Particle-Based Modeling of Reliability for Millimeter-Wave GaN Devices for Power Amplifier Applications

by

Alvaro D. Latorre Rey

A Dissertation Presented in Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy

Approved April 2018 by the
Graduate Supervisory Committee:

Marco Saraniti, Chair
Jennifer Kitchen
Stephen M. Goodnick
Trevor Thornton

ARIZONA STATE UNIVERSITY

May 2018

ABSTRACT

In this work, an advanced simulation study of reliability in millimeter-wave (mm-wave) GaN Devices for power amplifier (PA) applications is performed by means of a particle-based full band Cellular Monte Carlo device simulator (CMC). The goal of the study is to obtain a systematic characterization of the performance of GaN devices operating in DC, small signal AC and large-signal radio-frequency (RF) conditions emphasizing on the microscopic properties that correlate to degradation of device performance such as generation of hot carriers, presence of material defects and self-heating effects. First, a review of concepts concerning GaN technology, devices, reliability mechanisms and PA design is presented in chapter 2. Then, in chapter 3 a study of non-idealities of $AlGaN/GaN$ heterojunction diodes is performed, demonstrating that mole fraction variations and the presence of unintentional Schottky contacts are the main limiting factor for high current drive of the devices under study. Chapter 4 consists in a study of hot electron generation in GaN HEMTs, in terms of the accurate simulation of the electron energy distribution function (EDF) obtained under DC and RF operation, taking into account frequency and temperature variations. The calculated EDFs suggest that Class AB PAs operating at low frequency (10 GHz) are more robust to hot carrier effects than when operating under DC or high frequency RF (up to 40 GHz). Also, operation under Class A yields higher EDFs than Class AB indicating lower reliability. This study is followed in chapter 5 by the proposal of a novel Π-Shaped gate contact for GaN HEMTs which effectively reduces the hot electron generation while preserving device performance. Finally, in chapter 6 the electro-thermal characterization of GaN-on-Si HEMTs is performed by means of an expanded CMC framework, where charge and heat transport are self-consistently coupled. After the electro-thermal model is validated to experimental data, the assessment of self-heating under lateral scaling is considered.

ACKNOWLEDGMENTS

The successful completion of a Ph.D. program is, in my experience, a group effort that reflects the contributions and hard work of many people, not only the student leading the research. For this reason, I would like to begin this dissertation by gratefully acknowledging all of those who one way or another have been part of this work.

First, I want to thank my advisor Dr. Marco Saraniti for his robust support, guidance and patience. Doing research for a living guarantees a world of challenges and uncertainty. However, Dr. Saraniti provided me with a solid ground of technical and financial resources, allowing me to pursue my interests and successfully produce measurable contributions. Special thanks are due to Dr. John D. Albrecht, at Michigan State University, East Lansing - MI, for the continuous technical support, discussions and suggestions. I am very thankful for the technical support provided by Dr. Gregg Jessen at the Air Force Research Laboratory (AFRL), in terms of measurements of experimental devices used in this research. Also, I am very grateful for the funding that supported this work, provided by the Air Force Office of Scientific Research (AFOSR) and the AFRL. Finally, I would like to express my gratitude to the members of my committee: Dr. Stephen Goodnick, Dr Jennifer Kitchen and Dr. Trevor Thornton for disposing of their time and technical expertise in evaluating this work.

In a more personal note, I want to express my deepest gratitude to Cyndi, my most valuable finding during the Ph.D., my best friend and soon to be wife, for patiently walking with me through this strenuous path. Her love, kindness and support have been a light in the darkness, from which I have found the strength and encouragement to move forward. I am forever in debt for the unconditional love and support my parents and sister have given me throughout all these years. Without them none of this would have been possible. At last, but not least, I want to thank my friends and colleagues for making these experience a joyful and enjoyable one.

TABLE OF CONTENTS

Page

LIST OF FIGURES vii

CHAPTER

1 INTRODUCTION 1

1.1 Modeling of Semiconductor Devices 7

1.2 Overview of this Work 13

2 REVIEW OF GAN TECHNOLOGY AND POWER AMPLIFIERS 15

2.1 Polarization Effects in GaN 15

2.2 $AlGaN/AlN/GaN$ Devices 20

2.2.1 High Electron Mobility Transistors (HEMTs) 21

2.2.2 Hot Electron Transistors (HETs) 28

2.3 Reliability of GaN HEMTs 30

2.3.1 Hot Electron Effects 31

2.3.2 Self-heating Current Collapse 36

2.3.3 Threading Dislocations 38

2.3.4 Inverse Piezoelectric effect 40

2.3.5 Electrical-Stress Testing Conditions 42

2.4 Radio Frequency Power Amplifiers 47

2.4.1 Class A Operation 57

2.4.2 Class AB Operation 60

2.5 Conclusions 64

3 STUDY OF NON-IDEALITIES IN III-N ISOTYPE HETERJUNCTION DIODES 65

3.1 Introduction 65

3.2 Characterization of Experimental Device with Ideal Layout 68

CHAPTER Page

3.3 Mole Fraction Variations in the $Al_xGa_{1-x}N/GaN$ Barrier 72
3.4 Effect of Emitter and Base Schottky Contacts 77
3.5 Conclusions 86
4 HOT ELECTRON GENERATION UNDER LARGE-SIGNAL RF OPERATION OF GAN HEMTS 87
4.1 Introduction 88
4.2 GaN HEMT Model Validation 89
4.3 Large Signal RF Characterization 91
4.4 Hot Electron Generation 93
4.5 High Frequency RF Operation 98
4.6 High Temperature Operation 101
4.7 Comparison between Class A and Class AB Operation 105
4.8 Conclusions 109
5 A Π-SHAPED GATE DESIGN FOR REDUCING HOT ELECTRON GENERATION IN GAN HEMTS 111
5.1 Introduction 111
5.2 Performance of Π-gated HEMTs 113
5.2.1 Principles of Operation 113
5.2.2 DC, small-signal AC and large-signal RF Operation 115
5.3 Reduced Hot Electron Generation 119
5.3.1 Energy Distribution in DC Regime 121
5.3.2 Energy Distribution under RF Operation 125
5.4 Conclusions 127

CHAPTER Page

6 ELECTRO-THERMAL CHARACTERIZATION OF *GAN* HEMT ON SI THROUGH SELF-CONSISTENT ENERGY BALANCE-CELLULAR MONTE CARLO DEVICE SIMULATIONS 128
6.1 Introduction 129
6.2 Energy Balance Flux Approach 130
6.3 Electro-Thermal Device Simulations 134
6.4 DC Electro-Thermal Characterization 138
6.5 Impact of Lateral Scaling on Self-Heating 142
6.5.1 Gate-to-Drain L_{GD} Scaling 144
6.5.2 Source-to-Gate L_{SG} Scaling 147
6.6 Conclusions 152
7 SUMMARY 154
REFERENCES 158

LIST OF FIGURES

Figure Page

1.1 Comparison of semiconductor material properties which enable high power and high frequency operation [3]. 2

1.2 Diagram of Power-Frequency range of different technologies of semiconductor devices [4]. 3

1.3 Commercially available GaN HEMT products released by Cree Inc, showing (a) Discrete 50V GaN HEMT devices for radar and military communications and (b) High power MMIC for satellite communications. 5

1.4 Flowchart of the particle-based Cellular Monte Carlo algorithm (CMC). 10

1.5 Full electronic-band structure (a) and Phonon dispersion (b) of Wurtzite GaN (HCP) used in this work. 11

1.6 GaN scattering rates at $T = 300\,\mathrm{K}$ used in this work. 12

2.1 Crystal structure of GaN in Wurtzite form showing (a) Highlighted Hexagonal Closed Packed (HCP) sublattice for cation Ga, and (b) Typical growth planes, indicating the Ga-polar c-plane $\{0001\}$ and the semi-polar $\{11\bar{2}2\}$ plane. 16

2.2 Polarization vectors (Spontaneous P_{SP} and Piezoelectric P_{PE}) and equivalent interface charge for Ga-polar $AlGaN/GaN$ heterostructures under different growth conditions of the top layer: (a) AlGaN relaxed, (b) AlGaN Tensile strain and (c) GaN Compressive strain, while keeping the bottom layer relaxed. 19

2.3 Basic structure of a Ga-polar $AlGaN/AlN/GaN$ HEMT showing (a) device layout and (b) non-uniform grid used for CMC device simulations. 22

Figure Page

2.4 Band diagram of the simulated GaN-HEMT from figure 2.3 at DC bias of V_{DS} = 5 V and V_{GS} = -5 V. Panel (a) shows the 3D conduction band profile with the carrier concentration as contour, while panel (b) shows the vertical profile of the conduction band along with the carrier concentration, taken from a slice on the source-to-gate access region. . . . 26

2.5 DC characterization of the simulated GaN-HEMT from figure 2.3 showing (a) $I_{DS} - V_{DS}$ and (b) $I_{DS} - V_{GS}$ or output and transfer characteristics respectively. 27

2.6 Small-signal AC characterization of the simulated GaN-HEMT from figure 2.3 showing short-circuit current gain and unilateral power gain curves as a function of frequency, for extraction of f_t and f_{MAX}. 28

2.7 Hot Electron Transistor (HET) device showing (a) device layout and (b) 2 dimensional band diagram along the vertical direction. Also indicated in (b) a sketch of some carriers and their trajectories, along with the current components under typical DC operation. 29

2.8 Signature degradation of (a) drain current and (b) transconductance caused by hot-electron effects in the AlGaN barrier of an experimental GaN HEMT. The figure is reproduced from original research presented by Meneghini et al. [26]. 35

2.9 Signature degradation of self-heating current collapse observed in the I_{DS}-V_{DS} curves of an experimental GaN HEMT. The figure is reproduced from the white paper presented by Zanoni [23]. 37

2.10 Effect of thread dislocations (TDs) in the electrical performance of GaN HEMTs, shown with simulations of (a) I_{DS}-V_{GS} and (b) I_{DS}-V_{DS} curves at three different TDs density N_{TDs}. 40

2.11 Signature degradation of drain current (left axis) and gate leakage current (right axis) in an experimental GaN HEMT caused by the inverse piezoelectric effect. Figure reproduced from original research presented by del Alamo and Joh [69]. 41

2.12 Electrical stress characterization of GaN HEMTs, showing (a) I_{DS}-V_{DS} curves indicating stress bias points. Also, in (b) to (d) the electric field mapping is presented on 3D band diagrams of the device for 3 stress bias points. The spheres represent a snap-shot of real-space distribution of carriers with its corresponding energy as a contour. 44

2.13 Voltage or Current signal waveforms used in short-term life-time reliability testing, identifying the 3 main protocols (a) Stepped-Stress, (b) Step-Stress Recovery and (c) Stress-Recovery Cyclic Stress loading. 46

2.14 Schematic of Power Amplifier (PAs) circuits. 48

2.15 Sketch of figures of Merit for a typical power amplifier for single-tone input, showing output power P_{out} and transducer gain G_T on the left-axis, and power added efficiency PAE in the right-axis as a function of the input power P_{in}. Also indicated the 1dB compression point P_{1dB}. . . 52

2.16 Sketch of Inter Modulation Distortion (IMD) measurements for a power amplifier (PA) with a two-tone input signal separated by a small frequency $\delta = f_{\omega 2} - f_{\omega 1}$. Determining IIP3 and OIP3 can be done from measuring (a) output power delivered to fundamental and harmonic frequencies vs input power sweep, and (b) Output power spectrum of the PA for a given input. 54

