or more rows of fingers, where fingers are coupled in both directions along long and short edges. From a practical point of view, coupling along long edges has many advantages and is mainly used in practice. From a thermal point of view, coupling along short edges is more favorable due to the smaller sharing area. The question then arises: what is the disadvantage, thermally speaking, of using coupling along long edges?

It is noticed that the effect of coupling in T_{rise} is different for different geometries: it depends on finger size, spacing between fingers, and reciprocal emitter location. In order to further proceed with coupling effect on T_{rise} , first the contribution of coupling for emitters coupled with long and short edges is required. For both cases, the problem can be reduced at finding the temperature distribution along x or y axes in a multi emitter device, comparing the profile with a corresponding single finger profile, and calculating the effect of coupling on T_{rise} .

For a multi-emitter device, temperature at each finger is found by using linear superposition as mentioned in section 3.2.2, and the contribution from each source is added at all points in time and space. For example, the temperature profile (Fig. 4.3) can be used to calculate the coupling effect. In this case, if other sources were added and located at $x=\pm 5$ or $\pm 10 \mu\text{m}$, the thermal coupling between each of these sources and the source located in the center would increase the temperature with 27% or 16% respectively.

Figure 4.3 Temperature profile across the finger (single finger device).

In particular, T_{rise} at each finger in a multifinger array may be considered as a sum of two components: a direct T_{rise} due to its own power dissipation referred to as self-heating, which is always present and depends on its own geometry; a series of indirect T_{rise} terms due to thermal coupling with other sources which depends on the geometry of the other

sources and their spatial orientation relative to the original source. Applying this method, thermal coupling contribution in multi-emitter devices, when sources are coupled with each other along long and short edges, is calculated.

The effect of coupling along long edges (Fig. 4.4a) is calculated for three and five finger devices, finger size $0.5 \times 20 \mu\text{m}$, spacing $5 \mu\text{m}$. Temperature profiles are normalized to the maximum T_{rise} for a single finger. From the plots, the maximum T_{rise} for three and five fingers is higher than for a single finger, with a difference which represents the coupling effect. As shown, for a three finger device, coupling along long edges increases the maximum junction temperature by 60%, and for five fingers by 90%. Another effect of coupling is the non-uniformity of temperature distribution between fingers. The center finger is the hottest and the junction temperature in this finger is almost 11% higher than the other fingers for a three finger device, and 13.7% for a five finger device. For other fingers, junction temperature becomes progressively lower toward the outermost ones. This non-uniform temperature distribution is undesired because the base-emitter bias required to turn on a given current decreases with increasing junction temperature, and the central fingers conduct a larger amount of current than others [37]. This non-uniform current distribution reduces the maximum power output achievable compared to the operating condition, where all the fingers share roughly an equal amount of current.

The effect of coupling along short edges (Fig. 4.4b) is also calculated and temperature distribution along the emitter finger is shown for devices with one and three fingers coupled along short edge. Finger size and spacing are the same as per previous example. The maximum junction temperature is increased by 14% for a three finger device. This

number represents the coupling contribution when coupling along short edges is realized, and appears to be four times lower than the effect of coupling along long edges. Also, notice that the maximum temperature for two outer fingers is not located in the center of finger, but shifted toward the center finger.

Figure 4.4 Temperature profile for two different directions
a) across fingers and b) along fingers.

At this point, it can be concluded that coupling along long edges (which is the case in practical applications, since the base resistance becomes too large if coupling along short edges is applied) thermally represents the worst case and governs the temperature increase in a multifinger device. Based on this conclusion, thermal analysis performed in the remaining are limited for devices with fingers coupled along the long edges.

4.2.4 Lateral Temperature Distribution

Lateral temperature distribution (Fig. 4.5) is plotted for a three-finger device for $L=10, 20, 40 \mu\text{m}$. Temperature profiles look similar, but are only shifted in vertical direc-

tion. Operation temperature is increased with L because the amount of total power is increased. If temperatures are plotted in a normalized way (normalized at maximum junction temperature), a comparison between different profiles can be more feasible. The question raised is: can one model created for a specific structure be used to predict the thermal behaviour on other structures with different geometrical parameters, i.e. finger length and width? If the answer is positive, it means that structures with some similar characteristics can be grouped together in a 'family', and the thermal simulation results obtained for a typical 'member' or structure can be generalized and used for the other structures, thereby avoiding repeated simulations.

Figure 4.5 Temperature profiles for three finger device with different L .

To answer this question, first a target error should be specified for a comparison of the results and an evaluation of the method. A value of 2% error, approximately 1/3 of expected measurement uncertainty [35] on T_{rise} predicted is considered, which means that the difference between predicted temperature from a normalized approach and the simulated one at a given point should be inside this range. If this requirement is achieved, the method may be good enough to be used as a first order one.

After the target error is specified, the next step toward the solution is to compare normalized temperatures for a specific device at a different finger length. To illustrate, a single finger device is simulated for three different L : 10, 20, and 40 μm . As shown in Fig. 4.6a, temperature is normalized to maximum T_{rise} and plotted as a function of spacing (or distance in this case) in the lateral direction. Note that only half of the profile is shown here. Profiles are similar but clearly not identical and the results are promising. If the model for $L=20 \mu\text{m}$ is used to predict the temperature in other devices for example with $L=10$ and 40 μm , an error may be introduced due to the differences among profiles. To find the error in this case, the magnitude of difference between temperature profiles $|T_{20}-T_{10}|$ and $|T_{20}-T_{40}|$ is calculated at each point in the lateral direction and plotted (Fig.4.6b).

Considering temperature profiles obtained for $L=20$ and 40 μm (Fig. 4.6a), if T_{20} is used to predict the coupling effect in a three-finger structure with the $L=20 \mu\text{m}$ and spacing at 10 μm , the thermal effect of coupling with adjacent source is in the range of 17%, representing the increase in temperature above the one caused by self-heating. Following the same procedure, if T_{40} curve is used to predict the coupling effect on a similar device

($L=40\text{ }\mu\text{m}$, number of fingers and spacing are kept unchanged), the predicted coupling effect in T_{rise} will be 24%. If instead of T_{40} in the last case, T_{20} is used, the error made in the coupling effect prediction can be found from the magnitude of difference $|T_{20}-T_{40}|$, which for this example is in the range of 7%.

Figure 4.6 Normalized temperature as function of distance in lateral direction
a) Normalized temperature and b) Magnitude of error.

Generally speaking, if T_L is used to predict coupling effect in a device with finger length different from $L\text{ }\mu\text{m}$, a certain amount of error is introduced. According to Fig. 4.6b, the magnitude of error depends on the distance between fingers in x direction or, more specifically, from spacing between fingers. By increasing spacing, the amount of error is decreased because coupling contribution becomes weaker and the absolute value of the thermal coupling contribution in T_{rise} is decreased. For examples shown above, the maximum value of error is in the range of 7 and 8.2% respectively, and is reached at $x=5\text{ }\mu\text{m}$ which, unfortunately, corresponds with a practical desired spacing.

Compared with the error target specified, it can be noticed that this method of representation does not fulfill the requirement. One source of error is the fact that the temperature profiles are related to finger length, while spacings between fingers are not. If spacing is also expressed in terms of finger length, the error can be reduced.

4.3 Normalized Approach of Lateral Coupling

Inter-finger spacing can be expressed as a proportion of the finger length, leading in the normalized approach. In such case, both the temperature profile and the distance from the center of the finger will be related to the same parameter, that is finger length. Even though the temperature distribution for a given device at a fixed bias condition is unique, different temperature profiles can be obtained, depending on independent variable x in the relation $T(x)$. In the previous example shown in Fig. 4.6a, the distance from the center of the finger (in μm) is the independent variable, and $T(x)$ is the temperature profile. If instead, x is normalized and expressed as a function of finger length L with the relation $x_n = x/L^\alpha$, different profiles can be obtained depending on the value of the *normalization coefficient* α . The intention is to find a value for α and normalize the distance x in order to make the curves in Fig. 4.6a look similar, so that one can be used to predict the other. Based on this, in the following part of this section, different normalization coefficients are examined in order to minimize the error in thermal coupling prediction.

For three different $L = 10, 20$, and $40 \mu\text{m}$, normalized temperature profiles are plotted as a function of normalized distance. As a starting point, $\alpha=1$ was chosen, meaning that the distance is normalized to L , giving a straightforward relation, easy to be used in

practical calculations. This case may be considered as a particular case of general expression x/L^α . Temperature profiles and the magnitude of error $|T_{20}-T_{10}|$ and $|T_{20}-T_{40}|$ are shown in Fig. 4.7.

Figure 4.7 Normalized temperature as function of normalized distance
a) Normalized temperature and b) Magnitude of error.

By comparing the magnitude of error in this case with results shown in Fig. 4.6b, it can be observed that the value of maximum error is increased, reaching a 12% range for $|T_{20}-T_{10}|$ (compared with 8.2%). Interestingly, it is noticed that the location of maximum error is shifted to the left, that is the maximum is reached at $x/L=0.05$ or converted into distance from the center of the finger; for $L=10\mu m$ it gives $x=0.5\mu m$, and for $L=40\mu m$, $x=2\mu m$. In both examples, the distance where the maximum error is reached is less than the typical technological limit of spacing between fingers. On the other hand, if spacing between fingers is given, for example $5\mu m$, this corresponds to $x/L=0.5$ ($L=10\mu m$) and $x/L=0.125$ ($L=40\mu m$), and the error related is 3.7 and 6.1% respectively. These values are

still high, showing that the coefficient of proportionality used to relate spacing with finger length is not the optimum one.

In order to identify the optimum value for normalization coefficient, the magnitude of error should be minimized. Finding an analytic expression for the relation between magnitude of error and x/L^α is a difficult task. For this purpose, a graphical based method can be used. The magnitude of error for different values of α is calculated, plotted, and the results are compared to find the function that provides the minimum area under the curve. Normalization coefficient α that corresponds to the minimum error function can be considered as the optimum one.

Using the same example (three different L), magnitude of error for $|T_{20}-T_{10}|$ and $|T_{20}-T_{40}|$ for α between 0.58 and 0.72 are plotted in Fig. 4.8. Note that the interval where α is varied was selected by trial and error, based on the criteria of minimum area under the curve for both functions. Since these functions are independent of each other, a different value for α optimum is expected for each case. For practical purposes, only one value for α should be selected which will satisfy both cases, trading some accuracy in each of those. Finally, from the plots shown, $\alpha=0.62$ was considered the optimum value.

In both cases for $\alpha=0.62$, the maximum value of error is in the range of 5% and is reached approximately for $x/L^\alpha=0.1$. Referring to a practical spacing of 5 μm , if converted into a normalized term it gives different values for different lengths. For $L=40 \mu\text{m}$ example, corresponding normalized spacing is $x/L^{0.62}=0.507$ and the related error is in the range of 2% which can be considered as satisfactory since it is in the specification range.

For the length considered here, this case represents the worst one, since for shorter emitters, normalized spacing is greater, reaching the value of 1.19 for $L=10\ \mu\text{m}$, while the error associated with it is less than 0.5%. As can be noticed, by increasing spacing (in μm), the corresponding point in normalization axis will be shifted to the right which is valid for all emitter lengths. According to plots shown, if normalized spacing is greater than 1, the error associated remains less than 1% for all practical distances.

Figure 4.8 Magnitude of error for different values of $\alpha=0.58$ to 0.72
a) $|T_{20}-T_{10}|$ example, b) $|T_{20}-T_{40}|$ example.

Using the optimum normalization coefficient $\alpha=0.62$, for the same example with three different L , normalized temperature profiles in the lateral direction and the magnitude of error for $|T_{20}-T_{10}|$ and $|T_{20}-T_{40}|$ are plotted in Fig. 4.9. Profiles are very similar and the error is small in both cases.

Figure 4.9 Normalized temperature as function of normalized distance for $\alpha=0.62$
a) Normalized temperature and b) Magnitude of error.

For a first order estimation, another normalizing coefficient can be used in order to have a more practical normalization term, trading away from accuracy. If normalization coefficient is chosen to be $\alpha=1/2$, the normalization term can be expressed in the form of $x/\sqrt{L}$. Normalized temperature profiles and the magnitude of error for $|T_{20}-T_{10}|$ and $|T_{20}-T_{40}|$ in this case are shown in Fig. 4.10.

As expected, the magnitude of error is somewhat higher compared with the previous case where $\alpha=0.62$, particularly in the interval $x/\sqrt{L}=1$ to 4. However, even in this interval the magnitude of error does not exceed the 2% value, which can be considered as a good approximation for a first order estimation. With these results, it becomes evident that by normalizing the spatial temperature distribution by $\sqrt{L}$ or $L^{0.62}$, the distributions become similar, so that the results of 10 and 40 μm in the physical, non-normalized space,

can be predicted using the 20 μm result.

Figure 4.10 Normalized temperature as function of normalized distance for $\alpha=0.5$
a) Normalized temperature, b) Magnitude of error.

As already mentioned, when normalized approach is employed to predict T_{10} and T_{40} using T_{20} , the prediction error depends on the normalized distance between fingers. To evaluate the prediction error, for each specific location, finger location should be identified with the distance from the other finger; the distance in μm should be converted into normalized distance, and the magnitude of prediction error vs. normalized distance can be used to find the corresponding value for the error.

This procedure is applied in an example where the prediction error is found at four different locations (two different emitter lengths $L=10$ and $40\mu\text{m}$ and two different spacings between fingers, 5 and $10\mu\text{m}$). Using normalization coefficient $\alpha=0.62$, the magnitude of error shown in Fig. 4.11 will be the same with the plot shown in Fig. 4.9b, but the

interval of interest is magnified. Considering $x=L^{0.62}$ as unit of spacing U_s ($U_s=1$ for $x=L^{0.62}$ μm), four different normalized distances are obtained. Using the corresponding curve for each case, the error value is found and calculation results are summarized in Table 1.

Figure 4.11 Magnitude of prediction error vs. normalized length.

Table 1: The magnitude of error evaluated in different locations

Point #	Emitter length (μm)	x (μm)	$L^{0.62}$ (μm)	Unit of spacing $x/L^{0.62}$	Error (%)
1	40	5	9.85	0.51	1.1
2	40	10	9.85	1.01	0.1
3	10	5	4.17	1.19	0.4
4	10	10	4.17	2.4	0.3

The error value is lower than the required specified one, and this is true for all cases. The maximum value of error corresponds to a maximum L with minimum spacing representing the worst scenario. Even in this case, the error is in the range of 1.1% which is placed well within the specification.

Up to this point, the prediction error of thermal coupling effect was analysed for examples where two heat sources were (implicitly) employed. Applying this method in cases of more than two heat sources, the contribution of all heat sources at a given location are added up together, and the total error should be equal to the sum of all these errors, or, in other words, the error should be accumulated. In order to prove this statement, the following example is used. Assuming a multi emitter device with a schematic of cross section shown in Fig. 4.12. Fingers are spaced at U_s , two finger lengths (10 and 40 μm) are considered, and the other parameters remain unchanged from the device model explained in section 4.2.1. The requirement is to calculate the predicted error obtained when maximum T_{rise} at heat source zero (T_{10} or T_{40}) is found using T_{20} as a predictor.

Figure 4.12 Coupling effect and the error associated when maximum T_{rise} is predicted.

For simplicity, fingers are assumed to be added one by one, and the maximum T_{rise} and the error associated at finger 0 in each step are calculated using the plot shown in Fig. 4.9b. Starting by adding finger one; the presence of finger one will increase the temperature at finger 0. Since T_{20} is used instead of T_{10} (or T_{40}), the temperature obtained will have a certain degree of error which can be found at distance U_s in magnitude of error plot. Lets call this e_1 and plot the value in Fig. 4.13. Adding another finger, for example finger 2, the temperature at finger 0 will further increase due to coupling with source 2. This increment will be predicted and the error associated, e_2 in this case, can be calculated from the error plot at distance $2U_s$. The sum of $e_0=e_1+e_2$ gives the total error caused by sources 1 and 2. When source 3 is added as well, e_3 associated with this source is added, giving the total error at source 0 in the value of $e_0=e_1+e_2+e_3$, and so on. In general, when n sources are employed, the total error accumulated when maximum T_{rise} at source 0 is predicted as described, can be found as

$$e_0 = \sum_{i=1}^n e_i \quad (4.1)$$

If the device has a symmetry line as shown (Fig. 4.12) and the source 0 is considered to be located in the center of the array of sources, the total error will be twice e_0 , since sources from the other side shown with dashed lines will have their own coupling contribution in temperature increase.

Following the above procedure, accumulated error in maximum T_{rise} prediction for a center finger in a 23 finger device is computed, when T_{20} is used as a predictor. Two dif-

ferent values of normalization coefficient $\alpha=0.62$ and $\alpha=0.5$ were considered. Results are plotted for half plane (in the right side of symmetry line there are 11 fingers plus the middle one) and shown in Fig. 4.13. The expected prediction error is higher for devices with a large number of fingers. Also, by comparing both figures, accumulated error has a strong dependence on normalization coefficient α . For example, for 23 fingers and $L=10\text{ }\mu\text{m}$, the accumulated error is in the range of 5% for $\alpha=0.62$ and 15.5% for $\alpha=0.5$.

Figure 4.13 Accumulated error for 23 finger structure
a) for $\alpha=0.62$ and b) for $\alpha=0.5$.

After analysing the results obtained in this section, it can be stated that coupling between sources in the normal multi-finger configuration can be handled in a normalized way by expressing the inter-finger spacing as a proportion of the finger length. The normalized result for T_{20} is quite close to T_{10} and T_{40} , therefore, this case can be used as the predictor. Using the normalization coefficient $\alpha=0.62$, the prediction error is in the specification range even for structures with a large number of fingers. For practical applications,

even $\alpha=0.5$ can be used trading to some extent from accuracy.

4.4 Normalized Approach to Prediction of Maximum T_{rise}

Once the normalization coefficient is known, coupling effect can be handled in a normalized way. If fingers with a given length are located accordingly to spacing units or a proportion of this number, the maximum T_{rise} will have the same value for different lengths. This makes possible the use of the normalized approach in studying the effect of coupling for different devices and the impact it has on maximum T_{rise} .

4.4.1 Geometry Dependence of Maximum T_{rise}

As an example, the normalized approach is used to investigate the relation between maximum T_{rise} at the center finger and the number of fingers for different spacings between fingers. Spacing is calculated based on normalization coefficient $\alpha=0.62$ and varies in the range of 0.5 to 4 U_s . For $(L = 20) \Rightarrow (U_s = 6.4\mu m)$ and the distance between fingers will be varied in the range of 3.2 to 25.6 μm . Maximum device temperature found in each case is normalized to maximum temperature for a single finger device and the results are shown in Fig. 4.14.

Plots show that the operating temperature is significantly higher in a multiple finger device. Increasing the number of fingers, necessitates the increase in the maximum T_{rise} due to the effect of coupling between sources. In the first part of each curve, the maximum temperature increases fast with the number of fingers. Such increase can be explained by the fact that the total distance between center and outer finger is small, and

the contribution from outer finger in maximum T_{rise} in the center finger is strong. Further increase in the number of fingers brings a slower increase in temperature, which means that the coupling effect becomes weaker. Reason is that a larger distance diminishes the interaction between center and outer fingers. For further increase of the number of fingers, the curves are flattened and there is no increase in temperature. This means there is no practical effect on maximum device temperature. This can be noticed for a $4U_s$ curve for more than five fingers. In this case, the distance between center and outer source will be $2 \times 4U_s = 51.2 \mu m$.

Figure 4.14 Normalized center T_{rise} for $L=20 \mu m$.

Another factor that has a strong effect on maximum device temperature is the spacing between fingers. Maximum T_{rise} is higher for devices with small spacing between fin-

gers, which signals a stronger coupling effect being present. For example, for five fingers with spacing $0.5U_s$, maximum device temperature is 2.2 times higher than for a single finger device (compared with 1.15 for five fingers with a spacing of $4U_s$).

4.4.2 Effect of Normalization Case

Based on the results obtained in section 4.3, the effect of coupling for devices with different geometries is studied in a normalized way and the results are compared. Following the same procedure explained in the previous section, normalized maximum T_{rise} is calculated for six different structures varying the number of fingers from one to 11. For each structure, three different L were considered (10, 20, and 40 μm); spacing between fingers was set equal to U_s ; obtained results are shown in Table 2. For each case, both the absolute value of T_{rise} and the normalized one to a single finger are shown. Comparing the normalized results for each structure, the agreement between T_{20} , the predictor, and two other temperature values is strong.

Table 2: Maximum Trise comparison for structures with three different finger length

Finger length (um)	Trise (K) \ Norm Trise					
	Single	3-finger	5-finger	7-finger	9-finger	11
10	19.00\1	27.95\1.47	32.79\1.72	35.36\1.89	38.26\2.01	40.03\2.11
20	21.99\1	32.48\1.48	38.27\1.74	42.04\1.91	44.72\2.03	46.73\2.12
40	24.76\1	36.53\1.47	43.08\1.74	47.27\1.91	50.15\2.02	52.21\2.11

The analysis is extended for different spacings between fingers, and a normalized maximum T_{rise} for $L=10$ and 40 μm is calculated. The obtained results are superimposed

with results for $L=20$ from the previous section and plotted in Fig. 4.15.

Figure 4.15 Normalized center T_{rise} utilizing the normalized approach.

Again, the agreement between predicted and simulated results is excellent even for a wide variety of spacings, thus demonstrating the capability of applying the normalization approach for a larger ‘family’ of devices. Also, this agreement demonstrates that what stated at the end of section 4.3, is true for all practical spacings.

4.5 Verification Using Numerical Simulation

Up to this point, the analytical solver has been used to find the temperature distribution and/or maximum T_{rise} in different devices. In order to verify the results, numerical simulator Atar is used to simulate the same devices and find the temperature distribution.

Simulations were performed for few examples, and the results are compared with analytical ones.

A 3D numerical model for Atar is explained in general in section 3.3. Some details related with the specific device include the following: geometrical parameters, boundary conditions, and technological parameters for materials used should be specified into Atar input files. Although the device used is identical to the device described in section 4.2.1 for an analytical model, some modifications are required for its implementation in Atar. These changes are related to the method involved when numerical approach is used to solve heat conduction equation and to find temperature distribution in the device. In the next paragraph, changes made for simulation region, position of finger array into the simulation region, and boundary conditions applied are described.

The schematic of power HBT (Fig. 4.16) represents the layout, the 3D section of device, and the Atar model for numerical simulation. The geometrical assumption regarding infinite lateral and longitudinal substrate dimensions made for an analytical model can not be applied for a numerical one; instead, finite dimensions should be involved. The size of the simulation region was chosen based on the criterion of large size in order not to confine the heat distributed from the heat source toward the heat sink. A simple calculation was made based on the spreading angle recommended, and the effective spreading area in the substrate, when substrate thickness is known. Accordingly, a simulation region was set to $256 \times 256 \mu\text{m}$ (with thickness $75 \mu\text{m}$). A set of simulations for different simulation regions has also shown that the result was consistent and unaffected by larger region dimensions.

Device symmetry, with respect to x and y axes combined with the method of images, is used to scale down the model (Fig.4.16b). This symmetry of position causes heat flow to cancel exactly along the desired adiabatic planes, which in this case are x - z and y - z . The other top and side surfaces were also considered as adiabatic (non-conducting). An ideal 300 K boundary condition on the bottom (the same as per analytical model) was applied.

Figure 4.16 Power HBT model for numerical simulation
 a) Schematic of layout, b) simulated section, and c) Atar model.

The position of the thermal generation region specified in geometry file, was offset toward one corner of the simulation region (shown in x-y plane) which corresponds to the center of finger array, taking into account the symmetry of sources relative to the adiabatic boundaries. In z direction, the generation region was set near the top of the substrate, basically in the same position specified in the analytical model. The other parameters, as thermal generation per unit of area, temperature-independent thermal conductivity, and substrate material were kept the same as described for the analytical model.

The model described above was used to simulate numerically examples with three, five, and seven fingers, $L=10, 20$, and $40\text{ }\mu\text{m}$, and spacing $L^{0.62}$, while the results are plotted in Fig. 4.17.

Figure 4.17 Simulation results-numerical vs. analytical.

For a comparison of analytical results for the predictor case ($L=20\text{ }\mu\text{m}$ and spacing at U_s) are plotted with a solid line in the same axes with numerical ones. According to Fig. 4.17, the agreement between analytical and numerical results is very good. Both simulators show that when temperatures are normalized, and spacing between fingers is expressed in terms of finger length, the agreement between T_{20} , T_{10} , and T_{40} is excellent. The model created for a given device geometry can be used to predict the maximum temperature rise in similar devices with different geometrical parameters.

4.6 Design Scenario where Optimal Use of Area is Required

The relation between constant emitter array areas (*footprints*) and normalized T_{rise} curves can be used in the preliminary stage of design process in order to approximately define the configuration required. It should be noticed that the usage of normalized T_{rise} curves assumes that T_{rise} for the single finger device is known for every practical length. In this section, constant area curves for a different number of fingers and spacings are calculated and superimposed with the normalized T_{rise} curves. Using this method, few design scenarios where a trade-off between active area, spacing, number of fingers, and finger length is analysed.

4.6.1 Constant Footprints

Normalized T_{rise} plots given in Fig. 4.16 represent the maximum device temperature for a different number of fingers and spacing when length is kept constant at $L=20\text{ }\mu\text{m}$. Footprint can be found for each specific number of fingers and spacing using the sim-

ple relation

$$F_a = (N - 1)LS \quad (4.2)$$

where: F_a is finger array area (μm^2), N is number of fingers in the array, L is the finger length (μm), and S is spacing between fingers in the array (μm).

In expressing spacing in terms of finger length, Eq. (4.1) can be rewritten

$$F_a = (N - 1)L(L^\alpha) \quad (4.2a)$$

Considering $L^\alpha = U_s$ as *unit of spacing* and the product $LU_s = U_a$ as *unit of area*, the Eq. (4.2a) takes the following form:

$$F_a = (N - 1)U_a \quad (4.3)$$

This equation is used to calculate the emitter array area for a different number of fingers and spacings. An example, where two different finger arrays with equal area but different number of fingers, is shown in Fig. 4.18.

Figure 4.18 Example for equal area arrays with two and three fingers.

In this example, both arrays have the same footprint, but the maximum T_{rise} corresponding to these devices is different since they lie in two different curves as shown in Fig.

4.15 (for two fingers at U_s curve and for three fingers at $U_s/2$ curve). Similarly, equal emitter array areas, expressed in terms of unit of area, from curves with different spacing are calculated and connected together, thereby creating a family of curves representing constant finger array footprints. This set of curves is superimposed with normalized T_{rise} curves and shown in Fig. 4.19. This plot can be used to relate the number of fingers, length and spacing with the maximum T_{rise}

Figure 4.19 Equal area curves superimposed with normalized T_{rise} curves.