2.17 Example of the output spectrum of a PA with a single carrier WCDMA input signal band limited to 5 MHz, evidencing the spectral regrowth used in Adjacent Channel Power Ratio (ACPR) measurments. A typical spectrum mask is also included. 56

2.18 Schematic of a Class A power amplifier circuit topology, simulated with commercial software Advanced Design System (ADS) by Keysight.The active device is simulated by using the analytical model by Statz, and the input and output matching networks are identified. 58

2.19 Characterization of a Class A power amplifier, obtained with simulations of the schematic shown in figure 2.18. In the panels is presented (a) I_{DS}-V_{DS} space along with load-lines, (b) ac current waveform $i_{ds}(\omega)$, (d) ac drain voltage waveform $v_{ds}(\omega)$ and (c) typical figures of merit. . . 59

2.20 Schematic of a Class AB power amplifier circuit topology, simulated with commercial software Advanced Design System (ADS) by Keysight.The active device is simulated by using the analytical model by Statz, and the input and output matching networks are identified. 62

2.21 Characterization of a power amplifier in Class AB topology with tuned-load. The Results correspond to simulations of the schematic from figure 2.20, showing (a) I_{DS}-V_{DS} space along with load-lines, (b) ac current waveform $i_{ds}(\omega)$, (d) ac drain voltage waveform $v_{ds}(\omega)$ and (c) typical figures of merit. 63

3.1 Experimental device layout corresponding to the ideal structure of the emitter-base diode. The emitter and base contacts are included. 68

3.2 Profile of the conduction band in thermal equilibrium (V_{be}=0.00V), along with the concentration of electrons (right-axis) as obtained with CMC simulations. 69

3.3 Evolution with bias of the conduction band profile along the y-direction of the EB diode as obtained with CMC simulations. 70

3.4 Conduction Band in 3D and a snap-shot of the electron distribution in real space at V_{be}=1.25V. The color of each super-particle corresponds to its energy. 71

3.5 IV characteristic in semi-log scale of emitter-base diode, showing experimental data along with simulation results obtained by CMC, and TCAD Sentaurus simulations. 73

3.6 Vertical profile of the conduction band in equilibrium (left-axis) for $AlGaN$ with mole fractions x = 0.45 and 0.50. The corresponding concentration of electrons is also shown in the right-axis. 75

3.7 IV characteristic in semi-log scale, showing experimental data, along with simulation results for mole fraction values x = 0.45 and 0.50, at the $Al_xGa_{1-x}N/GaN$ barrier. 76

3.8 IV characteristic in semi-log scale of the EB diode showing experimental data, along with TCAD simulation results for different values of parasitic series resistance R_S in the contacts. 77

3.9 Vertical conduction band profile at equilibrium, where a Schottky contact is placed at the emitter with barrier height ϕ_2=0.45eV. 78

3.10 Vertical conduction band profile with and without emitter Schottky contact for a) $V_{be} = 1.0\,\mathrm{V}$ and b) $V_{be} = 1.5\,\mathrm{V}$, with Schottky barrier height $\phi_2 = 0.45\,\mathrm{eV}$, while keeping an ohmic base contact. 80

3.11 IV characteristic in semi-log scale, showing experimental data along with simulation results including the effect of a Schottky contact at the emitter, for different barrier heights. 81

3.12 Vertical conduction band profile at equilibrium, with a Schottky contact placed at the base with barrier height $\phi_2 = 0.65\,\mathrm{eV}$, while keeping an ohmic emitter contact. 82

3.13 Vertical conduction band profile with and without base Schottky contact for a) $V_{be} = 0.50\,\mathrm{V}$ and b) $V_{be} = 1.5\,\mathrm{V}$, with $\phi_2 = 0.65\,\mathrm{eV}$. 83

3.14 IV characteristic in semi-log scale showing experimental data, along with simulation results including the effect of a Schottky contact at the base, with different barrier heights. 84

3.15 Semi-log IV characteristic, showing experimental data along with the simulated curve corresponding to a structure with $Al_{0.50}Ga_{0.50}N$, $\phi_{base} = 0.65\,\mathrm{eV}$, $\phi_{emitter} = 0.45\,\mathrm{eV}$. 85

4.1 Device layout of the fabricated GaN HEMT used in the CMC model. . . 90

4.2 Experimental and simulated (a) output and (b) transfer (along with Transconductance G_m) characteristics. 91
4.3 Power amplifier characterization under Class AB (RF) operation, showing (a) dynamic load-lines and (b) FOM, for $f_{req} = 10\,\text{GHz}$ and tuned-load $Z_{Load} = 100\,\Omega$. 93
4.4 Electron energy distribution function for the DC bias points (1) to (6) indicated in figure 4.3 (a) with solid dots. 94
4.5 Close up under the gate of the profile along the channel of (a) x-component of the Electric field E_x, (b) electron concentration and (c) total scattering at DC bias points (4), (6) and (1) from figure 4.4, which trace the load-line in figure 4.3 (a). 96
4.6 Electron energy distribution function under RF Class AB operation for different input powers. Curves for DC peak generation (4) and LSOP (3) from figure 4.4 also shown for comparison. 97
4.7 Frequency study of power amplifier characterization under Class AB (RF) operation, showing (a) dynamic load-lines and (b) FOM, at frequencies of $f_{req} = 10\,\text{GHz}$ and 40 GHz and tuned-load $Z_{Load} = 100\,\Omega$. . . 99
4.8 Frequency study of the energy distribution function for Class AB operation at (a) $P_{out} = 26\,\text{dBm}$ and (b) $P_{out} = 30\,\text{dBm}$. 100
4.9 Electron energy distribution function for the DC bias points (I), (II) and (3), (4) indicated in figure 4.7 (a) with solid dots, corresponding to DC points along the load-lines at 40 GHz and 10 GHz respectively. 102

Figure Page

4.10 Effect of temperature in Class AB operation for $f_{req} = 10\,\text{GHz}$ and $Z_{Load} = 100\,\Omega$ showing (a) I_{DS}-V_{DS} curves along with load-lines, and (b) figures of merit. 103

4.11 Effect of temperature on (a) the total scattering rate and (b) the x-component of the electron velocity, shown in a region under the gate along the channel. 103

4.12 Temperature dependence of the distribution function for Class AB operation at $P_{in} = 5\,\text{dBm}$, for $f_{req} = 10\,\text{GHz}$ and $Z_{Load} = 100\,\Omega$. 104

4.13 Power amplifier characterization under Class A and Class AB operation for $f_{req} = 10\,\text{GHz}$ and tuned-load $Z_{Load} = 100\,\Omega$, showing (a) dynamic load-lines along with I_{DS}-V_{DS} curves. Also, the typical figures of merit are shown in (b) where lines with and without symbols correspond to Class A and Class AB respectively. 106

4.14 Electron energy distribution function under Class A (lines with symbols) and Class AB (lines) PA operation, at 3 input powers $P_{in} = 0\,\text{dBm}$, 10 dBm and 15 dBm. 107

4.15 Frequency study of the electron energy distribution function in Class A operation for $f_{req} = 10\,\text{GHz}$ (lines) and 40 GHz (lines with symbols) at input power of $P_{in} = 0\,\text{dBm}$ and 18 dBm. 109

4.16 Comparison of electron energy distribution function at high frequency operation, under Class A (lines with symbols) and Class AB (lines) topologies at input power of $P_{in} = 0\,\text{dBm}$ and 18 dBm. 110

5.1 Device layout showing (a) the experimental T-gate geometry from fabricated devices (b) the proposed Π-gate, including design rules. 114

5.2 Simulated DC curves (dashed and solid lines) showing (a) output and (b) transfer (along with transconductance G_m) characteristics for T-gate and Asymmetric-Π-gate devices. Experimental data (symbols) corresponds to the T-gate device. 116

5.3 Small-Signal AC parameters (a) Cut-off frequency f_t and (b) Maximum Oscillation Frequency f_{MAX} as a function of spacer region length d for T-gate ($d = 0$), symmetric-Π-gate ($L_{G1} = L_{G2} = 70\,\text{nm}$) and asymmetric-Π-gate ($L_{G1} = 100\,\text{nm}$, $L_{G2} = 40\,\text{nm}$), corresponding to two stem heights h. 118

5.4 Power amplifier characterization under Class AB (RF) operation, showing (a) dynamic load-lines for P_{in}=12 dBm at f_{req}=10 GHz along with $I_{DS} - V_{DS}$ curves and (b) P_{out}, Gain and PAE for $f_{req} = 10\,\text{GHz}$ and 40 GHz with Z_{Load}=100Ω. 120

5.5 Close up of the profile along the channel (x-component) of (a) the conduction band E_C and (b) the electron velocity at $V_{GS} = -3\,\text{V}$ and $V_{DS} = 30\,\text{V}$ for T-gate and Asymmetric Π-gate devices. 121

5.6 Electron energy distribution function for the DC bias points (1) and (2) shown in figure 5.4 (a) for T-gate and asymmetric Π-gate devices. . . 122

5.7 Profile along the channel of (a) x-component of the electric field, (b) electron concentration, (c) total scattering rate and (d) x-component of the electron velocity for T-gate and asymmetric-Π-gate at DC bias point (1) from figure 5.4. Solid bars in (a) show the position of the metallurgical gate contacts for each layout. 124

5.8 Electron energy distribution function under RF Class AB operation at (a) X-band 10 GHz and (b) K-band 40 GHz, for different input powers. 126

6.1 Flowchart of thermal energy transport in semiconductor devices. 131

6.2 GaN-on-Si HEMT device layout showing (a) Experimental Layout by Altuntas [112] and (b) Simulated device layout..... 135

6.3 3D conduction band of the HEMT with temperature profile for Acoustic phonon modes at DC bias of $V_{GS} = 2\,\text{V}$, $V_{DS} = 10\,\text{V}$. 137

6.4 3D conduction band of the HEMT with temperature profile for Optical phonon modes at DC bias of $V_{GS} = 2\,\text{V}$, $V_{DS} = 10\,\text{V}$. 137

6.5 Validation of the electro-thermal model with experimental curves corresponding to (a) I_D-V_{DS} for $V_{GS} = 0\,\text{V}$ and 2 V, and (b) I_D-V_{GS} for V_{DS}=6 V. Isothermal simulations with $T = 300\,\text{K}$ are also shown..... 139

6.6 Profile along the channel of (a) electron x-velocity and (b) total scattering at DC bias point $V_{GS} = 2\,\text{V}$, $V_{DS} = 5\,\text{V}$, comparing isothermal simulations at $T = 300\,\text{K}$ and self-heating curves. 140

6.7 Profile along the channel of (a) x- component of the electric field, (b) acoustic mode temperature and (c) optical mode temperature for V_{GS}=2V at three different V_{DS}. 141

6.8 Effect of gate-to-drain L_{GD} lateral scaling on the (a) I_{DS}-V_{DS} for $V_{GS} = 2\,\text{V}$ and (b) I_{DS}-V_{GS} for V_{DS}=6 V, also shown the transconductance G_m in right-axis of (b). Self-heating effects included in all simulations..... 145

6.9 Profile along the channel of (a) electric field (b) electron velocity for three different values of L_{GD}. The DC bias point $V_{GS} = 2\,\text{V}$, $V_{DS} = 5\,\text{V}$ corresponds to the point with constant DC power from figure 6.8..... 145

6.10 Effect of lateral scaling of L_{GD} on the profile along the channel of acoustic mode temperature (left-axis) and optical mode temperature (right-axis) for DC bias point from figure 6.8 corresponding to constant DC power. 147

6.11 Effect of source-to-gate L_{SG} lateral scaling, while keeping L_{GD} constant, on the (a) I_{DS}-V_{DS} for $V_{GS} = 2\,\mathrm{V}$ and (b) I_{DS}-V_{GS} for V_{DS}=6 V, also shown in right-axis of (b) the transconductance G_m. Self-heating effects included in all simulations. 148

6.12 Profile along the channel of the electric field for two different L_{SG} values. DC bias points along the line of constant DC power from figure 6.11 (b), implying different values of V_{GS} were used, while keeping V_{DS} and I_{DS} constant. 149

6.13 Effect of lateral scaling of L_{SG} on the profile along the channel of acoustic mode temperature (left-axis) and optical mode temperature (right-axis) for DC bias points from figure 6.11 (b) corresponding to the line with constant DC power. 150

6.14 Profile along the channel of (a) electric field (b) electron velocity for two different $L_{SG} = 200\,\mathrm{nm}$ and 400 nm where the DC bias points were chosen along the constant electric field line from figure 6.11 (b). 151

6.15 Profile along the channel of acoustic mode temperature (left-axis) and optical mode temperature (right-axis) under lateral scaling of L_{SG}, along the constant electric field line from figure 6.11 (b). 152

Chapter 1

INTRODUCTION

Wireless communications have become a ubiquitous element in modern society, resulting in the upsurge of the telecommunications and radio-frequency (RF) markets with a wide variety of applications and ever-more demanding requirements. In the field of personal mobile technology, internet of things (IoT) and broadband internet connections, one of the main goals is to provide high data transfer rates, such as 1 Gbps in 5G cellular technology [1] and up to 8 Gbps under the IEEE 802.11ad WiGig technology [2]. In addition, since most of these applications are battery-powered there is a strong need for high efficiency and portability. On the other hand, the evolution of airborne, aerospace, satellite and defense applications require not only high data rates, but also high power ratings and in the case of radar technologies high resolution.

The improved product-level performance correlates directly with stringent system level specifications, since high data rates require wider bandwidths, portability translates into smaller and more efficient components and the high power ratings call for higher reliability and efficiency. All these needs have in common that can be delivered in RF systems operating in super high frequencies (SHF) of 3 to 30 GHz and millimeter-wave (mm-wave) frequencies between 30 and 300 GHz. This is why considerable efforts have been directed over the last couple of decades to the development of the technology necessary to implement efficient and reliable high-frequency RF transmitters. In such systems, mm-wave transistors play an essential role as the power amplifier (PA) or the output stage that feeds the antenna [3].

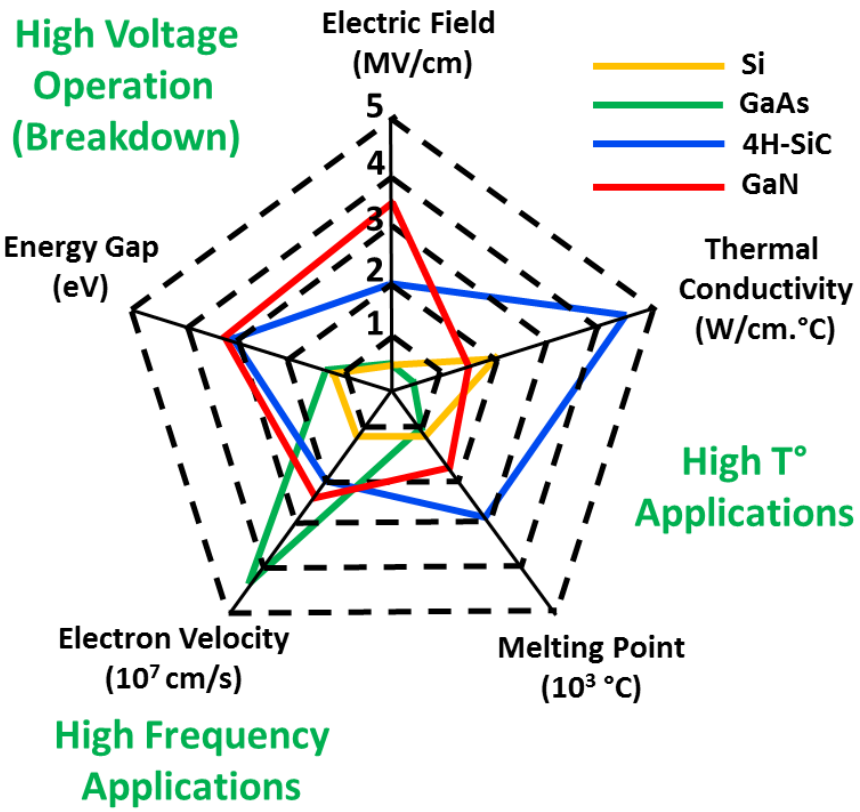

Figure 1.1: Comparison of semiconductor material properties which enable high power and high frequency operation [3].

Even though the semiconductor industry is heavily dominated by Silicon devices, the system requirements of state-of-the-art RF transmitters have driven the development of new microwave and mm-wave transistor technologies based on other materials such as SeGe, GaAs, SiC and GaN, which offer better physical properties to achieve the desired performance. Figure 1.1 shows a diagram comparing the relevant material properties that enable high frequency and high power operation [3].

For the latter, transistors must have high breakdown voltage, which is a process driven by the electric field. In this sense, wide bandgap materials like 4H-SiC and GaN are suitable options because of the high critical electric field required for impact-ionization. In addition, devices operating under high power conditions must be able to withstand and dissipate high temperatures, which demands for high thermal con-

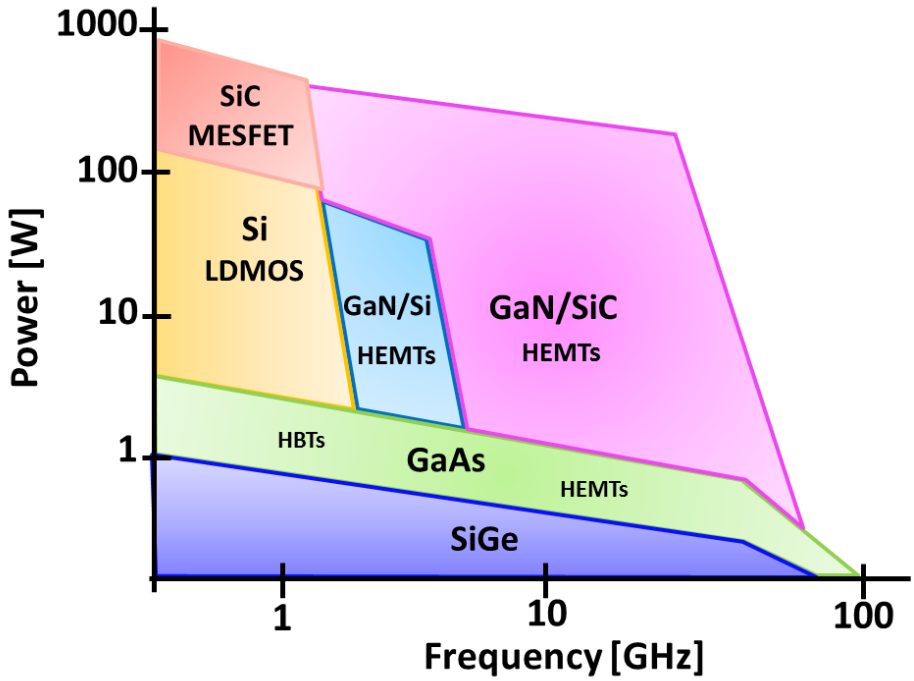

Figure 1.2: Diagram of Power-Frequency range of different technologies of semiconductor devices [4].

ductivity. From this point of view, SiC provides the best option. In terms of high frequency operation, materials with high electron velocities are the most adequate, and GaAs provides the highest, followed by SiC and GaN.

The intrinsic advantages provided by alternative materials have been exploited in devoloping different device technologies suitable for a wide range of power-frequency applications, as it is shown in figure 1.2 [4]. It must be noticed that even though Si is outperformed by all the other materials, advanced devices built with Si such as the laterally diffused MOSFET (LDMOS) transistors allow for improved performance with respect to conventional Si-MOSFETs, enabling its operation in frequencies below 3 GHz and high power between 10 and 100 W typical in power switching applications [5]. However, the properties of 4H-SiC are being exploited in cutting-edge metal-semiconductor field effect transistors or MESFETs providing efficient power switches with low on-resistance and significantly higher power operation [6].

In terms of PAs for hand-sets necessary in personal mobile communications, SiGe p-MOSFETs with low operating voltages of 2 V to 3 V for low power of 1 W to 2 W and high frequency are very attractive because of its integration potential with the dominant analog and digital Si CMOS technology. However, as frequency and linearity requirements have become more demanding, GaAs devices such as the Heterojunction Bipolar Transistor (HBT) and High Electron Mobility Transistor (HEMT) [7], have been adopted as the mainstream technology for applications up to 5 W.

For applications such as cellular base stations, satellite communications and radars for self-driving automotive or defense, typical RF amplifiers require high power above 10 W and frequencies of operation in the mm-wave range. In this case, GaN HEMTs fabricated on both Si and SiC substrates are the most prominent technology [8], enabled by the ability to form heterojunctions in the $AlGaN/GaN$ system. The wide bandgap of this heterostructure results in high breakdown voltage, enabling transistors to be biased at high drain voltages. This not only reduces the need for voltage conversion but also improves the power conversion efficiency critical for PAs [3].

In addition, unlike the $AlGaAs/GaAs$ heterostructure HEMTs that rely on a modulation doping layer to form a channel [9], III-nitrides heterojunctions like $Al_xGa_{1-x}N/GaN$ show high density of two-dimensional electron gas (2DEG) at the heterointerface due to the presence of a strong polarization effect [10], providing high carrier density and high electron mobility due to a reduced impurity scattering. The combined effects of high carrier density and high mobility result in high current and power densities even at high frequency operation [11; 12]. This in turn allows fabricating devices with small periphery resulting in high output impedance and easier matching with low loss circuits [13]. Furthermore, the high breakdown and high thermal conductivity of substrates such as SiC allow for high temperature operation and better noise performance than SiC MESFETs.