4.6.2 Design Scenarios

Let us assume a fixed finger width $W=0.5 \mu\text{m}$ set by technology constraints. Two different scenarios are investigated. For the first scenario: a fixed active area is given, for

example $125 \mu\text{m}^2$; since the width is known, the total emitter length can be calculated as $L_f = 125/0.5 = 250 \mu\text{m}$. One possible configuration for this example can be five fingers with $L = 50 \mu\text{m}$. Then, using the plots in Fig. 4.19, spacing between fingers S varies, and the footprint area F_a and T_{rise} corresponding to each spacing is found. Results are summarized in Table 3 where the unit of spacing $U_s = 50^{0.62} = 11.3 \mu\text{m}$, and the unit of area $U_a = 50^{1.62} = 565 \mu\text{m}^2$. Following the same procedure, other configurations for finger array can be examined. As it can be noticed, in order to decrease the maximum temperature, the footprint area should be increased representing the trade off area-temperature. Depending on the required operation temperature for the device, the designer can estimate the necessary spacing and the footprint area.

Table 3: Example of trade off between footprint area and T_{rise}

Parameter	Array configuration $5 \times 0.5 \times 50 \mu\text{m}$				
$S (\mu\text{m})$	$0.5 U_s$	$1 U_s$	$2 U_s$	$3 U_s$	$4 U_s$
$F_a (\mu\text{m}^2)$	$2 U_a$	$4 U_a$	$8 U_a$	$12 U_a$	$16 U_a$
T_{rise}	2.25	1.75	1.4	1.25	1.2

For the second scenario: a fixed footprint area is given, such as $2260 \mu\text{m}^2$, which expressed in a normalized way for finger length $L = 50 \mu\text{m}$ is $F_a = 4 U_a$. By varying the number of fingers and spacing, a different active area can be achieved, and, hence, different T_{rise} corresponding values. Using plots shown in Fig. 4.19, for configurations selected in these examples, spacing and temperature are found for each number of fingers, while results are summarized in Table 4.

Table 4: Example of trade off between active area and T_{rise}

Parameter	$F_d = 2260 \mu\text{m}^2 = 4 U_d$ (Finger area $0.5 \times 50 \mu\text{m}^2$)				
Number of fingers	3	4	5	6	9
$S (\mu\text{m})$	$2 U_s$	$1.3 U_s$	$1 U_s$	$0.6 U_s$	$0.5 U_s$
T_{rise}	1.25	1.5	1.75	2	2.8

Results show that by increasing the number of fingers, spacing is reduced and the temperature is increased. Depending on temperature limits, a thermal designer can use the footprint area available differently and place a different number of emitters inside the area (without changing the power density in the active area), thus trading from the operation temperature of the device.

4.6.3 Finger Geometry Variations at Constant Area

Another frequently asked question regarding practical design application problems is: should one split a longer emitter into fingers? If so, how are the device T_{rise} and finger array area affected by this splitting? According to [38], typical microwave power HBTs have multi-emitter fingers for the purpose of increasing the total device power output. If the fingers are short, more fingers are required for a given total emitter length. Such requirement increases the chip size, causing the circuit to suffer distributive effects as the size approaches the wavelength corresponding to the operating frequency [39]. However, the emitter length cannot be made too long, either. When the finger is lengthened to an extreme, the current distribution ceases being uniform, and a significant portion of an

emitter finger does not contribute to the transistor action. Therefore, though the total designed chip area is reduced, the effective emitter area and, hence, the total output power decrease. These arguments show that splitting a long finger is a requirement, and starting from this point, further investigation of a trade-off between temperature and array area is essential in order to reach a desired temperature with a given total emitter length.

In the example below, a given total emitter length L_f (two cases are presented with respectively $L_f=100\ \mu\text{m}$ and $L_f=50\ \mu\text{m}$) is split into different numbers of fingers, where the finger length in each case is calculated by dividing the total length with the number of fingers, maintaining the total length and the active area constant (since width is kept constant). Spacing between fingers is calculated with the intention of having a constant value of T_{rise} for the entire set of structures (one to ten fingers). Simulations were performed for three different operation temperatures, starting with a temperature equal to T_{rise} for a single finger, and proceeding with higher temperatures of 1.25 and $1.5 T_{rise}$ respectively. The results are shown in Fig. 4.20. A specific example to illustrate the procedure is shown, where $L_f=100\ \mu\text{m}$, number of fingers $N=5$, and $T_{rise}=T_{rise(100)}$ is required. To realize such a device, according to the corresponding curve (Fig. 4.20a), fingers should be spaced at $S=18\ \mu\text{m}$. If instead $T_{rise}=1.5 T_{rise(100)}$ is required, spacing should be $S=5\ \mu\text{m}$. In the same way, spacing can be found for different total finger lengths (Fig. 4.20b).

Since the coupling is function of spacing, in order to reduce the coupling component and maintain a constant temperature, such as one equal to maximum T_{rise} for a single finger (Fig. 4.20a), a large and unpractical spacing is required. The largest spacing value

of $34\text{ }\mu\text{m}$ is required, if the whole emitter is split into two fingers with a finger length of $50\text{ }\mu\text{m}$ each. This peak is related to a relatively high value of L at the region of $N=2$ to 4. Following that, the required spacing is decreased by increasing the number of fingers (more fingers mean shorter finger length), therefore reaching a moderate value of $S=10\text{ }\mu\text{m}$ for 10 fingers with $L=10\text{ }\mu\text{m}$. This observation suggests that from the T_{max} point of view, a larger number of fingers should be more preferable. Another interesting result is that by allowing a higher operation temperature, for example moving from $T_{rise}=27.7^\circ\text{C}$ to $1.5T_{rise}=41.5^\circ\text{C}$, required spacing is reduced drastically, and the peak value (or the largest spacing required) is shifted to the right at five fingers.

Figure 4.20 Spacing vs. number of fingers for two different total L .
a) $L_f=100\text{ }\mu\text{m}$ and b) $L_f=50\text{ }\mu\text{m}$.

The information obtained from plots in Fig. 4.20 is not complete, since only spacing between fingers is involved. The next step toward a more practical situation, is to investigate the relation between finger array area and finger splitting. To illustrate, struc-

tures with different L_f (200, 100, and 50 μm) are split into different numbers of fingers using the same technique as per previous example, and maintaining the total length constant. Simulations were performed for constant operation temperatures equal to maximum T_{rise} for a single finger. The finger array area is found using Eq. (4.2) which gives the relation between spacing found from simulation, the number of fingers, and the finger length for each case. Normalizing this area to the single emitter one, the ratio of area increase, later referred to as *array factor*, can be found. The relation between array factor and the number of fingers is shown in Fig. 4.21 for the examples mentioned above.

Figure 4.21 Array factor vs. number of fingers for constant temperatures.

There is a similarity between spacing and array factor curves. Both plots show that splitting a long finger is associated with a larger array area and again, splitting into a small

number of fingers is the worst case. Interestingly, when the total length is reduced by a factor of two, spacing and/or array factor are reduced approximately by the same factor, suggesting the idea of handling this analysis in the normalized way and generalizing the results. In this case, the results found for one specific structure with a given total emitter length can be used in thermal design process to estimate the geometrical parameters and predict the operation temperature for the entire group of devices sharing similar characteristics, without a need for specific calculations in each case.

For an investigation of this approach, spacing between fingers for structures with $L_t=200, 100, \text{ and } 50 \mu\text{m}$, and for three constant temperatures (T_{rise} , $1.25T_{rise}$, and $1.5T_{rise}$), is normalized to corresponding finger length (each specific number of finger has a corresponding length), and the results are shown in Fig. 4.22.

Figure 4.22 Normalized spacing vs. N for three different constant T_{rise} .

For a given operation temperature, curves are very similar, and simulation results found for a structure with emitter length $L_f=100\text{ }\mu\text{m}$ can be used to estimate or predict some design parameters for structures with emitter length 200 and 50 μm . The difference in magnitude between results obtained for emitter lengths 200 and 100 μm , and 100 and 50 μm represent the prediction error. The error is in the range of $\pm 5\%$ which may be considered acceptable. This error is higher than the error in the previous case (recall 2% value was considered for normalized approach in section 4.3). The reason for such higher error is that the range of length variation considered here is much larger than in the previous case, therefore a higher level of error should be tolerated.

4.6.4 Use of the Normalized Approach for Optimization

If the overall optimization of temperature or layout is required, there are potentially many different geometrical constraints and/or degrees of freedom, which may lead to different optimal solutions. In general, if a certain value of absolute T_{rise} is to be maintained in the device, first of all this temperature should be converted into normalized temperature that depends on the reference temperature. Should be mentioned here that in the work presented in this thesis, the temperature in a single finger is considered to be the reference for addition of more fingers or splitting into smaller fingers, and that in an optimization scenario, reference would have to be chosen with respect to which normalized temperature could be expressed. After that, assuming that the temperature corresponding to each finger length and power density is known (simulated for the device with single finger and different finger lengths), normalized temperature for each finger length can be

found, and the problem can be transferred into normalized temperature domain; and finally, the optimization methods presented in sections 4.6.2 and 4.6.3 can be used.

In addition, the application to prediction of absolute temperatures in a given technology requires specific R_{th} values for a set of representative structures. If this is the case, based on the assumption of constant power density and the R_{th} model for each finger described in 2.2.3, the absolute value of temperature can then be found.

4.7 Width Representation

At high aspect ratios, the dependence on finger width is low, thus it is expected that the methods presented in this chapter are applicable to varying the width W without modification. To prove this statement, devices with $W=0.5$, 1, and 1.5 μm , and fixed L are simulated and the normalized temperature as function of the number of fingers is plotted (Fig. 4.23).

The agreement between results is not perfect, but acceptable for a first estimation method. If 1 μm width is used instead of 0.5 μm (or 1.5 μm), the error increases with the number of fingers, reaching the highest value of 4.4% for 11 fingers. Although this error value is high, its elimination in a design process is not difficult, because the width is usually a fixed parameter for a given technology and does not require a constant change. More accurate results can be obtained by setting the real width and getting the corresponding curves.

Figure 4.23 Finger width as function of number of fingers.

4.8 Summary

In this chapter, the nature of thermal coupling in a multiemitter HBT device was investigated, coupling in lateral direction was found to be the most significant, and is considered further. A new normalized approach to model lateral coupling in a device was introduced and normalized results were examined to find the possibility of predicting thermal behaviour in a entire 'family' of devices, without the need to simulate each specific structure. Structures with three different emitter lengths have been used to demonstrate the method, and the error in prediction was calculated. Less than two percent error in temperature prediction suggests that coupling between sources in the normal multi-finger configuration can be handled in a normalized way by expressing the inter-finger spacing in

proportion to the finger length. The normalized result for T_{20} is quite close to T_{10} and T_{40} , thus this case can be used as the predictor. Results have been verified using a numerical simulator, and the agreement between analytical and numerical is in the range of $\pm 2\%$. Based on this approach, some design scenarios have been investigated, where 'optimal' use of area or finger strategy is required. In addition, some guidance in design process has been suggested.

Chapter 5

The Effect of Interfinger Metal on Temperature Distribution

5.1 Heat Removal Techniques in Power HBTs

As it was described in section 3.1, the high power potential of HBT is limited in a practical device operation by thermal handling capability [23] rather than the electronic properties of the semiconductor. In other words, the temperature rise due to self-heating and thermal coupling between fingers in a multi-finger emitter device limits the device performance of a GaAs HBT before the electronic limitations are reached [22]. Because of this, the thermal designer is, therefore, most often concerned with removing heat from heat-dissipating components. Based on the fact that heat is removed from the components mainly by conduction, the scope of this work is focused on conduction methods, and a brief overview of different techniques to reduce the maximum device temperature is presented.

The effect of emitter size and spacing between fingers on maximum device temperature is described in detail in previous chapter. Those are the first considerations for the device designer. From a thermal point of view, according to [19], the ideal emitter fingers are narrow and long, and divided by a large space in between. These techniques are closely related with the chip area which is limited in size. Combining these techniques with heat removal methods will improve operation temperature of the device significantly.

Recently, the plated heat sink (PHS) structure [7] and bathtub structure [40] have

been used to improve the thermal handling capability in GaAs power devices. These structures are mainly developed to reduce the junction temperature by thinning the substrate [21]. They generally use substrate thinned down to several tenths of microns. On the other hand, a thin substrate has some technical difficulties. According to [16], any thickness below 50 μm becomes progressively more difficult to handle.

Variations of the standard MMIC process have been proposed to reduce the self-heating thermal resistance. One of the approaches used is based in the idea of localized thinning of the substrate [7]. In this technique, the substrate is etched to about 5 μm only at a location directly below the device, but retains its typical thickness elsewhere.

A second alternative to reducing the self-heating thermal resistance is to connect all the emitter fingers using metal interconnection and then mount the chip face down to the heat sink (Fig. 5.1). This idea is usually embodied in the “Flip-Chip” approach [7]. The heat dissipated travels from a dissipation volume to the gold (instead through the substrate) and then to the heat sink.

Figure 5.1 Simplified schematic of “Flip-Chip” approach.

Another alternative of the topside idea is to lay a thin layer of high thermal conductivity material placed immediately over the device, and surrounding the top surface to

spread the heat. A diamond-like (electrically insulating and low dielectric constant) thin film could be used as a high thermal conductivity material [19]. The heat is allowed to spread outlying top surface GaAs areas and then move through the 75 μm thick substrate to the heat sink.

Another factor that reduces the maximum temperature in a device is the metal layer that connects the fingers together electrically (Fig. 5.2). It creates the conditions for heat to be redistributed through the metal. In this case, the heat flow in metal layer is a by-product which will help to redistribute the heat between fingers, hence reducing the maximum temperature in the device. Metals such as gold or aluminum, with high thermal conductivity are often used for a realization of the metal interconnection. Although all heat is eventually dissipated through the substrate, some heat generated in the middle finger, with the highest temperature, is initially transported through the metal interconnection to the other fingers where the temperature is lower, and from there through the substrate to the heat sink. In this case, the metal layer will create the conditions for the heat flux to flow in all directions from the hottest to the coolest point, and the result will be lower maximum temperatures in the middle finger area and higher temperatures in other fingers.

Figure 5.2 Three finger emitter structure with metal interconnection.

Since this processing technique is widely used in the MMIC multi-finger transistor technology, the thermal aspects related with this structure were studied in more depth. In addition, only full numerical models are available to simulate and characterize the thermal effect of metal interconnection structure in device junction temperature.

The main goal of this part of the work is to create a simple, first order model, capable of estimating quickly the device temperature for a wide variety of structures with different geometries and parameters, and that will also be able to capture the fundamental nature of heat flow in the structure.

5.2 Temperature Distribution from Numerical Simulations

In this section, the thermal analysis for a multi-finger HBT device with metal interconnection structure is performed using the numerical simulator Atar. A brief review of theoretical concepts used in this analyse is presented; the heat flux distribution inside the structure is calculated for different planes, and the effect of materials and geometry in maximum device T_{rise} is studied.

5.2.1 Theoretical Background

A detailed analysis of theoretical concepts used in this chapter can be found in [15], [30], and [41], while a brief overview is presented here. As stated in section 2.2.2, according to Fourier's law, the Eq. (2.1) represents the derivative of temperature with respect to coordinate axis x , which is the heat conduction for a one-dimensional system.

Since the material of plate in Fig. 2.5 was assumed homogeneous, then k_{th} would

represent the thermal conductivity over the thickness L of the plate. The quantity of heat transferred per unit time across a unit area is called *heat flux* q'' . The unit of heat flux is watts per square meter (W/m^2) in the SI system. Eq. (2.1) now can be written in terms of heat flux as

$$q'' = \frac{q}{A} = -k_{th} \frac{dT}{dx} \quad (5.1)$$

Eq. (5.1) can be interpreted as stating that if there is a negative temperature gradient dT/dx at a location x , then there will be a heat flow in the positive x direction across the isothermal surface at the same location, and the magnitude of the rate of heat flow per unit area of the isothermal surface is given by Eq. (5.1). Heat flow on the positive x direction is considered positive; therefore, the minus sign in Eq. (5.1) is necessary to meet the requirement that heat must flow from a higher to a lower temperature.

The foregoing discussions can now be extended to a medium with a two-or three-dimensional temperature distribution. Fig. 5.3 shows a set of isothermal surfaces in a solid body, each differing in temperature by a small amount ΔT . The heat flux across the isothermal surface at point P can be readily expressed as

$$q'' = -k_{th} \frac{\partial T}{\partial n} \quad (5.2)$$

where $\partial/\partial n$ represents the differentiation along the normal to the isothermal surface, which is characterized by the unit direction vector $\mathbf{n}$ pointing in the direction of decreasing temperature. Eq. (5.2) can also be written in the rectangular coordinate system as

$$q'' = -k_{th} \left(\frac{\partial T}{\partial x} \frac{\partial x}{\partial n} + \frac{\partial T}{\partial y} \frac{\partial y}{\partial n} + \frac{\partial T}{\partial z} \frac{\partial z}{\partial n} \right) \quad (5.3a)$$

or in the form of

$$q'' = -k_{th} \left(\frac{\partial T}{\partial x} \cos \alpha + \frac{\partial T}{\partial y} \cos \beta + \frac{\partial T}{\partial z} \cos \gamma \right) \quad (5.3b)$$

where (α, β, γ) are the direction cosines of the unit vector $\mathbf{n}$, that is

$$\hat{n} = \hat{i} \cos \alpha + \hat{j} \cos \beta + \hat{k} \cos \gamma \quad (5.4)$$

Figure 5.3 Isothermal surfaces and heat flow paths in a solid body

Therefore, Eq. (5.3a) can be rewritten as

$$q'' = -k_{th} \left(\hat{i} \frac{\partial T}{\partial x} + \hat{j} \frac{\partial T}{\partial y} + \hat{k} \frac{\partial T}{\partial z} \right) \cdot \hat{n} \quad (5.5a)$$

or in a short form

$$q'' = -k_{th} \nabla T \cdot \hat{n} \quad (5.5b)$$

where

$$\nabla T = \hat{i} \frac{\partial T}{\partial x} + \hat{j} \frac{\partial T}{\partial y} + \hat{k} \frac{\partial T}{\partial z} \quad (5.6)$$

is the *gradient of temperature* distribution at point P . This vector is normal to the isother-

mal surface passing through P and pointing in the direction of increasing temperature as shown in Fig. 5.3. Based on Eq. (5.5b) the *heat flux vector* normal to the isothermal surface at P and pointing in the direction of decreasing the temperature can be defined by

$$\mathbf{q}'' = \hat{n} q''_n \quad (5.7)$$

then it can be shown from Eq. (5.5b) that

$$\overline{\mathbf{q}''} = k_{th} \nabla T \quad (5.8)$$

which is the vector form of Fourier's law.

The magnitude of the heat flux across an arbitrary surface passing through P and having the unit direction vector s as its normal, will be equal to the magnitude of the component q'' in s direction. In the rectangular coordinate system, the three components of the heat flux vector q'' are given by

$$q''_x = -k_{th} \frac{\partial T}{\partial x} \quad q''_y = -k_{th} \frac{\partial T}{\partial y} \quad \text{and} \quad q''_z = -k_{th} \frac{\partial T}{\partial z} \quad (5.9)$$

which are the magnitudes of the heat fluxes at P across the surfaces perpendicular to the directions x , y , and z respectively.

These partial differential equations can be solved using analytical or numerical methods. The vast majority of problems encountered in practice, however, can not be solved analytically as they usually involve irregular geometries with mathematically inconvenient boundary conditions. In such cases, numerical methods often provide the answer. For example, if the discrete approximation of heat flux for a given structure is required, using the numerical methods, first the structure model should be discretized, the nodal temperature should be calculated, and the derivatives at each grid point should be

approximated. To illustrate this, the finite-difference formulation of a steady state problem in rectangular coordinates is considered. A solid body with a two-dimensional temperature distribution in rectangular coordinates is shown in Fig. 5.4. Thermal conductivity of the solid is assumed to be constant and there are no internal energy sources or sinks. Under steady-state conditions, the temperature distribution satisfies Laplace's equation

$$\frac{\partial^2 T}{\partial x^2} + \frac{\partial^2 T}{\partial y^2} = 0 \quad (5.10)$$

which gives a continuous temperature distribution within the solid. When these differential equations are expressed in a finite-difference form, the resulting system of equations are satisfied only at certain points.

Figure 5.4 Network for numerical analysis of 2-D heat conduction [41]
a) grid points b) System around grid point for derivation of finite-difference equation.

Figure 5.4a shows the solid body in both x and y directions, forming a nodal network. Eq. (5.10) could be approximated at each grid point by replacing the derivatives with their approximate finite-difference equivalents providing a simultaneous algebraic equation for each grid point.

The Fourier law of heat conduction for unit depth can be approximated as

$$q''_{m-1,n} \equiv -k_{th}\Delta y \frac{T_{m-1,n} - T_{m,n}}{\Delta x} \quad \text{and} \quad q''_{m+1,n} \equiv -k_{th}\Delta y \frac{T_{m-1,n} - T_{m,n}}{\Delta x} \quad (5.11a)$$

Similarly, for other fluxes

$$q''_{m,n-1} \equiv -k_{th}\Delta x \frac{T_{m,n-1} - T_{m,n}}{\Delta y} \quad \text{and} \quad q''_{m,n+1} \equiv -k_{th}\Delta x \frac{T_{m,n-1} - T_{m,n}}{\Delta y} \quad (5.11b)$$

5.2.2 Structure Description-General

A three-finger HBT power device with a metal interconnection structure, as shown in Fig. 5.2, was used to model the device and simulate using the numerical simulator Atar. The structure can be easily described considering that the entire device is composed of two parts: a device without a metal structure referred as *device*, and the metal interconnection structure referred as *metal structure*. This convention is used throughout this material unless otherwise noted. Since the simulation structure and the assumptions made for the device are identical with the structure and assumptions described in the previous chapter, (see section 4.5), they are not repeated here.

Metal structure includes the emitter contacts and the metal layer that connects the contacts together. As shown in Fig. 5.2, emitter contacts are located on the substrate surface above the thermal generation region (see section 4.2.1). Contact dimensions would correspond to the width and length of the emitter stripe. Interconnection layer dimensions would correspond to width and length of the finger array and depend on width and length of the emitter stripe as well as on the spacing between fingers. Thickness for both contacts and interconnection layer is related with the technology and can be set independently in

the *technology file*. For structures simulated throughout this work, metal thickness was varied from 1 to 15 μm which encompasses modern technology limits. Two types of materials were considered for the metal structure: gold and aluminum with respective thermal conductivities 3.17 and 2.37 W/cmK [36].

The device symmetry with respect to x and y axis, combined with the method of images explained in section 4.2.1, is extended and used for the metal structure as well. The structure for the numerical model, after scaling it down by using the symmetry around x and y axes is shown in Fig. 5.5.

Figure 5.5 Power HBT with metal interconnection numerical model

a) layout b) simulated section, and c) Atar model

Boundary conditions on the bottom surface of the substrate represent the chip/package interface. For the work presented here, the simple fixed temperature condition equal to heat sink temperature (ambient temperature T_A) is used. The remaining boundary condition to be defined is for the top substrate surface (where no contact is present), top metal structure surface, and all side surfaces for the entire structure. A zero flow condition across these surfaces was assumed.

5.2.3 Heat Flux and Isothermal Contours

The structure simulated in this case has the following parameters: three-finger emitter; emitter length $L=10\text{ }\mu\text{m}$; fingers spaced at $S=4.16\text{ }\mu\text{m}$; substrate material GaAs; interconnection structure material Au; emitter contact thickness $1\text{ }\mu\text{m}$, and metal layer thickness $1\text{ }\mu\text{m}$.

Substrate surface x-y plane for $z=0$ was considered, and the temperature distribution was calculated for an area that corresponds to half-length of the finger in y-direction ($y=0:5$) and one finger and a half in x-direction ($x=0:4.42$). A mesh was created and the temperature at each point was printed in a output file, resulting in the two dimensional temperature distribution at $z=0$. The x and y components of temperature gradient at each point were calculated using Eq. (5.6) for the whole surface. Using the temperature gradient distribution and Eq. (5.8), the heat flux vector was calculated at each point for the whole area and plotted in Fig. 5.6. Each arrow represents the flux vector at a corresponding point. Arrow length is linearly scaled and represents the magnitude, and the arrow direction represents flux direction at this point.

Figure 5.6 Heat flux and isothermal contours x-y plane for $z=0 \mu\text{m}$.

As it can be seen, flux vectors are normal to the isothermal contours and directed away from heat sources toward the center line between fingers. This flux represents the coupling effect, and the result is higher temperature for the center finger. It is interesting that flux vectors along the center line between fingers are small. This can be explained by the fact that x-y components at these points are small, although this is not the case for the z component which is normal to the x-y plane, directed toward the heat sink, and not shown in this plot. The isothermal contours are more dense near the power dissipation area (at the edge of fingers) which is related to bigger values of temperature gradient and, hence, bigger values for the magnitude of heat flux. Flux direction at the edge of the array shows that the effective spreading area is larger than the finger array area. A vertical plane

y-z can better show the spreading area, and in relation to this, the spreading angle as well.

Following the same procedure as per previous paragraph, the heat flux vectors are plotted (Fig. 5.7) in y-z plane for $x=0$, covering two materials, which are GaAs for substrate and gold for metal structure with different thermal conductivities. Plotting area is a vertical plane with $z = -2:2 \mu\text{m}$ and $y=0:5 \mu\text{m}$. In this case, y and z components of the temperature gradient were calculated at each point and flux vectors were computed.

Figure 5.7 Heat flux and isothermal contours y-z plane for $z=-2$ to $2 \mu\text{m}$.

Clearly, two different areas can be distinguished: inside the substrate where heat flux is directed toward the heat sink, and the metal structure where heat moves up from the hottest point which is the center point of the emitter finger, in the metal, and from there

changing the direction pointing down at the coolest point of the metal structure which is the edge of finger, and toward the heat sink. Even though the amount of heat redistributed through the metal is smaller than heat flowing through the substrate, the vector lengths indicate that heat flow is non negligible, but plays an important role in maximum temperature reduction. The same conclusion can be reached if isothermal contours are analyzed instead. Isothermal contours are more dense inside the substrate, which means temperature gradient dT/dx in Eq. (5-1) is higher. Another aspect that could lead in some interesting results is the direction of flux vectors. Inside the substrate in the center of the finger, flux vectors are perpendicular with x-y plane. Moving away from the center in y-direction, the vectors begin to change the direction, and the angle is different from 90^0 , reaching the range of 45^0 at the finger edge. This value represents the spreading angle and agrees with the value recommended in literature [36] for similar conditions. The spreading angle is related to the effective spreading area, thus for a given spreading angle the ratio between array of finger area and effective spreading area depends on geometry (finger length and array width) and plays a significant role in maximum temperature reduction. The amount of heat flowing from the center to the end of the finger represents the effect of temperature redistribution along the finger that takes place upon the use of metal contact.