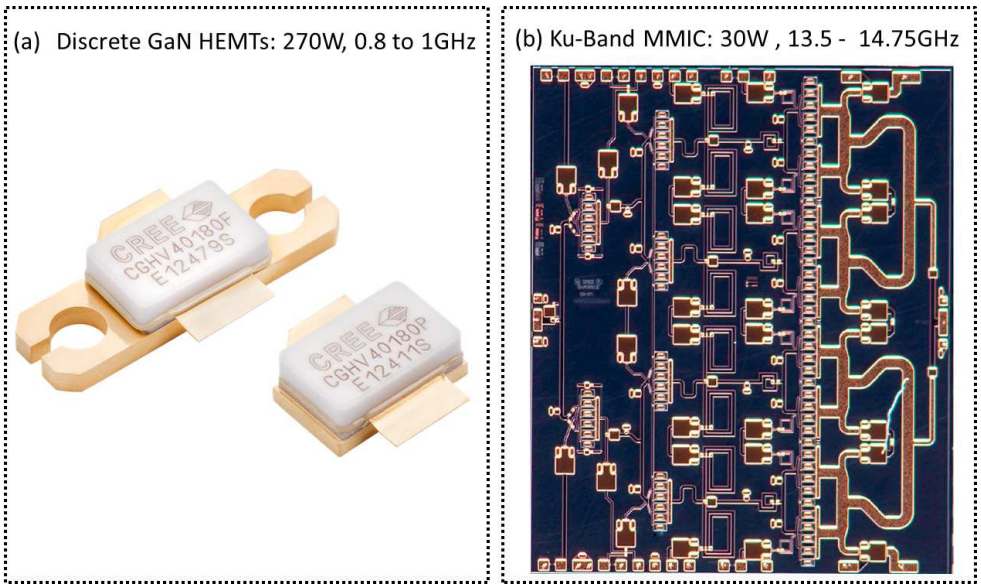

Figure 1.3: Commercially available GaN HEMT products released by Cree Inc, showing (a) Discrete 50V GaN HEMT devices for radar and military communications and (b) High power MMIC for satellite communications.

In the last decade, extraordinary progress both at industrial and research level have made $AlGaN/GaN$ HEMTs technology commercially available [14; 15], with a market size valued at USD. 974.9 million in 2016 [16], including both GaN on Si HEMTs for power switching and GaN on SiC HEMTs for mm-wave PAs. A couple of commercially available products recently released by Cree are shown in figures 1.3 (a) and (b) corresponding to discrete 50 V GaN HEMT devices for radar and military communications and a high power MMIC (monolithic microwave integrated circuit) for satellite communications [17].

The development of the technology necessary to fabricate these devices is due to major progress in improving the material growth quality and reducing its cost [18]. In addition, process flows for fabrication of devices were developed including passivation techniques that minimize current dispersion [19], along with advanced contact engineering including field plates that increase the RF power density [11]. The leap in GaN processing technology is such that it has enabled the development at a research

level of devices for sub-millimeter RF operation, in particular Hot Electron Transistors (HETs) based on vertical $AlGaN/GaN$ heterojunction diodes are being studied as suitable candidates for terahertz operation [20; 21].

In terms of robustness of the technology, remarkable mean-time-to-failure values of 10^7 h and greater have been reported [22], and some HEMT products have been certified matching Si CMOS manufacturing metrics and process control [14]. Despite this significant achievement, reliability concerns are still a roadblock to fully deploy the technology. This issue arises from the fact that niche applications of GaN HEMTs often imply harsh operating conditions such as high temperature and power dissipation, resulting in high electric fields and current densities both under DC and high-frequency operation. In this sense, reliability concerns are still a challenge and the subject of active research [23].

In terms of device failure of intrinsic GaN HEMT devices, two main degradation mechanisms are studied: trap generation induced by hot-electrons and high temperature effects, and electric field-induced current degradation attributed to the inverse piezoelectric effect [24; 25]. Under reverse-bias stress, HEMTs show performance degradation in terms of high gate-leakage current, typically attributed to defect-generation through inverse piezoelectric effect [22]. Under forward bias, current collapse is associated to self-heating effects and defect generation induced by hot-electrons [26]. Even though these phenomena have been widely studied experimentally by means of DC life-time reliability testing, a deeper understanding of the underlying physical mechanisms responsible for device degradation is still necessary, in particular when the electrical stress is induced in RF operation. It is in this context that advanced techniques for simulation of semiconductor devices and materials have made important contributions and have been established as a powerful tool for the study of physical mechanisms of operation and reliability.

1.1 Modeling of Semiconductor Devices

Numerical simulations are widely used in industry and research because they allow for a higher level of accuracy than compact models, necessary to capture and predict highly non-linear behavior of devices operating under high fields. In addition, simulations are a traditional surrogate for experiments in the presence of cost, time or even experimental feasibility constrains. Furthermore, computer aided design (CAD) provides an accurate way to advance the development of new technologies. From this point of view, numerical simulations in the context of the broad topic of computational science have been established as the third paradigm of science, adding to theory and experiment [27].

Simulation of semiconductor devices involves several factors. First, it is necessary to obtain the charge and potential distributions self-consistently, which is the realm of electromagnetic theory in the form of Poisson's equation, for which elegant and efficient computational solutions have been developed [28]. Then, the transport of carriers in the semiconductor material, or the way charge moves in the presence of electric fields and potentials through a crystal, requires a physical model to be solved in a self-consistent manner with the electrostatic problem. Finally, for accurate device simulations boundary conditions must be included in order to capture the effect of surfaces and contacts.

Under the semi-classical approach, a complete description of carrier transport can be done in terms of the single-particle distribution function $f\left(\vec{r},\vec{k},\vec{t}\right)$, which gives the probability of finding one carrier with crystal momentum $\vec{k}$, at position in real space $\vec{r}$ at time $\vec{t}$. This definition can be extended to an ensemble of non-interacting particles [29]. The distribution function describes the average distribution of carriers in phase-space, e.g. in real and momentum space, and can be used to calculate

meaningful physical quantities such as carrier concentration, energy, velocity and currents by calculating averages on the ensemble [30]. In order to obtain $f\left(\vec{r},\vec{k},\vec{t}\right)$, the Boltzmann Transport Equation (BTE) given by 1.1 must be solved [31]:

$$\frac{\partial f}{\partial t} = -\vec{v}\cdot\nabla_{\vec{r}} f - \dot{\vec{k}}\cdot\nabla_{\vec{k}} f + \left(\frac{\partial f}{\partial t}\right)_{coll}, \tag{1.1}$$

where $\vec{v}$ is the carrier velocity, and the last term on the right hand side of 1.1 is the rate of change of f due to collisions. In simple terms, the BTE describes the time evolution of the distribution function in momentum-space, accounting for all possible mechanisms by which it can change, such as driving forces given by applied fields and dissipative scattering forces. However, the BTE is a multi-dimensional, non-linear differential equation and solving it is a cumbersome task [32]. The strategy used to do so, be it analytical or numerical, defines the characteristics of a device simulator.

A widespread technique is the drift-diffusion method [33] based on the first moment of the BTE, where approximations applied to the BTE yield the well known drift-diffusion current equations, which can be self-consistently solved with the electrostatic problem, taking into account analytical carrier statistics [34]. The main advantage of this method is its relative computational efficiency at the expense of relying on significant simplifications that fail under relevant operating conditions of state-of-the-art devices.

The hydrodynamic model is an improved version of the drift-diffusion approach which allows to study non-equilibrium conditions, however it still relies heavily on parameters such as energy relaxation times, mobility amd diffusion coefficients among others, whose accuracy determine the validity of the predictions and results obtained under this approach [35]. Nevertheless, this method is widely accepted as the industry standard and several CAD packages such as Sentaurus TCAD tools by Synopsys are commercially available.

On the other hand, particle-based simulation based on the stochastic Monte Carlo methods have been established as a highly accurate way to solve the BTE [36], because it mimics the physics of carrier transport, providing an insightful microscopic description of device operation [30]. In this strategy, a population of carriers is represented by an ensemble of particles whose individual trajectories are tracked in both momentum and real space as they move through a device following Newton's laws, under the presence of electric fields and dissipative forces. The latter are the scattering rates, which correspond to the last term on the RHS of 1.1, and are formulated as a probability density function according to a quantum mechanical treatment [29].

Monte Carlo methods provide an exact statistical solution of the BTE, and from the resulting distribution function physical observables can be derived through averages of the ensemble. Moreover, if the number of particles (or trajectories) considered in the simulation is large enough, the averaged quantities of the ensemble (concentration, energy, velocity and currents) are a good approximation to the average behavior of the carriers in a real device. The flowchart of the particle-based Cellular Monte Carlo framework (CMC) used in this work [37] is presented in figure 1.4.

The algorithm begins with initializing carriers in real and momentum space, within the simulation domain representing the semiconductor device. Then the total charge is calculated by imposing the charge neutrality condition considering mobile carriers, ionized impurities and other fixed charges, which is used to solve Poisson's equation to obtain electric field and potential maps in real space. Then, the field is kept constant over a short time-step called the free-flight and carriers are allowed to move ballistically following Newtown's laws updating the carrier position in both real and momentum space. The following step corresponds to the carrier dynamics portion of the method, where a stochastic Monte Carlo procedure is used to select if the carrier scatters, and if it does then the final momentum state is selected and updated.

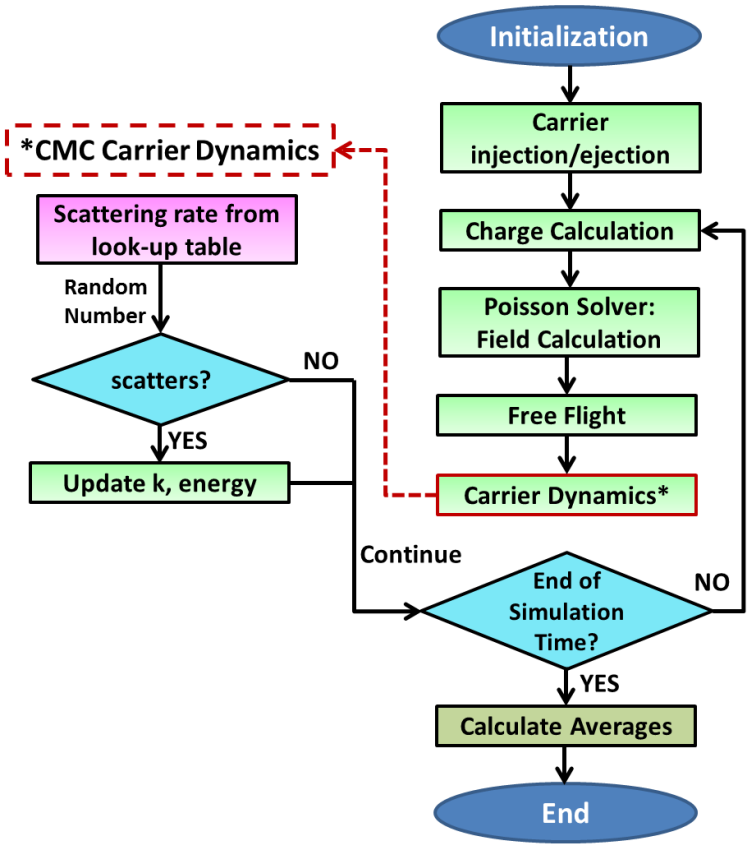

Figure 1.4: Flowchart of the particle-based Cellular Monte Carlo algorithm (CMC).

This process is iterated over all particles in the ensemble, which typically requires 4 to 5 free-flight steps before the charge and electric field are re-calculated, implying the use of two different time-steps one for the free-flight and one for the Poisson Solver. Choosing the free-flight time-step most be done so that all scattering processes can be resolved, whereas the Poisson time-step should be chosen so that numerical artifacts such as plasma oscillations are avoided [36]. The whole loop is iterated over a simulation time long enough for the system to evolve into steady-state, where physical observables are obtained by calculating averages over the ensemble during additional simulation time.