To create a better idea of the proportion of heat flowing through the metal, a closer look inside the metal structure itself is required. The intention is to find out the heat flux configuration and distribution in different sublayers inside the metal. For this purpose, temperature profiles along emitter finger for different values of z in metal are shown in Fig. 5.8.

Figure 5.8 Temperature profiles along emitter finger for $z=0:1 \mu\text{m}$.

The temperature difference center-end of the finger has the maximum value for $z=0$ which corresponds to contact surface between substrate and metal. Increasing z , the distance from heat generation region is increased as well, while the temperature difference is decreased considerably. Due to the relation between heat flux and temperature gradient, the amount of heat moving toward the end of finger decreases too. This decrease suggests that the lowermost part of the metal strongly affects heat redistribution. In other words, the contribution of different 'sublayers' of metal in temperature redistribution is different and depends on its distance from the heat source. In addition, because of the lower significance of the effect of upper 'sublayers' on heat redistribution, increasing the metal layer thickness beyond a certain value will have no practical effect on temperature reduction.

Again, heat flux and isothermal contours are plotted in y - z plane (Fig. 5.9), where z was varied in the interval -0.05 to $2\text{ }\mu\text{m}$ (heat sources are located at $z=-0.05\text{ }\mu\text{m}$). Isothermal contours inside the substrate are very dense which can be explained by the high thermal resistance of substrate material. The direction of heat inside the metal structure can be observed.

Figure 5.9 Heat flux and isothermal contours y - z plane for $z=-0.05:2\text{ }\mu\text{m}$.

Since the vectors in substrate area are much larger than vectors in metal, it is difficult to estimate the magnitude of flux per each sublayer in metal. For a better estimation, z is changed inside the metal only, zooming the area of interest. Figure 5.10 illustrates the heat flux and isothermal contours calculated with the same set of parameters assumed for Fig. 5.9, except that $z=0:2$ was considered. In this case, z covers metal contact ($1\text{ }\mu\text{m}$) and metal layer ($1\text{ }\mu\text{m}$). The portion of flux passing through the upper sublayers is much

smaller compared to the bottom sublayers. This plot suggests that the metal thickness of 2 μm seems to be effective, but increasing it above this value will decrease its effectiveness.

Figure 5.10 Heat flux and isothermal contours y-z plane for $z=0:2 \mu\text{m}$.

For a complete view, the heat flux and isothermal contours in x-y planes for different values of z in metal is calculated. In a 1 μm thick metal interconnection layer, the bottom $z=1.1 \mu\text{m}$ and top $z=1.9 \mu\text{m}$ planes are analyzed. The same structure is simulated and the calculations are similar with what was shown in the previous paragraph. Figure 5.11 represents x-y plane for $z=1.1 \mu\text{m}$. This is the bottom 'sublayer' of the metal that connects fingers. The center finger is located at $x=0$, and the next one at $x=4.16 \mu\text{m}$. The heat flux is flowing from the point with highest temperature $x=0, y=0$ following mainly two directions to the points with lowest temperature. For the area shown here, there are two points to

which the flux is directed: one at $x=4.16 \mu\text{m}$, $y=5 \mu\text{m}$ which has the lowest temperature and the flux is flowing diagonally, following the shortest path; and the next one located at $x=0$, $y=5 \mu\text{m}$ and flux flowing in the y -direction along the finger. The diagonal path could be considered as the main 'stream'. The amount of flux following that direction seems to be greater than in the previous one. As a result, the effect of temperature redistribution that takes place when metal interconnection is used could be better justified.

Figure 5.11 Heat flux and isothermal contours x - y plane for $z=1.1 \mu\text{m}$.

The portion of heat flux flowing through upper sublayers in metal is smaller compared with the bottom one. To support this idea, another 'sublayer' close to the top surface of the metal (at $z=1.9 \mu\text{m}$) is picked and heat flux and isothermal contours (Fig. 5.12a), as well as Atar temperature contours (Fig. 5.12b) are shown. Since the heat flux vectors are

not scaled in both plots (Fig. 5.11 and Fig. 5.12a), their magnitude can be compared. The difference noticed is that the amount of heat flowing in bottom part is greater than the heat flowing in the upper part.

**Figure 5.12 Isothermal contours x-y plane for $z=1.9 \mu\text{m}$.
a) Heat flux and b) temperature contours captured from Atar simulation**

5.2.4 The Effect of Materials and Geometry on Maximum T_{rise}

Based on the qualitative results from the previous section, it can be concluded that the device maximum temperature is reduced when the metal layer is used. In addition, having a metal layer thicker than a certain value did not affect the maximum temperature any longer. In order to obtain a quantitative analysis, different structures were modeled and simulated, and the effect of some of the device parameters on maximum temperature was found. Five parameters were considered and varied: substrate material, emitter length, number of fingers, metal thickness, and spacing between fingers. As the first step, the contribution of metal on heat spreading in y direction was simulated using the model for one finger structure. The heat flux was expected to flow from the center toward the edge of the finger due to temperature difference that exists along the emitter finger.

As substrate material, Si, InP, and GaAs were used. Based on several observations in this chapter, where the amount of flux at $z=2\text{ }\mu\text{m}$ was negligible compared with bottom sublayers, metal contact thickness was varied from 0 (no metal) to $3\text{ }\mu\text{m}$. The emitter length was varied $L=10$ to $40\text{ }\mu\text{m}$. Maximum T_{rise} for each structure simulated were normalized to no metal corresponding maximum T_{rise} and plotted in Fig. 5.13.

The first simulation results show that the heat flux flowing in y-direction plays an important role in maximum device temperature. Basically, this is the same conclusion reached qualitatively when heat flux and isothermal contours were plotted in y-z plane while x-component of temperature gradient was not taken into account. The model simulated here creates the same conditions; the heat flow in x-direction was excluded because

the device contains only one finger. This creates the conditions for separating the contribution of y-z components from x-component and investigating it independently. Simulation results (Fig. 5.13) can be used to analyse quantitatively maximum device temperature as a function of parameters mentioned when emitter contact is used.

Figure 5.13 The effect of metal contact on temperature reduction for single emitter structure: a) Si substrate b) InP substrate, c) GaAs.

The relation between finger length and device maximum temperature shows that metal is more effective when used in short finger devices. Generally speaking, there are two factors affecting this relation: temperature difference center-edge of finger, and heat spreading angle. These factors contribute to opposite directions which means that by increasing the length, the difference of temperature center-edge of finger will be increased allowing more heat to flow toward the edge of the finger; and in opposite with that the corner contribution on heat spreading in substrate will be decreased, hence increasing the substrate thermal resistance (relatively speaking).

The relation between substrate material and maximum device temperature shows that metal layer is more effective for devices with low thermal conductivity substrate material. Low thermal conductivity in substrate is related to higher maximum temperature and with a higher difference of temperatures (absolute value) between center and edge of finger, thereby creating the conditions for more heat to flow through the metal which, in this case, is an alternative path. This in general, will improve heat redistribution and lead to lower maximum temperature. From simulation results, for the example of GaAs, the temperature reduction is 6% which is almost twice the temperature reduction of Si (3.2%). The same proportion stands for other structures too.

Eventhough increasing the metal thickness leads to lower thermal resistance and, hence, lower maximum temperature, simulation results repeatedly show that increasing thickness more than 2 μm will have a slight effect on maximum temperature value. For example, referring to the same structure (GaAs-10 μm), increasing thickness from 0 to 1 μm , temperature is decreased by 4.8%, for thickness 1 to 2 μm by 1%, and for 2 to 3 μm

by 0.2% only. The same behaviour can be noticed for other structures but with different magnitude.

In order to enhance the quantitative analysis, a model with three fingers was simulated. Again, metal contacts posted on top of the emitter area were not connected with each other, meaning that the heat flow in x direction in the metal structure was not included. Coupling between fingers is realized through the substrate only. In Fig. 5.14, the cross sectional view of this structure, Atar model, and Atar temperature contours are shown. Finger length is varied from 10 to 40 μm , and metal contact thickness is varied from 0 to 5 μm .

Figure 5.14 Structure model, device with metal contacts only
a) cross sectional view, b) Atar model, and c) Simulated temperature contours.

Simulation results are normalized to maximum T_{rise} found for a structure without metal and plotted in Fig. 5.15. Maximum T_{rise} is similar with the results shown in Fig. 5.13c. As observed, metal thickness beyond 3 μm has a negligible effect on maximum temperature.

Figure 5.15 Normalized temperature vs. contact thickness.

By including the metal interconnection layer, the structure will be similar to the one in Fig. 5.2. Metal contact thickness was set to 1 μm , and interconnection layer thickness was varied from 0 (contacts only) to 5 μm , L was set to 20 μm , and spacing to $L/4$. Three different intervals can be distinguished (Fig. 5.16): for the first interval, a structure with contacts only, thickness 0-1 μm ; for the second interval, interconnection layer is included, thickness 1-2 μm ; and for the third interval, metal layer thickness is increased 2-5 μm . A discrepancy point can be noticed which corresponds to the interfinger metal layer

insertion. The slope for the first part is quite different from the second part, meaning that temperature decreases much faster when the interfinger metal is present. The third part shows the same tendency observed before: increasing metal thickness more than 2 μm has no practical effect on temperature reduction. For a comparison, one of the plots of Fig. 5.15 ($L=20\text{ }\mu\text{m}$) is shown again in Fig. 5.16. The metal interconnection layer significantly reduces maximum device temperature reaching 91.5% (compared with 96% in structure without interconnection). This reduction was expected because of the introduction of a third component in temperature gradient. Now heat can also flow in between fingers, redistributing the temperature in x-direction as described in Fig. 5.11, 5.12.

Figure 5.16 The effect of metal interconnection layer on maximum temperature.

The relation between spacing and maximum temperature was investigated by simulating a 3-finger emitter with GaAs substrate and $L=20\text{ }\mu\text{m}$. Spacings between fingers were varied from $L/4$ to L . Metal layer thickness was varied in a wider range from 0 (no metal) to $16\text{ }\mu\text{m}$. Maximum T_{rise} was normalized to temperature without metal, and plotted in Fig. 5.17, showing that the metal is more effective for structures with small space between fingers. This is related to the coupling effect which is stronger when fingers are placed close to each other, with metal thermal resistance, and also with substrate thermal resistance. To illustrate, when spacing between fingers is small, maximum device temperature is higher and the temperature difference center-edge of finger is higher hence more heat flows through the metal. Also, metal thermal resistance is smaller due to the shorter distance in x-direction in the interconnection layer.

Figure 5.17 Normalized temperature vs. metal thickness for different spacings.

5.3 Structure Representation in Circuit Simulation

The requirement for this section is to create a simple model for a multifinger emitter HBT device with a metal interconnection structure. Although Atar gives a very accurate result, a long time is required to simulate devices with a metal interconnection structure, particularly with a high number of fingers, or long fingers located far from each other. This new approximate model will predict the maximum device temperature when the metal interconnection structure is used, and will also allow for a quick estimation of the effect of metal for design and modeling purposes. Related to this, the main concerns in the new approximate model are considered simulation time and result accuracy.

In a first attempt, a lumped model is developed and the results are shown. After the conclusions are drawn from these results and based on observations made so far, a new and more accurate model is developed. Following the first results, the new model is simplified and used to identify the relations between device maximum temperature and some design parameters as number of fingers, spacing, material thermal conductivity etc. Results from the new model are compared with results obtained from Atar for similar devices.

5.3.1 Lumped Model

The lumped model represents one of the simplest models that can be used to predict the effect of interfinger metal on maximum T_{rise} in a multifinger bipolar transistor. The model is created based on the analogy between thermal and electrical circuits [36] and Eq. (2.7). The current lumped models use a simple resistor network between single R_{th}

models. Electrically speaking, the model performance is assumed to be determined primarily by the maximum finger temperature, while ignoring temperature distribution by using only lumped representation. In other words, thermal resistance calculated for each heat source reflects finger temperature distribution.

To create the model, the entire device with a metal interconnection structure is partitioned into two parts as mentioned in section 5.2.2: the device, and the metal structure. For each part, the model is created and the sub-network is extracted. Both sub-circuits are connected together and the circuit is solved using an external simulator.

As an example, a three-finger device with $L=20\ \mu\text{m}$ and $S=5\ \mu\text{m}$ with metal structure is considered. The device lumped model is created using a simple circuit for each source which contains an independent current source and R_{th} , as shown in Fig. 5.18.

Figure 5.18 Lumped model representation
a) device cross section view b) electric circuit extracted.

The value of current in current source is equal to the value of power dissipated in this source, and is calculated using the power dissipation given per unit area. R_{th} is calculated using Eq. (2.6) in the form of $R_{thi} = (T_{maxi} - T_A)/P_i$, where T_{maxi} (K) represents maximum finger temperature and can be found using analytical solver tsource, and P_i is

total power dissipated in this finger. From the model created, thermal subnetwork is extracted, representing the device

The metal structure is modeled using only four different blocks: emitter contacts, corners, interfinger, and center as shown in Fig. 5.19. The thermal resistance between two adjacent blocks is calculated using the common cross sectional area, the size of two blocks, and the local thermal conductivity. Note that this method is explained in more detail in section 5.3.6 later in this chapter. The thermal sub-network for the metal structure is extracted.

Figure 5.19 Lumped model for metal structure.

Two sub-networks are connected creating the equivalent thermal network shown in Fig. 5.20, where R_{ths} , R_{thc} , R_{thA} , R_{thI} , and R_{thc} are thermal resistance for substrate, contact block, angle block, interfinger block, and center block respectively. Nodes 1, 2, and 3 represent the finger temperature points.

Figure 5.20 Thermal network extracted from lumped model.

This equivalent circuit is solved and the maximum temperature for each finger is shown in Table 5. Temperatures obtained using Atar, for the same nodes, are also shown. It can be noticed from the table that the agreement between two models is not satisfactory. The direction of predicted temperature for two outermost fingers is also incorrect. Furthermore, maximum temperature is wrongly increased instead of decreasing. Finally, the predicted temperature for the middle finger is far from the expected Atar value.

Table 5: Maximum finger temperature predicted with lumped model

Node #	Atar no metal T (K)	Atar with metal		Lumped model with metal	
		T (K)	$\Delta T(\%)$	T(K)	$\Delta T(\%)$
1	47.92	44.69	-6.7	48.84	1.9
2	51.64	46.73	-9.5	49.66	-3.8
3	47.92	44.69	-6.7	48.84	1.9

Such poor results are related to the fact that lumped representation can not reflect the temperature distribution, or, in other words the temperature distribution is ignored and,

hence, the heat flowing in y direction, as shown in Fig. 5.10, and also in the diagonal direction from the hottest (center of middle finger) to the coolest point (edge of outermost finger) shown in Fig. 5.11 are also ignored. To support this idea, Atar simulation results are plotted in Fig. 5.21, where the temperature profiles along x direction for $y=0$ and $y=5$ μm , for a device without and with metal are illustrated. When a metal structure is used, the maximum finger temperature in the center of the finger is decreased for three fingers, but the decrement is larger for the middle one, creating a more uniform distribution between fingers and reducing the maximum junction temperature. In contrast, at the edge of the finger the temperatures increase because of the heat flowing from the center toward the edges. The increment for the middle finger is slightly smaller than that for the outer ones, which also creates a more uniform temperature distribution in the finger array area.

Figure 5.21 Temperature profiles across the fingers for structure with and without metal.

The same conclusion can be reached by looking into the temperature distribution along emitter finger in y direction at $x=0$ and $x=5\text{ }\mu\text{m}$ for a device without and with a metal structure (Fig. 5.22). It shows that when the metal structure is included, temperature distribution along the finger is more uniform, thus improving the device performance.

Figure 5.22 Temperature profiles along the fingers for structure with and without metal.

Due to the lack of such details when the lumped model is used, the conclusion reached is that the lumped representation and, hence, the lumped model can not represent thermal behaviour and can not be used to predict the maximum T_{rise} in a multifinger HBT with interfinger metal structure. This conclusion reinforces the requirement of a more accurate model which is developed in the following sections of this chapter.

5.3.2 Model Development

According to the flow chart of model developed (Fig. 5.23), the entire structure is divided into two parts: device and metal structure, and each of these is modeled in a different way. The analytical method is used to model the device, and a simplified numerical one is used to model the metal structure. From the models created, using the analogy with electrical circuit, two different thermal networks are extracted. Connecting these two sub-circuits properly, a thermal network for the whole device is created. This circuit is simulated using an external circuit simulator. Once the circuit is solved, the temperature at each node for the whole structure can be found. In the following paragraphs, each block of the flow chart is explained in more detail.

Figure 5.23 Model development flow chart.

5.3.3 Input Data

Parameters included in data input file are variables that can be changed in order to obtain different device configurations. This file contains information about geometry and material properties for the device and metal interconnection structure. The substrate thickness was specified without a lateral dimension as per analytical model described in section 4.2.1. Finger dimensions were specified with width along the x-axis and length along y-axis. Finger width and length were considered as design variables, although different values could be given. Generation area can be specified with a single thermal source or with an array of sources where the number of sources is specified in both x and y directions. Source location was specified in x-y plane with a center array coordinate, and in z direction as depth from the substrate surface. Spacing between sources was expressed in terms of finger length L . Power dissipated P ($\text{mW}/\mu\text{m}^2$) at these sources was considered uniformly distributed in the emitter area. Boundary conditions were defined based on assumptions made in section 5.2.2. Material properties were specified with thermal conductivity K (T_{backside}) (W/cmK), density ρ (g/cm^3), and specific heat c [J/gK]. According to [28] ρ and c do not influence the steady state temperature if integration time is long enough ($t \rightarrow \infty$).

5.3.4 Device Model

After defining the parameters in the input file, t_{source} is used to model the device. Heat sources are discretized creating a set of elementary sections for each source. Temperature inside an elementary section is assumed to be constant and equal to the temperature

at the center of the section. A set of points, representing elementary sections where the temperature should be known, is defined and then the temperature at each given point is calculated. Using power dissipation and temperature at each point, thermal resistance for each elementary section can be found. Based on the assumption that the thermal conductivity is constant, thermal resistors are also constant. Thermal capacitance between blocks was not considered because the steady-state device temperature was calculated. Based on these considerations, the substrate thermal network is built. This network consists of constant thermal resistors and independent current sources.

The first step toward creating the device model is heat source discretization. The substrate surface x - y plane where the heat sources are located is considered an area of interest. Since temperature distribution across the heat source is shown to be relatively constant, its effect on heat flow is neglected, and the temperature across the emitter width is considered uniform and equal to the temperature in the center of the emitter $x_0(0,0)$ as shown in Fig. 5.24. From Fig. 5.8, temperature profile along finger (y -direction) is shown to be not uniform, therefore the difference in temperature center-edge of the heat source can not be neglected.

Figure 5.24 Schematic of single emitter structure with metal.

To model the temperature profile along y-axis, temperature is calculated at different points, and a discretized profile is obtained. In other words, the source is divided into elementary sections and the temperature at each section is considered equal to the temperature at the center of the section. In Fig. 5.24 an example of three sections per source is shown where (x_0, y_0) , (x_0, y_1) , and (x_0, y_2) are center points, and T_{y0} , T_{y1} , and T_{y2} corresponding temperatures. These temperatures are plotted in Fig. 5.25 in the same axis with a continuous temperature profile obtained for the same device. For the center block, approximation is acceptable; however, this is not the case for two other blocks where the temperature difference between block edges is much larger and approximation is far from a real profile. In conclusion, the accuracy of temperature representation depends on the number of sections per source or, in other words, section size.

Figure 5.25 Temperature profile along emitter finger.

Due to the steep temperature profile at the edge of the finger (blocks y_1 and y_2), and almost flat for center section y_0 , the section size near edge must be smaller without decreasing the size of the center section. By creating more independent blocks, a better approximation to the real profile will be provided. In the example shown, for a middle section with length equal to one third of the total emitter length, the temperature drops only 5% of the total temperature drop, unlike other sections. The strategy adapted in the program is as follows: starting with three sections, three equal sized sections with $L_i=3.33 \mu\text{m}$ for total emitter length $L=10 \mu\text{m}$ are created, increasing the number of sections, according to the technique applied, means dividing the outermost sections in two equal sized sections (center section kept in original length). As an example, for five sections, the section length will be 1.65, 1.65, 3.33, 1.65, 1.65 μm as shown in Fig. 5.26 and so on, creating a set of points in y-direction at x_0 . Finally, for a single finger device the set will contain one vector with n points in y-direction at $x=x_0$. In a multi-finger device, for each finger a set of n points will be created in y-direction at $[x_0, x_1, \dots, x_i, \dots, x_m]$ where x_i is a vector with n elements corresponding to i -th finger in x axis, and m is the total number of fingers. The distance $x_i x_{i+1}$ represents spacing between fingers. The set in this case will be a matrix with dimensions $[m, n]$.

From another viewpoint, knowing the section length (width is also known), power dissipation per each section is calculated based on power dissipation per unit of area and the area of each section. The set of temperature points and power dissipation per each section were included in the input file for external simulator tsource, and the structure was

simulated from inside the main program. Simulation results from the output tsource file were fed back into the main program, providing the value of temperature at each point specified in the input file.

Figure 5.26 Sectioning the heat source.

As mentioned above, the number of sections per finger plays an important role in simulated temperature profile; increasing the number of sections means better approximation of temperature profile, hence more accurate results, but larger computation time. To demonstrate this, approximated temperature profile for 17 sections per finger and continuous one are plotted in Fig. 5.27 where a slight difference can be observed. Note that the method used for source sectioning requires an odd number of sections per finger, which introduces a constraint.

Figure 5.27 Temperature profile approximation for 17 sections/finger.

5.3.5 Metal Structure Model

The basic idea behind the metal structure model is to create a complex mesh geometry using rectangular blocks of different size. Fig. 5.28 shows a simplified model of a metal interconnection for a three-finger emitter device with finger length L , finger width W , and spacing between fingers S . To define the length of each block in y-direction, the same idea as per source sectioning is applied. Five sections per finger and two different block widths (in x-direction) were considered for this example. Blocks located above the finger area, including metal contacts and sections of interconnection layer, have the same width as the finger does (W). Blocks located in the “between fingers” area have width equal with the spacing between fingers, which is the distance edge to edge shown as $S-W$

in Fig. 5.28. Such large blocks between fingers are chosen because the gradient of temperature in the metal layer is small, and two resistors per block in x-direction are sufficient to cover the temperature difference in this area. This can be justified by looking at isothermal contours in Fig. 5.11 and 5.12 where the temperature difference between two furthestmost points in the metal layer is 1.3 K and 0.9 K for the bottom and top sublayers in metal respectively.

Figure 5.28 Simplified model for metal structure.

5.3.6 Extraction of the Thermal Network

Following the creation of the models for the device and the metal structure, the next step is to extract the thermal network and create the equivalent electrical circuit. Thermal resistance R_{thi} per each section of generation region is found using corresponding power dissipation, temperature rise, and Eq. (2.6). Because the device model contains $[xnum, ypoints]$ elementary sections with each one of them having its own power dissi-

pated P_i and thermal resistance R_{thi} calculated using Eq. (2.6), the equivalent electrical sub-network for the device owns the same number of elementary circuits. Each of these circuits has an independent current source numerically equal to P_i and an electrical resistor numerically equal to R_{thi} , therefore creating the first 'layer' with nodes and resistors connected to ground. An equivalent electrical sub-circuit extracted for the example of three sections per finger is shown in Fig. 5.29.

Figure 5.29 Device model without metal structure and thermal network extracted.

The thermal network for the metal structure model, shown in Fig. 5.28, consists of linear thermal resistors calculated using the common cross sectional area, the size of two blocks, and the local thermal conductivity. Thermal capacitance between blocks once again, is not considered. Using the topology of connections with adjacent blocks, a network of resistors to the neighboring nodes is created. $R_{thj,k}$ can be found using

$$R_{thj,k} = \frac{1}{k_{thj}} \frac{L_{j,k}}{A_j} \quad (5.12)$$

where: $R_{thj,k}$ is the thermal resistance between blocks j and k , k_{thj} the local thermal conductivity, $L_{j,k}$ the length between center j th and k th blocks (m), and $A_{j,k}$ is the cross-sectional area between two adjacent blocks, perpendicular with the heat flux direction.

An example of $R_{thj,k}$ calculation is shown in Fig. 5.30. Using Eq. (5.12), thermal resistances for each block are calculated. The bottom layer with resistors is calculated differently since these resistors connect nodes in two different materials with different thermal conductivities, in this case Au and GaAs. Referring to electrical analogy, when different materials are stacked in series, the equivalent or total thermal resistance becomes the sum of the local thermal resistances, analogous to several resistors in series. Each of these local thermal resistances is related to its local thermal conductivity. Eq. (5.12) can be applied for each section of block separately, and the total thermal resistance for each block in this layer can be found as a sum of two local thermal resistances. Having all the sub-circuit nodes and resistors connected in between, the equivalent electrical sub-circuit for the metal structure is created.

Figure 5.30 Example of thermal resistance calculation.

Connecting both sub-circuits (one for the device and one for the metal structure), the equivalent electrical network for the entire device is extracted. Figure 5.31 shows a simplified circuit including both parts. This circuit could be simulated using an external circuit simulator such as Hspice.

Figure 5.31 Interconnection between two networks.

Although the basic concept involved in metal structure modeling is similar with the concept used in Atar, in the new approximated model the metal structure sub-circuit is simplified and the size is reduced. On the other hand, the network for substrate is extracted in a different way, which is a more efficient technique of substrate model representation. Both of these factors reduce the circuit size and computation time.

Short computation time is achieved at the cost of some limitations in the spectrum of devices covered, and in the result accuracy. These limitations are related to assumptions made in the model. For example, the aspect ratio length/width of an emitter finger is assumed to be at least ten, which does not allow the use of the model without modifica-

tions for lower aspect ratio devices. Another limitation is attributed to limited number of sections per finger, producing coarse discretization of temperature profile along the finger for both center and edge areas. The center section has a length equal to one third of the total finger length, giving an unsmooth approximation for this area. For the edge of the finger, the approximation is also reduced due to the limited number of sections. These two approximations will introduce some errors, especially for long emitter devices, as it will be shown later in this chapter. Another limitation is related to the assumption that the heat sources are considered being located close to the substrate surface, restricting the use of the model for planar or quasi planar devices. If sources are located somewhat deeper into the substrate, the effect of coupling between sources through the upper part of the substrate takes place and the model cannot represent this effect.