Traditional Ensemble Monte Carlo (EMC) simulators have the disadvantage of being inefficient or computationally expensive because the position of carriers in

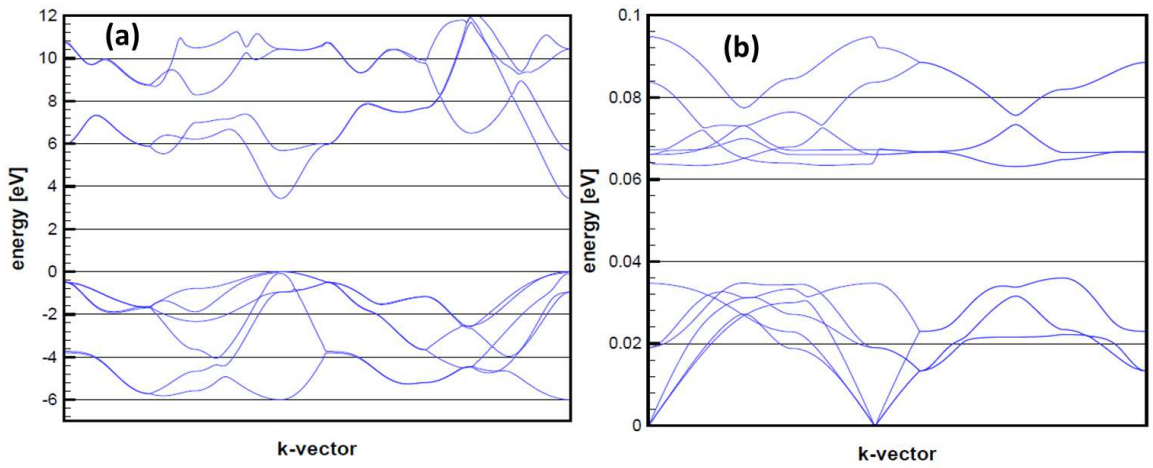

Figure 1.5: Full electronic-band structure (a) and Phonon dispersion (b) of Wurtzite GaN (HCP) used in this work.

momentum-space have to be recalculated on the fly during the simulation, for every particle in every scattering step, which requires searching for the final momentum states that conserve energy throughout the entire Brillouin Zone (BZ). Additional inaccuracies are introduced in EMC simulators because with the aim of reducing the computational burden, parabolic band structures are used in the scattering rate calculations. These issues are overcome by the Cellular Monte Carlo method (CMC) [37], where the total scattering probability rates due to all mechanisms, from every initial state to every final state are previously calculated and stored in look-up tables. This approach reduces the selection of the final state to the generation of a single random number, increasing the efficiency in the scattering step of the algorithm reducing the simulation time at the expense of higher memory requirements.

Furthermore, within the CMC framework full electronic-band structures and phonon dispersion are used in the calculation of scattering rates for each material of interest, resulting in high accuracy in the description of highly non-linear, out of equilibrium transport of carriers under high electric fields. The calculation of band structures is done with the empirical nonlocal pseudopotential method [38], whereas the phonon

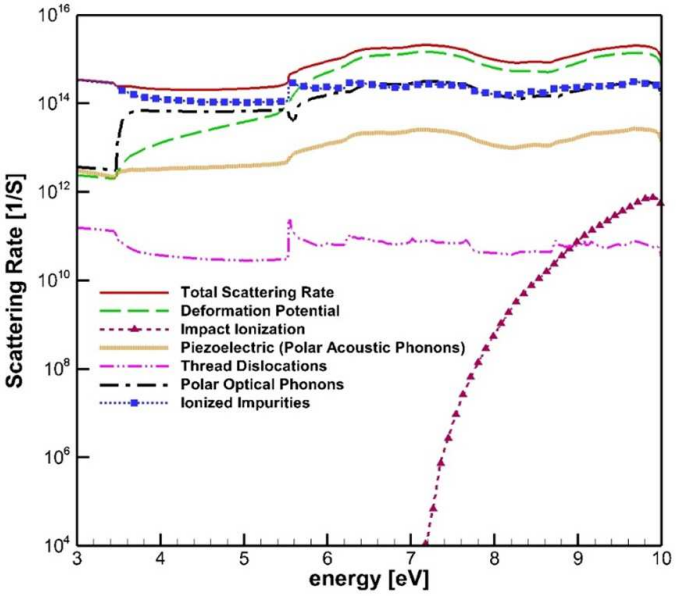

Figure 1.6: GaN scattering rates at $T = 300\,\mathrm{K}$ used in this work.

dispersion relations are computed with the volume-shell approach [39]. Figures 1.5 (a) and (b) show the calculated band structure and phonon disperssion of GaN-HCP respectively. Concerning the scattering rates, within the CMC simulator scattering processes due to deformation potential phonons, polar optical phonons, piezoelectric (polar acoustic) phonons, ionized impurities, impact ionization and thread dislocations are included in all simulations. The corresponding scattering rates for GaN are shown in figure 1.6. Finally, quantum mechanical effects due to carrier confinement in the 2DEG at the heterojunction interface, are accounted for by including the effective potential approach [40], allowing for effects such as charge setback and energy level quantization [41].

It must be emphasized that under high electric field conditions, carrier transport in semiconductors lead to significant non-linear behavior of devices, which arise from major changes in the shape of the carrier distribution function. The changes in $f\left(\vec{r},\vec{k},\vec{t}\right)$ correspond to far from equilibrium conditions which violate most of the approximations used in methods other than Monte Carlo [29]. Additionally, the

accuracy of the description of transport is directly correlated to the material properties considered in simulations. Since GaN HEMTs often operate in high electric field regime, the CMC framework constitutes an appropriate tool for modeling and characterization of these devices.

1.2 Overview of this Work

In this work, the assessment of physical mechanisms associated to reliability concerns in mm-wave GaN devices for power amplifier (PA) applications is presented. The study is performed by means of a full band Cellular Monte Carlo particle-based device simulator (CMC) [37], which allows for the systematic characterization of the performance of GaN devices operating in DC, small-signal AC and large-signal radio-frequency (RF) conditions. From the device's response under different operating regimes, the microscopic properties relevant for transport are correlated to effects such as deviations from ideal electrical response, generation of hot carriers, material defects and self-heating associated to degradation of performance.

With the aim of providing the main concepts used throughout this study, Chapter 2 is dedicated to a brief review of the material properties of wurtzite GaN, followed by the discussion of basic operation of GaN devices. In addition, the main reliability concerns associated to this technology is discussed, and the fundamentals of power amplifier operation and design are introduced. Next, chapter 3 is focused on the the study of non-idealities in III-N heterojunction diodes that explain the limited performance of the experimental IV characteristics with respect to those obtained by simulations of the ideal layout, showing that variations of mole fraction in the barrier and the presence of Schottky contacts are the dominant parameters responsible for the deviations from ideal response.

In chapter 4, a hot electron generation study in GaN HEMTs is presented in terms of the electron energy distribution function (EDF), calculated under large-signal RF power amplifier operation. While most studies are done for DC bias only, here simulations are carried out under realistic PA conditions including the effect of frequency and temperature variations. The results suggest that degradation due to hot electrons is higher in DC than in RF Class AB PAs when the frequency is in the X-band (10 GHz) independently of the temperature, but the trend is reverted for mm-wave frequencies (up to 40 GHz). Also, it is shown that Class A operation would lead to lower reliability than Class AB, because of a higher generation of hot carriers.

The methodology presented in chapter 4, is then used in chapter 5 to benchmark a new Π-shaped gate contact proposed here for improved reliability of GaN HEMTs which reduces the generation of hot electrons in DC and large-signal RF operation without significantly degrading the device's performance. The results show that in all operating regimes the EDF is lower with a Π-gate contact than with a typical T-gated device, which suggests improved robustness to failures from the generation of traps induced by hot carriers. Design rules along with projections of small-signal AC parameters were presented as guidelines for design and fabrication.

Finally, the self-heating effects observed in GaN-HEMTs on Si substrates is studied by means of electro-thermal simulations of the DC characteristics, obtained with an expanded CMC framework. In this new approach, the thermal effects are included through an energy-balance equation for phonons which self-consistently couples the charge and heat transport. After validating the electro-thermal model with experimental data showing excellent agreement, the model is used to assess the impact of lateral scaling of the device in terms of self-heating effects, showing that on scaled devices the improved electrical characteristics are accompanied by stronger self-heating effects which require the development of thermal management strategies.

Chapter 2

REVIEW OF GAN TECHNOLOGY AND POWER AMPLIFIERS

Group III-nitride semiconductors have drawn interest in research and industry in the last decades [42; 43; 18]. The wide band gap of GaN and AlN of 3.4 eV and 6.2 eV respectively, in addition to the ability to form lattice matched heterojunctions in the $AlGaN/AlN/GaN$ system, makes this a very attractive technology for applications in power electronics and mm-wave power amplifiers (PAs). These materials become intrinsic at higher temperatures than narrow band gap semiconductors and they can sustain large current and voltage due to their high breakdown field. In addition, III-nitrides heterojunctions show high density of two-dimensional electron gas (2DEG) at the heterointerface without the need of modulation doping, due to the presence of a strong polarization effect. This translates in higher electron concentration and mobilities which makes the system suitable for high power and high frequency applications [10]. In this chapter, a brief review of the fundamental concepts relevant for GaN technology is presented. First the polarization effects and principles of operation of $AlGaN/AlN/GaN$ devices is discussed, followed by a review of the main aspects of HEMTs reliability. Finally, the principles of operation and design of PAs is introduced.

2.1 Polarization Effects in GaN

In order to understand the material properties that allow high current densities and high frequency operation in $AlGaN/GaN$ devices, making GaN one of the most studied and utilized wide band gap semiconductors, it is necessary to dwell in detail into the characteristics of its crystal structure. While it is possible to grow GaN

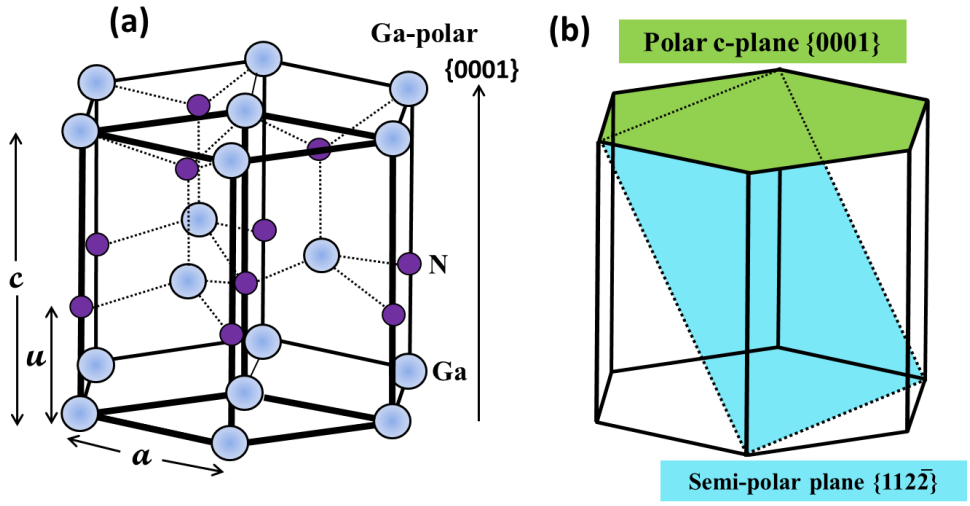

Figure 2.1: Crystal structure of GaN in Wurtzite form showing (a) Highlighted Hexagonal Closed Packed (HCP) sublattice for cation Ga, and (b) Typical growth planes, indicating the Ga-polar c-plane $\{0001\}$ and the semi-polar $\{11\bar{2}2\}$ plane.

in zincblend, the equilibrium crystal under conventional growth conditions is the hexagonal wurtzite structure, as it is shown in figure 2.1 (a). Due in part to the fact that GaN is difficult to grow in bulk form because of its high melting point of 2500 °C and the high equilibrium pressure of nitrogen, its growth is done on affordable substrates like Si, SiC and sapphire (Al_2O_3), which impose the growth direction, typically along the $\{0001\}$ plane, known as the c-axis, which offers good stability and morphology in diverse growth environments, as well as a strong spontaneous polarization effect [44].