5.3.7 Circuit Simulation

After the extraction of the thermal (electrical) network, Hspice netlist was created, and the circuit was simulated and solved. The voltage at each node represents the T_{rise} at this point. The number of elements in this circuit depends on the number of fingers, number of sections per finger, and number of sublayers per metal layer. For example, for a three-finger structure, with 17 sections per finger, and three sublayers per 1 μm of metal layer (or contact), there are 459 nodes, 51 independent current sources, and 1047 resistors. For comparison purposes, the same structure was simulated using Atar, and it was found that the thermal network consists of 16,286 nodes, 480 independent current sources, and 58,298 resistors.

5.3.8 Model Specifications and Simulation Results

To demonstrate the approximate model and to compare the simulation results with results obtained from Atar, examples with three and five emitter fingers were simulated using this new approximated model. Device parameters are identical with parameters presented in section 4.2.1 but are not shown here. In addition, three different emitter lengths ($L=10, 20$, and $40 \mu\text{m}$) were defined, spacing between fingers was set to $L^\alpha \mu\text{m}$, and the array of sources was specified with two dimensions: number of sources in x direction $xnum=3$ (or 5), and number of rows in y direction $ynum=1$ (in total one row with three or five sources). Based on the approximated temperature profile shown in Fig. 5.27, the number of sections per finger was set to $ypoints=17$.

Metal structure was defined using width, length, and thickness in x , y , and z -direction respectively. Metal contacts have the same width and length as heat sources do, thus $W=0.5 \mu\text{m}$ and L and posted above the heat sources on top of substrate surface as shown in Fig. 5.2, and thickness was set to $1 \mu\text{m}$. For the interconnection layer, the width was defined using spacing between fingers and finger width ($9.4 \mu\text{m}$ for three-finger $L=10 \mu\text{m}$ case); length was set equal to L , and thickness to $1 \mu\text{m}$, covering the whole area of finger array and posted on top of the contacts as shown in the model in Fig. 5.28. Gold was used as the metal structure material. In a first attempt, block thickness in metal was chosen to be $0.33 \mu\text{m}$, meaning that $2 \mu\text{m}$ metal was divided into six sublayers, creating six layers with blocks. Other parameters not mentioned here are identical with those in section 5.2.2.

For an evaluation of the approximated model, first, a target error should be specified. According to [35], a two percent error is considered as a target, representing a quarter

of a typical measured thermal resistance uncertainty. Simulation results from Atar and the new approximated model (Table 6) are close, and the accuracy reached is within the specification range, but more than sufficient for a first order estimation model, which suggests the possibility of a further simplification in the approximated model.

Table 6: Simulated maximum T_{rise} approximated model vs. Atar

Number of fingers	Finger length (μm)	T_{rise} max Atar	T_{rise} max New model	Difference (%)
3	10	38.37	38.54	0.4
3	20	44.98	44.76	0.5
3	40	51.86	50.67	2.2
5	10	44.91	44.82	0.2
5	20	52.63	52.21	0.8
5	40	60.07	58.66	2.3

5.4 Model Simplification

In order to simplify the model and reduce the number of elements in the circuit, two different approaches are considered: block thickness in metal structure, which is related to the number of elements in the metal structure sub-circuit, and number of sections per finger which affects the whole circuit size.

Increasing the *sublayer thickness* B_l (see Fig. 5.28) means fewer blocks in the metal structure model and fewer nodes and resistors in the corresponding circuit. In order to achieve this, first the difference of temperature between two adjacent sublayers must be investigated. Investigation is performed for the nodal structure shown in Fig. 5.32, where

the whole metal structure is divided into six sublayers with $B_f=0.33\ \mu\text{m}$, creating six layers with blocks in z direction, and the T_{rise} at each node of the grid is found and plotted on Fig. 5.33. There are six sets of temperatures each of them representing the temperature profile for the corresponding sublayer in the metal along the finger. The middle finger was chosen because of its higher temperatures compared to other fingers.

Figure 5.32 Nodal structure for the middle finger (five sections/finger).

The difference of temperature inside one sublayer (Fig. 5.33) is very small; the maximum value is reached at the bottom sublayer and is in the range of 0.2 K (compare with 0.04 for top sublayer). Between two adjacent sublayers, the temperature difference is also small, and it becomes smaller when moving toward the top metal sublayer (far from source). Even between the two furthestmost sublayers top and bottom, the difference is only 0.1 K. This difference suggests that increasing the sublayer thickness or decreasing the number of sublayers in metal will not compromise the result accuracy. Based on this investigation, the number of sublayers was decreased to four, which corresponds to metal block thickness $B_f=0.5\ \mu\text{m}$.

Figure 5.33 Temperature profiles along emitter finger for different sublayers in metal.

Simulation was run for three and five finger structures, and the results for both approximated models and Atar are compared (Table 7). The difference between two approximated models is small and randomly distributed, and may be considered as a meshing noise. The number of elements in the second approximated model is even smaller. Now the circuit has 323 nodes, 51 independent current sources, and 715 resistors. Another element of interest is the intersection point between curves which represents the temperature contour line perpendicular with x-y plane shown in Fig. 5.10.

Table 7: Simulated maximum T_{rise} comparison between two approximated models

Number of fingers	Finger length (μm)	Atar	New model S1 six sublayers	New model S2 four sublayers	Mag. of Difference (%)	
					S1	S2
3	10	38.37	38.54	38.63	0.45	0.67
3	20	44.98	44.76	44.99	0.47	0.03
3	40	51.86	50.67	50.74	2.29	2.15
5	10	44.91	44.82	44.91	0.18	0.01
5	20	52.63	52.21	52.30	0.79	0.62
5	40	60.07	58.67	58.73	2.35	2.22

The second approach in reducing the circuit size is to reduce the number of sections per finger. If the number of sections per finger is decreased, the model will be simplified, presumably trading the result accuracy and/or consistency. According to section 5.3.4, 17 sections per finger were found to generate a reasonable approximation of the continuous temperature profile along the finger. Based on this, the model was initially built and simulated with 17 sections per finger. To find the relation between maximum T_{rise} in the model and the number of sections per finger, a three-finger structure with finger lengths of $L=10, 20$, and $40 \mu\text{m}$ was simulated. The results are normalized to maximum T_{rise} for 17 sections. This number is chosen because it ended up representing a limit after which the result no longer depends on the number of sections per finger. Normalized results (Fig. 5.34) show that by reducing the number of sections in the range of seven, the uncertainty introduced will be in the range of one percent, which is within the range of the target of error specified. Note that a model with six sections per finger is also within the

specification, but recall that only odd numbers of sections can be used. Finally, the circuit size after simplification (2 sublayers in 1 μm metal layer and 7 sections/finger) is shown in Table 8. For a comparison, the number of elements in an Atar circuit are also listed.

Figure 5.34 Result consistency as function of number of sections per finger.

Table 8: Circuit size-approximated model vs. Atar

Elements	Atar	Approximated model
Nodes	16,286	133
Resistors	58,298	285
Current sources	480	21
Simulation time	15'	1'

5.5 Maximum Temperature Analysis, Approximated Model vs. Atar

The approximated model can be used to investigate the dependencies between maximum device temperature and different design parameters. Calculation time, even when a very large model is simulated, is in the range of a few minutes with Ultraspark 1 workstation. Using Atar for the same device requires an unpractical amount of time.

5.5.1 Interfinger Material Dependence of Metal Coupling Effect

Two different metal structure materials were considered: gold and aluminum with respective thermal conductivity [36]. The model for five fingers, $L=20\text{ }\mu\text{m}$, and $S=20^{0.62}\text{ }\mu\text{m}$ is simulated using both Atar and approximated model s2, and the maximum T_{rise} was found. As shown in Table 9, the maximum temperature is higher when aluminum is used. This is expected because the thermal conductivity of aluminum is lower than gold: in other words, since aluminum is less conductive, the amount of heat flux redistributed through the metal layer will decrease and the maximum temperature will increase. Both models agree at this point. The effect, in general, is not strong because the thermal conductivities for both metals do not differ significantly.

Table 9: Maximum temperatures for different metal structure material

Metal Structure Material	k_{th} (W/cmK)	Atar Max T_{rise} (K)	New model Max T_{rise} (K)
Gold	3.17	52.63	51.96
Aluminum	2.37	53.10	52.29
Temperature Difference		+0.47	+0.33

5.5.2 Substrate Material Dependence of Metal Coupling Effect

The approximated model s2 is used to simulate the structure for different substrate materials. For this demonstration, Si, InP, and GaAs are utilized as substrate materials. Initially, the model for a three-finger emitter, $L=10\text{ }\mu\text{m}$, and $S=10^{0.62}\text{ }\mu\text{m}$ was created and simulated without a metal interconnection using Atar. Simulation was run for the three different substrate materials mentioned above, and the results are listed in Table 10. Once the metal interconnection model was included in the device, the whole model created was simulated using Atar and s2. Maximum device temperature obtained from both models is computed by normalizing to no metal maximum T_{rise} and also shown in Table 10. Results found using s2 and Atar differ in the range of two percent and the new approximated model predicts the effect very well. The conclusions outlined before for a three-finger case ($L=20\text{ }\mu\text{m}$ and different spacings Fig. 5.16 and 5.17), are confirmed here: the effect of metal interconnection is stronger when used in structures with poor thermal conductivity materials like GaAs. Temperature reduction for GaAs (9.2%) is higher compared with Si (7.9%).

Table 10: Maximum temperature rise for three different substrates

Substrate material	k_{th} (W/cmK)	Atar no metal Max T_{rise} (K)	Atar with metal Max T_{rise} (K)	New model with metal Max T_{rise} (K)	Atar Norm T_{rise} (%)	New model Norm T_{rise} (%)
Si	1.5	12.31	11.65	11.34	94.65	92.10
InP	0.69	26.79	24.71	24.50	92.23	91.42
GaAs	0.44	42.22	38.22	38.34	90.50	90.81

5.5.3 Geometry Dependence of Metal Coupling Effect

The effect of geometry in maximum T_{rise} is also investigated. Three design parameters are considered: emitter length, number of fingers, and spacing between fingers.

To find the relation between maximum T_{rise} and L , when metal structure is used, a three-finger structure was simulated, considering two options: spacing between fingers is function of finger length ($S=L^{0.62}$ μm); and spacing is set to 6.4 μm and kept constant. Simulations were performed in models with and without metal interconnection, corresponding maximum T_{rise} were found, and temperature reduction ΔT was computed from

$$\Delta T = \frac{T_1 - T_2}{T_1} 100 [\%] \quad (5.13)$$

where: ΔT - is temperature reduction when metal structure is used (%), T_1 - maximum T_{rise} for device without metal (K), T_2 -maximum T_{rise} for device with metal (K) as shown in Fig. 5.35.

Temperature reduction ΔT (%) as a function of the emitter length for both options mentioned above is plotted in Fig. 5.35. Three factors affect temperature reduction: F1- difference of temperature between center-edge of finger array ; F2- angular spreading contribution; and F3- metal structure thermal resistance. These factors will affect ΔT in different ways for different finger lengths.

Figure 5.35 Temperature reduction vs. L for constant and variable spacing.

When constant spacing is used, in the interval $L=10 \mu m$ to $30 \mu m$, the effect of coupling between fingers is strong. Such an effect will generally increase the device temperature and the difference of temperature between center-edge of finger array (Fig. 5.36), where T_c and T_e are temperatures for center-edge of finger array for a device without metal structure. The increment is very high in this interval. It is expected to generate a higher heat flux flowing through metal interconnection, hence, temperature reduction ΔT will increase. Conversely, because the fingers are short, the second and third factors are weak (ratio is still high and metal thermal resistance is still low), their effect on ΔT is weak. In general, the first factor will govern the ΔT for this interval. The second interval starts at

$L=40\text{ }\mu\text{m}$. The difference of temperature between center-edge of finger array continues to slowly increase in proportion to finger length (Fig. 5.36b). This increase will lead to a weaker effect of temperature difference. At the same time spreading angle and metal thermal resistance begin to play an important role. Increasing the size of the metal structure, thermal resistance in the heat flow path in the metal will increase, reducing the amount of heat redistributed. In general, the second and third factors will manage the ΔT for this interval.

(a)

(b)

Figure 5.36 Temperature difference center-edge vs. L for device without metal structure
a) finger array-top view and b) the dependence.

If spacing between fingers is function of length, the factor of temperature difference center-edge is not strong enough, since length and spacing increase simultaneously. In this case, only the angular spreading and metal structure thermal resistance will be in effect and will govern the ΔT .

The effect of number of fingers on temperature reduction is shown in Fig. 5.37.

The model simulated has finger length $L=20\text{ }\mu\text{m}$, and $S=6.4\text{ }\mu\text{m}$. ΔT was found using Eq. (5.13). The same speculation made in the previous paragraph is effective here. In the interval 1 to 7 fingers, the effect of coupling is high, and the temperature difference center-edge of finger array governs ΔT . In the next interval, starting at 7 fingers, ΔT is dominated from the metal interconnection thermal resistance.

Figure 5.37 Temperature reduction vs. number of fingers for device with metal structure.