The wurtzite crystallographic structure is composed of two interpenetrating hexagonal closed packed (HCP) sublattices, one made up of anions and the other occupied by cations, these being N and Ga respectively. The sublattice formed by cations is displaced with respect to the anion sublattice by a factor $u \cdot c$ along the c-axis, where c is the lattice parameter along the $\{0001\}$ direction and u corresponds to a reduced coordinate. If the HCP structure is not distorted, then the ratio c/a is equal to $\sqrt{8/3}$, with a being the length of the hexagonal edge of the HCP cell structure, and u =3/8.

However, in wurtzite crystals the value of u departs from the ideal 3/8. Additionally, since along the c-axis direction planes of cations alternate with planes of anions, the material has a direction of polarity with low symmetry. When crystals have a unique polarity direction, e.g. a direction without symmetry elements relating the two ends, they are pyroelectric and exhibit spontaneous polarization P_{SP} [18], which is exactly the case for wurtzite GaN grown along the polar c-plane in the $\{0001\}$ direction, as shown in figure 2.1 (b).

The direction of the P_{SP} depends on the direction of polarity chosen during growth, which for wurtzite GaN can be either Ga-polar or N-polar obtained along the directions $\{0001\}$ and $\{000\bar{1}\}$ in that order [45]. Concerning the magnitude of the spontaneous polarization, it increases as the parameter u deviates from its ideal value, because it gives rise to a nonzero dipole moment per unit volume, which is independent of external electric fields. Specifically for GaN, AlN, InN and ZnO, the reported P_{SP} values are -0.029, -0.081, -0.032 and -0.074 C/m^2 respectively [18]. For ternary alloys of the type $Al_xGa_{1-x}N$, where x is the Al mole fraction, the spontaneous polarization is proportional to x and it can be found as the linear interpolation between the GaN and AlN values [45]. Even though the polar c-plane is the preferred direction of growth for a variety of applications, in the field of optoelectronics the presence of a polarization vector is deleterious, hence other directions of growth are favored, such as the semi-polar plane $\{11\bar{2}2\}$ shown in figure 2.1 (b).

In addition to the pyroelectric nature of wurtzite GaN, pseudomorphic $AlGaN/GaN$ heterostructures also exhibit strong piezoelectric polarization effects due to strain in the epitaxial layers induced by the mismatch of the lattice constants and thermal expansion coefficients of AlN and GaN. During crystal growth, if the epitaxial layer grown on top of the buffer is kept under a critical thickness, it undergoes strain along the basal plane leading to a piezoelectric vector P_{PE} parallel to the direction

of growth, i.e. the c-axis in Ga-polar heterostructures. Following the formalism proposed by Ambacher et al. [45], the piezoelectric polarization is given by:

$$P_{PE} = 2\frac{a - a_0}{a_0}\left(e_{31} - e_{33}\frac{c_{13}}{c_{33}}\right), \tag{2.1}$$

where a_0 is the lenght of the hexagonal edge of the relaxed HCP cell, a is the lattice constants of the strained layer, e_{31} and e_{33} are the piezoelectric coefficients, while c_{13} and c_{33} are the elastic constants. The positive direction of P_{PE} is defined from the anion N towards the cation which can be either Ga or Al. Furthermore, since the inequality

$$\left(e_{31} - e_{33}\frac{c_{13}}{c_{33}}\right) < 0, \tag{2.2}$$

is valid for all mole fractions in the $AlGaN$ alloy, then the piezoelectric polarization is negative or parallel to the P_{SP} vector under tensile strain, and positive or anti-parallel to P_{SP} under compressive strain. Furthermore, the parameters of equation 2.1 are proportional to the mole fraction x of the alloy and can be found by semi-empirical linear or quadratic interpolation expressions between the GaN and AlN values [45]. Then, the total polarization is given by:

$$P = P_{SP} + P_{PE}, \tag{2.3}$$

which in the case of high polarization discontinuities across heterointerfaces, can be associated to a polarization charge density defined as:

$$\sigma_P = -\nabla P\,[C/m^2]. \tag{2.4}$$

Considering Ga-polar heterostructures in the system $AlGaN/GaN$, 3 possible cases of polarization induced charged at the heterojunction can be identified follow-

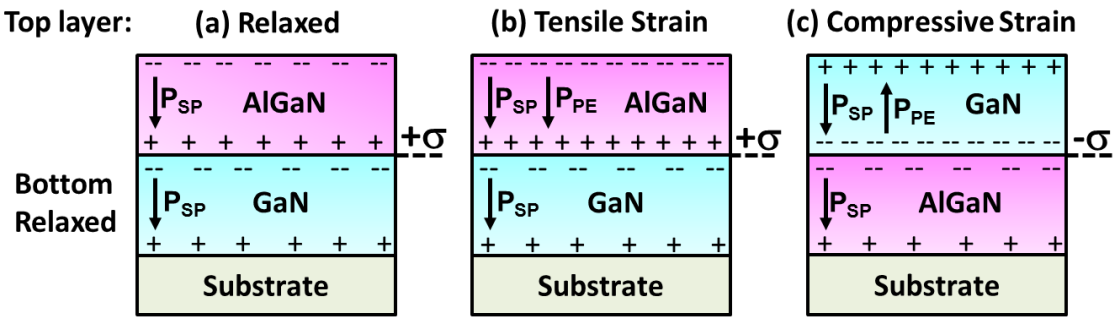

Figure 2.2: Polarization vectors (Spontaneous P_{SP} and Piezoelectric P_{PE}) and equivalent interface charge for Ga-polar $AlGaN/GaN$ heterostructures under different growth conditions of the top layer: (a) AlGaN relaxed, (b) AlGaN Tensile strain and (c) GaN Compressive strain, while keeping the bottom layer relaxed.

ing the previous description, and these are shown in figure 2.2. If a thick epitaxial $AlGaN$ layer is grown on top of a relaxed GaN buffer as shown in figure 2.2 (a), then both layers are relaxed, resulting in only P_{SP} vectors pointing towards the substrate. Since $P_{SP}|_{AlGaN} > P_{SP}|_{GaN}$ then an equivalent positive polarization charge σ is induced at the heterointerface. On the other hand, if a thin $AlGaN$ layer is grown on top of a relaxed GaN buffer, then the top layer will be under tensile strain resulting in both spontaneous and piezoelectric polarization, with parallel vectors P_{SP} and P_{PE} which will induce a strong positive polarization charge at the interface, as shown in figure 2.2 (b), significantly larger than that of case (a), because of the extra P_{PE} component. Finally, if a thin epitaxial layer of GaN is grown on top of a relaxed $AlGaN$ buffer, as shown in figure 2.2 (c), then the top GaN layer will be under compressive strain resulting in anti-parallel P_{SP} and P_{PE} vectors, which in turn induce a total negative charge density at the heterointerface.

During fabrication of heterojunctions in wurtzite GaN, in particular of the type Ga-polar $AlGaN/GaN$ under tensile stress of figure 2.2 (b), the strong positive polarization charge $+\sigma$ induced on the $AlGaN$ side at the interface is compensated by free electrons mostly provided by surface donor states [18]. These free electrons

reach high densities, creating a conductive confined layer typically referred to as a 2-Dimensional Electron Gas or 2DEG. The polarization induced 2DEG can reach concentrations higher than $10^{13}\,\mathrm{cm}^{-2}$ without the need of doping, resulting in high electron mobility >1000 $\mathrm{cm}^2\mathrm{V}^{-1}\mathrm{S}^{-1}$ due to a reduced impurity scattering [10] which can be exploited in the design of high frequency devices. Herein lies the relevance of the polarization effects observed in wurtzite GaN heterostructures, which provide a strong advantage with respect to the system $AlGaAs/GaAs$ where the 2DEG is formed by carrier injection from the $AlGaAs$ barrier which is heavily doped.

2.2 $AlGaN/AlN/GaN$ Devices

The development of technologies such as Metal Organic Chemical Vapor Deposition (MOCVD) and Molecular Beam Epitaxy (MBE), made possible the fabrication of heterojunctions in the 1970s, which was motivated by the early studies of carrier transport parallel to a heterojunction done in the late 1960s [46]. In addition, the first report on mobility enhancement in 2DEGs induced at the interface of heterojunctions was reported for the $AlGaAs/GaAs$ system in 1978 [9], which lead to developing devices based on heterojunctions. In particular, the High Electron Mobility Transistor or HEMT was first proposed in the $AlGaAs/GaAs$ system in 1980 [47], where the 2DEG at the interface is formed by electrons injected from the heavily doped $AlGaAs$ barrier. Even though research on the growth of GaN had been done since the 1960s, the lack of substrates halted the progress of this field. In fact, the first report on mobility enhancement of carriers in polarization induced 2DEGs in $AlGaN/GaN$ heterojunctions was presented in 1991 [48]. It was not until 1994 that the microwave capabilities of $AlGaN/GaN$ HEMTs was demonstrated for the first time [49]. From this point, the development of devices based on the $AlGaN/AlN/GaN$ system has skyrocketed. In the following subsections the principles of design and operation of

GaN HEMTs is discussed, followed by a brief description of the novel Hot Electron Transistors (HETs), which are related to chapter 2 of this study.

2.2.1 High Electron Mobility Transistors (HEMTs)

HEMTs based on the $AlGaN/GaN$ heterostructure are the most emblematic devices exploiting the advantages of wide band gap materials and strong polarization effects. Figure 2.3 (a) shows the basic structure of a typical GaN-HEMT, which will be used in this section to explain device performance by means of CMC simulations using the real-space grid of figure 2.3 (b). The epitaxial stack consists of a Ga-polar $Al_xGa_{1-x}N/AlN/GaN$ heterostructure, where the Al mole fraction of the AlGaN alloy is $x = 0.29$. A thin AlN interlayer, typically 1 nm to 3 nm thick, is incorporated because it provides a stronger polarization charge inducing higher concentration of carriers in the 2DEG. Furthermore, $AlGaN/GaN$ heterostructures without AlN interlayers suffer from of alloy scattering, which is caused by the perturbation of the periodic potential due to the substituting atom in the alloy crystal structure, affecting electron transport in the channel because of the partial penetration of the wave function of the carriers into the alloy layer. In this sense, adding the AlN layer increases the distance between the channel and the $AlGaN$ barrier resulting in higher mobility by reducing the effect of alloy scattering [50]. The top GaN cap layer is used to provide a high quality surface and also to reduce gate leakage. Traditionally the $AlGaN$ barrier is unintentionally doped, i.e. no impurities are added to increase the 2DEG, however in some new structures the AlGaN layer is doped with Si to increase the 2DEG concentration and therefore the current density [51]. It must be noted that since the 2DEG is intrinsically formed in the heterostructure, Ga-polar HEMTs are depletion devices. This represents no significant concern for mm-wave PAs, but it constitutes a challenge for power-switching applications.

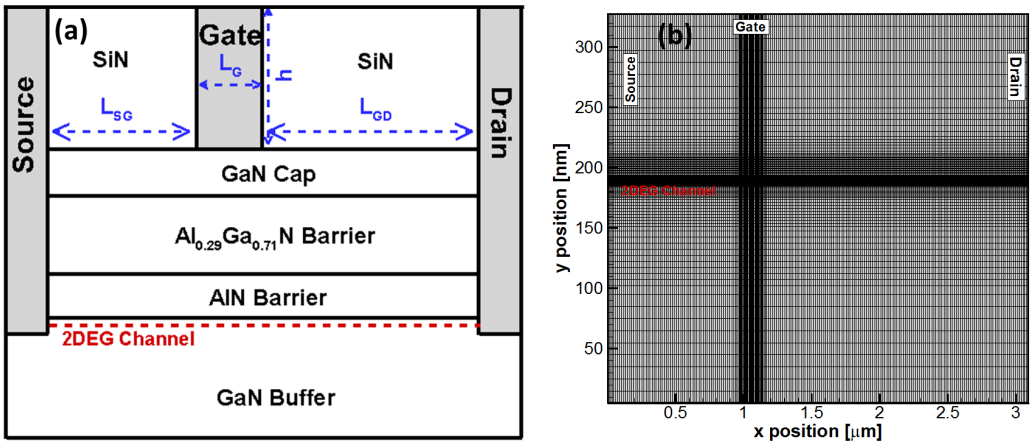

Figure 2.3: Basic structure of a Ga-polar $AlGaN/AlN/GaN$ HEMT showing (a) device layout and (b) non-uniform grid used for CMC device simulations.

The epitaxial stack is terminated with a SiN layer for passivation of the surface, which is a crucial step in HEMT design and fabrication [51]. As it was explained in section 2.1, the strong polarization charge at the heterojunction interface is compensated by free electrons that are donated from surface state donors during fabrication of the epitaxial stack forming a 2DEG. If the donor surface states are left electrically active, charge trapping /de-trapping can occur during device operation, producing variations of charge at the surface that can accumulate or deplete the 2DEG causing dispersion of the conductivity and therefore the current of the device. In addition, some surface states are acceptor-like which when empty are negatively charged limiting the concentration of the 2DEG. These effects can be eliminated by plasma treatment of the surface and by depositing dielectrics like SiN for passivation [13].

Even though it is not shown in figure 2.3 (a), because all the simulations performed in this work focus on the active region of the device, $AlGaN/GaN$ is grown either by MBE or MOCVD on thick substrates, typically 4H-SiC which provides high thermal conductivity and therefore good thermal dissipation. Also Si substrates are of interest for integration purposes. However, the lattice mismatch between these substrates and GaN require the careful design of nucleation layers in order to grow high quality

crystals. This is done by depositing AlN layers on top of the SiC or Si, before growing the wurtzite GaN [13]. The effect of the lattice mismatch is the formation of thread dislocation defects which reduce the mobility of carriers in the channel, as it is explained in the reliability section 2.3.3 of this chapter. An essential technological aspect regarding buffer design, is the compensation of n-type conductivity of the buffer by means of Fe or C doping, which increases the 2DEG confinement and reduces the short-channel effects [52].

Regarding the contacts, HEMTs are fabricated with ohmic source and drain contacts employing Ti/Al/Ni/Au or Ti/Al/Ti/Au layers, that once deposited undergo Rapid Thermal Annealing (RTA). On the other hand, the gate contact is Schottky type made using Ni/Au or Pt/Au layers [13; 51]. Due to the inherently high electric fields present in GaN HEMTs, either under equilibrium because of polarization dipoles or in DC and RF operation, the shape of the gate contact is a crucial design parameter. The state-of-the-art devices are fabricated using a T-shaped gate contact, which consists of a single stem which provides the channel length and a metal top with lateral plates which help reduce the gate resistance. In order to increase the current capability or total output current (not the intrinsic current density), the layout can include multiple gate fingers and interdigitation [51].

Modern structures for high power devices also include field plates that can be connected to the source or the gate which diffuse the peak electric field in the channel, enabling higher breakdown voltages [11]. In addition, shifting the gate placement towards the source contact i.e. making $L_{SG} < L_{GD}$ also reduces the peak electric field and allows for higher breakdown. In the case of HEMTs fabricated on conductive Si substrates, the source is ohmic-contacted to the substrate while the back-side of the device undergoes metallization to ground the buffer, which is known as source-via-grounding or SVG. The use of SVG designs reduces the on-resistance and the

source parasitic inductance by eliminating source wire bonding for packaging while also improving thermal dissipation, but more importantly it acts as a backside field-plate which enhances breakdown voltage capabilities by reducing the peak electric field in the device [53].

Now that a general picture of the complexity associated to the layout of GaN HEMTs has been discussed, it naturally follows the study of DC and small-signal AC characterization of a device. This is done here by means of CMC simulations of the device layout presented in figure 2.3 (a), which despite of following the general structure of an experimental device, it is not the intention of this section to provide a full fit to a particular experimental device but merely a functional model of a HEMT. This is done with the aim of providing a more detailed explanation of the simulation approach of this work as well as the basic operation of the device. If the reader is familiar with both of this aspects, the remaining of this section can be skipped without missing essential contributions of the work.

In all the simulations presented in this work, the main focus is on the active area of the device, meaning that detailed substrate layers are not included. The polarization effect of $AlGaN/AlN/GaN$ heterostructures is included as sheets of charge placed at the corresponding interfaces, with densities calculated following Ambacher's formalism [45] as explained in section 2.1. The effect of passivation of surface states is also included as an additional sheet of charge placed at the SiN/GaN interface, which can be used to calibrate the 2DEG concentration and fit experimental devices. Even though experimental studies to extract the polarization fields and charge values of $AlGaN/GaN$ heterojunctions have shown that Ambacher's theoretical framework lead to an overestimation of about 15% of the polarization charge [54], the accuracy of this simulation approach has been demonstrated by electron holography studies [55].

With respect to the simulation domain in real space, it is discretized with the grid shown in figure 2.3 (b) which is highly inhomogeneous in both the vertical and horizontal directions. This strategy allows to enhance the resolution in critical regions such as the $AlGaN/AlN/GaN$ barrier, the 2DEG region and the channel under the gate, while reducing the computational effort in less relevant areas, i.e. areas with no transport of carriers, such as the SiN passivation layer or deep down the buffer. The dimension of the cells have to be carefully chosen to avoid numerical artifacts such as plasma oscillations, which is done by keeping the maximum length shorter than the Debye length calculated with the local carrier concentration [36]. Also, with the aim of facilitating the convergence of the Poisson solver the variation between two adjacent cells is no larger than 2x the length in any direction [28].

The contacts of the device are modeled as equipotential regions, meaning they serve as boundary conditions for the Poisson solver. In the case of the source and drain contacts, lateral regions are defined across the barrier to directly contact the 2DEG emulating the ohmic behavior, while the doping is used to calibrate the experimental contact resistance when a full fit to experimental data is intended. Even though in the layout of figure 2.3 (a) a simple stem gate contact was used, complex structures such as T-shaped gates can be defined and simulated. In order to capture the buffer related effects such as threading dislocations and conductivity compensation, sheets of charge and doping concentration of the buffer region can be used as parameters for calibration with experiments, in particular of the threshold voltage of the device.

Figure 2.4 presents the band diagram of the simulated GaN HEMT at a DC bias condition of drain voltage V_{DS} = 5 V and gate bias V_{GS} = -5 V. In particular, the left panel identified as (a) shows the 3D conduction band profile with the carrier concentration as a contour, and it can be seen the accurate simulation of the barrier profile and the formation of the 2DEG obtained with the CMC model. Furthermore, in fig-

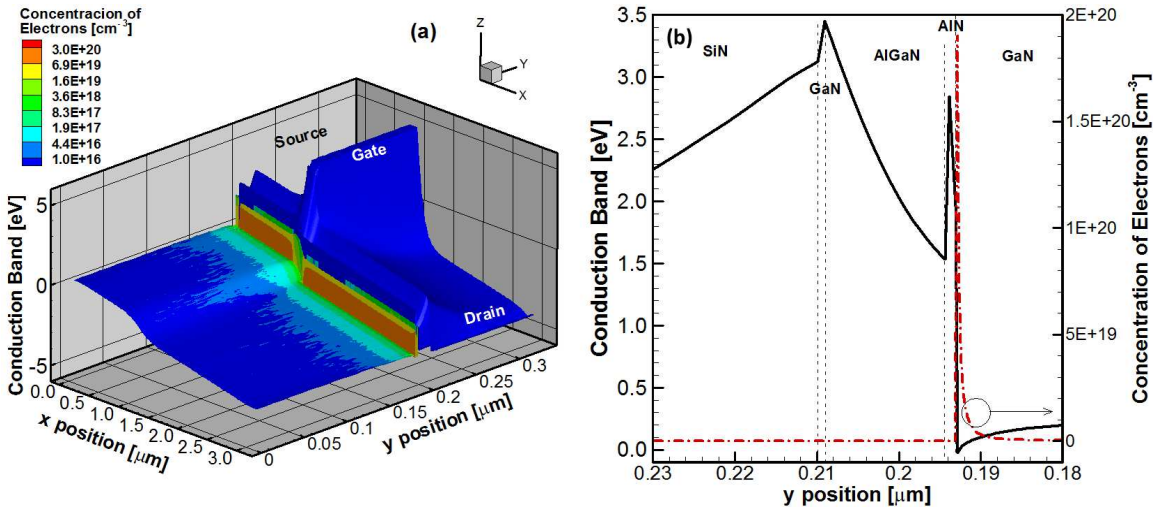

Figure 2.4: Band diagram of the simulated GaN-HEMT from figure 2.3 at DC bias of V_{DS} = 5 V and V_{GS} = -5 V. Panel (a) shows the 3D conduction band profile with the carrier concentration as contour, while panel (b) shows the vertical profile of the conduction band along with the carrier concentration, taken from a slice on the source-to-gate access region.

ure 2.4 (b) is plotted the vertical profile of the conduction band along with the carrier concentration, corresponding to a slice taken in the source-to-gate access region. In particular, the effect of adding the AlN interlayer in the $AlGaN/GaN$ barrier profile is captured, characterized by a thin barrier with height of 2.8 eV and also a high density 2DEG with a carrier concentration of 1.9x10^{20} cm^{-3}, resulting in a density in the order of 10^{13} cm^{-2}.

A full DC characterization is presented in figures 2.5 (a) and (b) corresponding to the $I_{DS}-V_{DS}$ and $I_{DS}-V_{GS}$ curves respectively, also known as the output and transfer characteristics. The first property that is observed is the depletion mode behavior of the device, which requires V_{GS} = -10 V to turn-off. This value is particularly high because no buffer compensation was used in the simulations as it would be done in an experimental device. Also, it must be noticed the high current densities achieved as well as the high transconductance G_m shown in the right y-axis of panel (b).