The effect of spacing between fingers on temperature reduction is shown in Fig. 5.38. A three-finger emitter device with finger length set to $L=20\text{ }\mu\text{m}$ and different spacings was simulated. Minimum spacing, in terms of unit of spacing (recall section 4.3), was taken $0.75 U_s$ where $U_s=20^{0.62}=6.4\text{ }\mu\text{m}$, and maximum $7U_s$.

Figure 5.38 Temperature reduction vs. spacing for device with metal structure.

Temperature reduction ΔT is calculated using Eq. (5.13). By increasing spacing between fingers, the metal structure becomes less effective, also coupling effect decreases and device temperature in general and specifically $T_c - T_e$ decrease, resulting in less heat flux flowing through the metal. Because the metal interconnection thermal resistance will be higher, and the effective spreading area in the substrate will increase, more heat flux will flow toward the heat sink. All these factors are governing ΔT in this case.

5.6 Summary

In this chapter a new approximated model for a multifinger HBT device with a metal interconnection structure was introduced. Thermal analysis was performed using a numerical simulator. Based on this analysis, a simplified model was created. The model size is very small compared to the numerical one. Using this new model, simulation time can be reduced dramatically, making possible the prediction of maximum T_{rise} in the devices with a large emitter area. To demonstrate this, the effect of different design parameters in maximum T_{rise} when the metal structure is used was investigated. Results showed that the new model predicts the effect adequately, with a prediction error of less than two percent. This model can be used successfully in the thermal design process where many parameters must be varied in order to optimize device thermal performance.

Chapter 6

Conclusions

6.1 Thesis Summary

In this thesis the nature of thermal coupling in a multiemitter HBT device with and without metal interfinger connection was investigated, with the intention of creating new simplified models for maximum temperature prediction. Proposed models can be used successfully in the initial stages of the thermal design process where many simulations should be performed in order to approximately define the appropriate device structure(s), which later, in the design process, can be simulated using more sophisticated numerical models.

For a multifinger device without a metal interconnection, a new normalized approach to model thermal coupling was proposed. This model was used to predict the thermal behavior in a entire ‘family’ of devices with different finger lengths. To demonstrate this idea, a model for a device with emitter length $L=20\text{ }\mu\text{m}$ was utilized to predict the temperature in other devices ($L=10$ and $40\text{ }\mu\text{m}$). Less than a two percent error in temperature prediction suggests that coupling between sources in the normal multi-finger configuration can be handled in a normalized way by expressing the inter-finger spacing as a proportion of the finger length. Using this new approach, the effect of coupling for different devices and its impact on maximum T_{rise} was studied. The new approach shows that T_{rise} depends very strongly on geometrical parameters (number of fingers, finger length,

and spacing) as well as on the substrate thermal conductivity. Operating temperature is significantly higher in a multiple finger device, especially for small spacings. Also, some design scenarios were presented and the trade-off between area and temperature was studied.

For a multifinger HBT device with interfinger metal, a new simplified model was proposed. A simple model based on lumped representation was found inadequate in providing a physical representation of the thermal aspects in the device, therefore a more accurate model (but simpler than a numerical one) was created. Results obtained from the simplified model are compared with numerical ones and a good agreement (within the range of 2%) was found. Since the model size is very small compared to the numerical one, simulation time was reduced dramatically, making possible the prediction of maximum T_{rise} in devices with a large emitter array size. Results show that the design parameters and material properties strongly affect the maximum temperature. Based on the examination of simulation results, some practical conclusions can be observed: metal interfinger layer reduces the maximum device temperature in the range of 10-15%, it works better for short emitters and small spacings, it is more effective for low thermal conductivity substrate materials, like GaAs, and increasing the metal thickness more than 3 μm has no practical effect on temperature reduction.

Temperature reduction in the 10-15% range predicted for the device with metal structure is potentially important, however, such a range does not preclude the use of the model proposed for the device without metal structure (chapter 4), which can still be used to predict the temperature in a initial stage of the design process.

6.2 Future Research

Although the simplified models described did achieve the final goal, a number of issues remain to be examined, most of them related to assumptions made about the models. Throughout this work, thermal conductivity was assumed to be constant (independent from temperature), which is regarded as a reasonable assumption for a certain temperature range. However, in applications where temperature rise is very significant, this effect should be taken into account. Since for different substrate materials, thermal conductivity depends on temperature in different ways, drawing a line where this effect should be counted for is difficult. Based on this point, one direction for the future work becomes evident: to test both models in the presence of dependent thermal conductivity and compare the results with the results shown in this work. Completing this stage will further improve the models. According to some early stage simulations, results are promising, but more is required in order to reach a conclusion about that.

As shown throughout the material presented in this thesis, predicted temperatures have an excellent agreement with numerical simulation, but comparison/calibration to measurements is still open, initiating another direction for future work. In this way, it will be very useful to compare theoretical results with temperatures measured in practical device.

According to the simplified model, the metal structure reduces the maximum device temperature in the range of 10-15 percent. A question raised at this point is how will this reduction affect the circuit performance? The answer to this question leads to another direction for future work.

References

- [1] D. A. Doane and P. D. Franzon, *Multichip Module Technologies and Alternatives-The Basics*, New York, NY: Van Nostrand Reinhold, 1993.
- [2] W. Shockley, U.S. Patent No. 2,569,347, 1951.
- [3] H. Kroemer, "Theory of a wide-gap emitter for transistors," *Proceedings of the IRE*, vol. 45, no.11, pp. 1535-1537, 1957.
- [4] D. K. Jaus and D. L. Feucht, "The realization of GaAs-Ge wide band gap emitter transistor," *IEEE Transaction on Electron Devices*, vol. ED-16, no. 1, pp. 102-107, 1969.
- [5] B. Bayraktaroglu, "GaAs HBT's for microwave integrated circuits", *Proceedings of the IEEE*, vol. 81, no. 12, pp.1762-1785, 1993.
- [6] B. K. Oyama and B. P. Wong, "GaAs HBT's for analog circuits", *Proceedings of the IEEE*, vol. 81, no. 12, pp.1744-1761, 1993.
- [7] B. Bayraktaroglu, "HBT Power Devices and Circuits," *Solid-State Electronics*, vol. 41, no. 10, pp. 1657-1665, 1997.
- [8] N. L. Wang, N. H. Sheng, M. F. Chang, W. J. Ho, G. J. Sullivan, E. A. Sovero, J. A. Higgins, and P. M. Asbeck."Ultrahigh power efficiency operation of common-emitter common-base HBTs at 10 GHz," *IEEE Trans. Microwave Theory & Tech.*, vol 38, pp. 1381-1389, 1990.
- [9] M. F. Chang, ed, *Current Trends in Heterojunction Bipolar Transistors*, Singapore: World Scientific, 1996.
- [10] J. J. Liou, *Principles and Analysis of AlGaAs/GaAs Heterojunction Bipolar Transis-*

- tors. Norwood, MA: Artech House, 1996.
- [11] T. Won and H. Markoc, "Self-aligned $\text{In}_{0.52}\text{Al}_{0.48}\text{As}/\text{In}_{0.53}\text{Ga}_{0.47}\text{As}$ heterojunction bipolar transistors with graded interface on semi-insulating InP grown by MBE," *IEEE Electron Device Lett.*, vol. 10, pp. 138, 1989.
- [12] C. A. King, J. L. Hoyt, and J. F. Gibbons, "Bandgap and transport properties of Si_{1-x}Ge by analysis of nearly ideal $\text{Si}/\text{Si}_{1-x}\text{Ge}/\text{Si}$ heterojunction bipolar transistors," *IEEE Trans. Electron Devices*, vol. 36, pp. 2093-2104, 1989.
- [13] J. R. Houser, "The effect of distributed base potential on emitter-current injection density and effective base resistance for stripe transistor geometries," *IEEE Transactions on Electron Devices*, vol. ED-11, pp. 238-242, 1964.
- [14] F. Ali and A. Gupta, eds, *HEMTs and HBTs: Device, Fabrication, and Circuits*, Norwood, MA: Artech House, 1991.
- [15] M. Necati Ozisik, *Heat Conduction*, 2nd ed., New York, NY: John Willey & Sons, 1993.
- [16] W. Liu, *Handbook of III-V Heterojunction Bipolar Transistors*. New York, NY: John Wiley & Sons, Inc, 1998.
- [17] P. R. Stricland, "The thermal equivalent circuit of a transistor," *IBM Journal*, 1959.
- [18] D. T. Zweidinger, R. M. Fox, J. S. Brodski, T. Jung, and S.-G. Lee, "Thermal impedance extraction for bipolar transistors," *IEEE Transactions on Electron Devices*, vol. 43, no. 2, 1996.
- [19] J. A. Higgins, "Thermal properties of power HBT's," *IEEE Transactions on Electron Devices*, vol. 40, no. 12, 1993.

- [20] G. B. Gao, M. Z. Wang, X. Gui, and H. Markoc, "Thermal design studies of high-power Heterojunction Bipolar Transistors", *IEEE Transaction on Electron Devices*, vol. 36, no. 5, pp. 854-862, 1989.
- [21] C.-W. Kim, N. Goto, and K. Honjo, "Thermal behavior depending on emitter finger and substrate configurations in power heterojunction bipolar transistors," *IEEE Transactions on Electron Devices*, vol. 45, no. 6, pp. 1190-1195, June 1998.
- [22] L. L. Liou and B. Bayraktaroglu, "Thermal stability analysis of AlGaAs/GaAs Heterojunction Bipolar Transistors with multiple emitter fingers," *IEEE Trans. Electron. Dev.*, vol. 41, no. 5, pp. 629-636, 1994.
- [23] W. Liu, S. Nelson, D. Hill, and A. Khatibzadeh, "Current gain collapse in microwave multifinger heterojunction bipolar transistors operated at very high power densities," *IEEE Transaction on Electron Devices*, vol. 40, no. 11, 1993.
- [24] A. Kager and J. J. Liou, "Two-dimensional numerical analysis of AlGaAs/GaAs Heterojunction Bipolar Transistors including the effect of graded layer, setback layer and self-heating," *Solid-State Electronics*, vol. 39, no. 2, pp. 193-199, 1996.
- [25] T. J. Smy, D. J. Walkey, and S. K. Dew, "A 3D simulation tool for integrated devices," *IEEE Trans on CAD*, vol. 20, no.1, pp. 50, 2001.
- [26] W. Liu and B. Bayraktaroglu, "Theoretical calculations of temperature and current profiles in multifinger heterojunction bipolar transistors," *Solid-State Electronics*, vol. 36, no. 2, pp 125-132, Feb. 1993.
- [27] J. Dell, T. S. Kalkur, Z. Meglicki, A. G. Nassibian, and H. L. Hartnagel, *Int. J. Electron.* vol. 57p. 155, 1984.

- [28] R. C. Joy and E. S. Schlig, "Thermal properties of very fast Transistors," *IEEE Transaction on Electron Devices*, vol. 17, no. 8, 1970.
- [29] W. Zhou, S. Sheu, J. J. Liou, and C. I. Huang, "A multiemitter finger AlGaAs/GaAs HBT model including the effect of two-dimensional temperature distribution on emitter fingers," *Solid-State Electronics*, vol. 42, no. 5, pp. 693-698, 1998.
- [30] H. S. Carslaw and J. L. Jaeger, *Conduction of Heat in Solids*, 2nd ed., London: Oxford University Press, 1959.
- [31] D. J. Walkey, T. J. Smy, D. Marchesan, H. Tran, and M. Schroter, "A scalable thermal model for trench isolated bipolar devices," *Solid-State Electronics*, vol. 44, pp. 1373-1379, 2000.
- [32] J. J. Liou, L. L. Liou, and C. I. Huang, "Analytical model for the AlGaAs/GaAs multiemitter finger HBT including self-heating and thermal coupling effects," *IEEE Transaction on Electron Devices*, vol. 40, p. 1570, 1993.
- [33] A. Kager, J. J. Liou, L. L. Liou, and C. Huang, "A semi- numerical model for multiemitter finger AlGaAs/GaAs HBTs," *Solid-State Electronics*, vol. 37, no. 11, pp. 1825-1832, 1994.
- [34] J. Higgins, "GaAs HBT: A second generation microwave power amplifier transistor," *Microwave J.*, vol. 34, no. 5, p. 176, 1991.
- [35] D. J. Walkey, T. J. Smy, H. Tran, D. Marchesan, and M. Schroter, "Prediction of thermal resistance in trench isolated bipolar device structures," *IEEE BCTM*, vol. 12, no. 3, pp. 207-210, 1998.
- [36] J. E. Sergeant and A. Krum, *Thermal Management Handbook for Electronic Assem-*

blies, New York, NY: McGraw-Hill, 1998.

[37] R. L. Prichard, *Electrical Characteristics of Transistors*, pp. 593-599. New York, NY: Mc Graw-Hill, .

[38] A. Khatibzadeh, B. Bayraktaroglu, and T. Kim, *Monolithic Circuits Symposium*. Albuquerque, New Mexico, pp. 47-50, 1992.

[39] W. Liu, "Note, the temperature and current profiles in an emitter finger as a function of the finger length," *Solid-State Electronics*, vol. 36, no. 12, pp. 1787-1789, 1993.

[40] F. Bonani, *at al*, "Large-scale, computer-aided thermal design of power GaAs integrated devices and circuits," *IEEE GaAs IC Symp. Dig.*, pp. 141-144, 1995.

[41] S. Kakac and Y. Yener, *Heat Conduction*, 3rd ed., Washington, DC: Taylor & Francis, 1993.