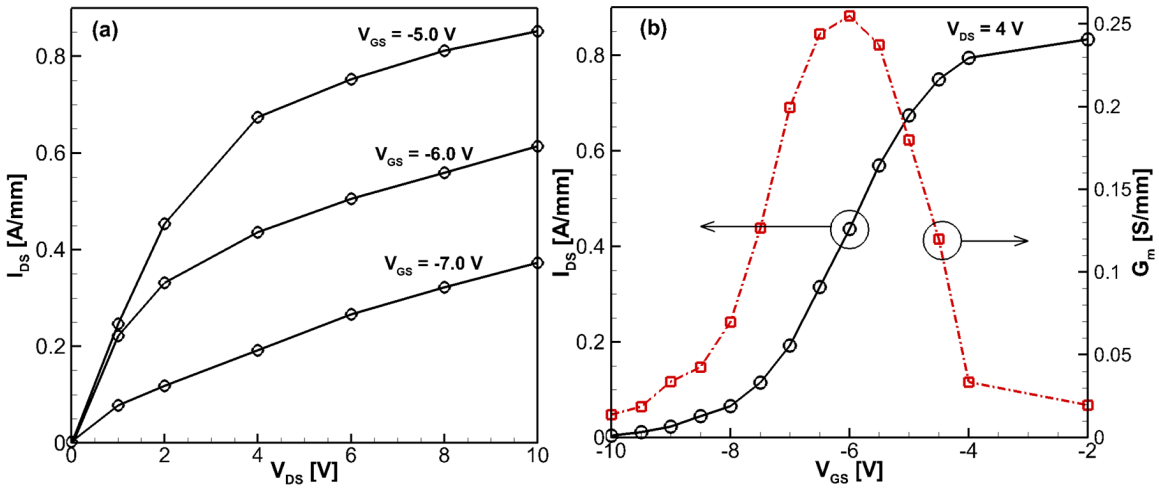

Figure 2.5: DC characterization of the simulated GaN-HEMT from figure 2.3 showing (a) $I_{DS}-V_{DS}$ and (b) $I_{DS}-V_{GS}$ or output and transfer characteristics respectively.

With respect to the AC simulation method, it consists in calculating the Y-parameters of the device over multiple frequencies. This is done by performing two separate simulations where an AC small-signal, meaning small amplitude with respect to the DC bias, is applied to either the gate or the drain of the device while keeping the other contact with only a DC bias, which emulates an AC short-circuit. The type of signal used is a multisinusoid, which consists in a linear combination of sines each centered at a desired frequency while keeping the summation of their amplitudes in the small-signal condition. Then, a fourier decomposition method is used to extract the individual contributions, and calculate the current and power gain [56].

Finally, the small-signal AC characterization of the device is presented in figure 2.6 where the current gain and unilateral power gain are plotted as a function of frequency calculated at a DC bias point corresponding to peak G_m. From this plot it is possible to extract the cut-off frequency f_t = 95 GHz, defined as the frequency for unity current gain, and the maximum oscillation frequency f_{MAX} = 123 GHz which corresponds to unity power gain. These parameters constitute the most relevant metrics in terms of frequency response, because f_t measures the maximum frequency at which the

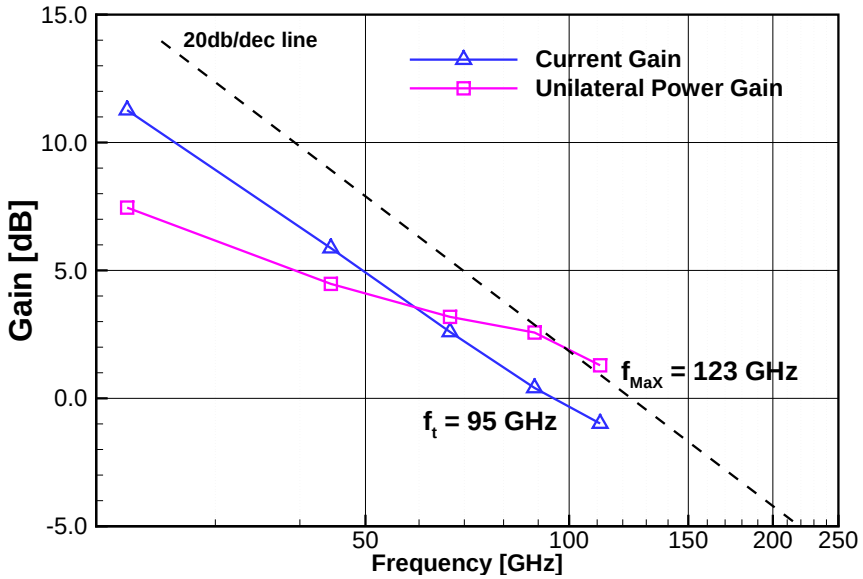

Figure 2.6: Small-signal AC characterization of the simulated GaN-HEMT from figure 2.3 showing short-circuit current gain and unilateral power gain curves as a function of frequency, for extraction of f_t and f_{MAX}.

device can be switched, whereas f_{MAX} represents the maximum frequency at which the device can provide RF power. In both cases, this particular HEMT demonstrates the high frequency capabilities of the GaN-HEMT technology.

2.2.2 Hot Electron Transistors (HETs)

In recent years there has been a resurgent interest in the development, at least at a research level, of vertical GaN devices for ultra-high frequency operation in the range of terahertz. The most prominent device in this field is the Hot Electron Transistor (HETs) based on $AlGaN/GaN$ heterojunctions, and significant progress has been made both in fabrication and transport analysis [20; 21]. The layout of a HET is shown in figure 2.7 (a). It is a 3 terminal device that can be understood as formed by two back-to-back III-N hetorojunction diodes vertically stacked. It helps to think of the HET as an unipolar version of a heterojunction bipolar transistor or HBT. The top diode is formed by an $AlGaN/GaN$ junction and it works as the emitter-base

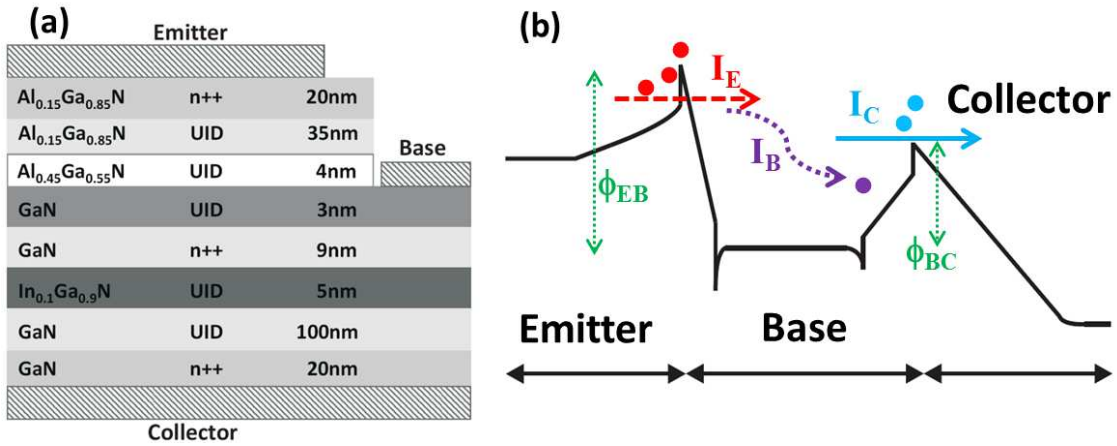

Figure 2.7: Hot Electron Transistor (HET) device showing (a) device layout and (b) 2 dimensional band diagram along the vertical direction. Also indicated in (b) a sketch of some carriers and their trajectories, along with the current components under typical DC operation.

barrier, whereas the bottom diode is made of an $InGaN/GaN$ stack and constitutes the base-collector barrier.

In order to understand the principles of operation of this device, it is convenient to examine the simulated conduction band profile of the stack shown in figure 2.7 (b), where the sketch of a few carriers and their trajectories has been added. When a positive voltage V_{BE} is applied, the emitter-base barrier ϕ_{EB} is lowered allowing electrons to be injected from the emitter into the base through tunneling and thermionic emission, creating an emitter current I_E. As the carriers cross the heterojunction interface they gain high kinetic energy becoming hot, hence the name Hot Electron Transistor, which allows them to traverse through the base. Here carriers can experiment scattering, losing energy and getting stuck in the base represented by the purple dot creating the base current I_B, or they can reach the base-collector barrier with energy higher than ϕ_{BC} as indicated by the blue dots, and cross that second barrier being collected, creating the collector current I_C. For this last process to happen the base-collector barrier must be biased with a positive V_{CB} voltage. In practice, biasing HETs is done by forcing a small I_B and applying a positive V_{CE} as done in HBTs.

In theory, the main advantage of a device as the HET here described rises from the fact that a short base, in the range of 12 nm allows for extremely short transit times enabling terahertz operation. The III-N heterostructure offers fundamental advantages for the implementation with respect to the $AlGaAs/GaAs$ system, because the polarization induced 2DEG at the top $AlGaN/GaN$ emitter-base barrier provides a highly conductive base without doping, reducing the scattering and increasing the collection of electrons in the base-collector barrier. It must be highlighted that in HETs the transport is perpendicular to the 2DEG unlike what is pursued in HEMTs. Also, the wide band gap of AlN and ternary $AlGaN$ alloys, as well as the high energy separation between the Γ and the satellite valleys in GaN, allows for the engineering of ϕ_{EB} barriers which inject high energy electrons into the base without these being transferred to satellite valleys, enhancing the transfer of carriers to the collector.

A comprehensive study of hot electron transistors is out of the scope of this research work. Nevertheless, the non-idealities of fabricated emitter-base diodes affecting HETs performance are indeed studied in detail, and the results are presented in chapter 3. The motivation for this work was the discrepancies observed between CMC simulations and experimental data of the $I_E - V_{BE}$ curves of emitter-base diodes, which are correlated to the shift observed between CMC simulations and experimental data of the $I_C - V_{CE}$ curves of HETs. In chapter 3 it is demonstrated that such discrepancies can be explained in terms of mole fraction variations in the $AlGaN$ barrier and including non-ideal Schottky contacts at the emitter and base.

2.3 Reliability of GaN HEMTs

Due to the high power and frequency performance of GaN HEMTs, the technology has become a prominent alternative in the field of power amplifiers for satellites, radars and base stations for both civilian and military applications. This applica-

tions demand for robust devices capable of operating under harsh conditions and no maintenance, making reliability requirements a crucial aspect. Operation of GaN HEMTs as power amplifiers bring about device failure mechanisms not observed in Si or other compound semiconductors [23], since under typical conditions devices are subjected to high temperature and power dissipation, high electric fields, high current and high-frequency. In addition, GaN technology development relies on heterostructures exhibiting strong polar effects, that are grown in non-native substrates such as Si, SiC or sapphire (Al_2O_3), resulting in high density of defects, along with thermal and interfaces mismatches. In this section, the main degradation mechanisms of GaN HEMTs as well as stress testing strategies are discussed.

2.3.1 Hot Electron Effects

One of the aspects that all technologies of modern semiconductor devices have in common, from Si CMOS for digital applications to GaN HEMTs for RF power amplifiers, is the fact that under typical operating conditions they are subjected to extremely high electric fields. Let's consider for example Intel's Si-FINFET for digital applications. Currently, Intel's node has a critical gate length L_G =14 nm and an operating drain voltage V_{DS} = 0.4 V resulting in an electric field of 2.8x10^5 V/cm in the active region of the device. Even though HEMTs are orders of magnitude larger with typical access region length of L_{GD} = 1 μm, they also operate at significantly larger voltages, say for example V_{DS} = 28 V yielding the same peak electric field of Intel's Si-FINFET.

Under high electric fields, devices operate far from equilibrium and their transport is highly non-linear. Furthermore, carriers traveling through the channel are accelerated by the high electric field region and may gain significant fractions of an electron volt in kinetic energy. From a transport point of view, it can be shown that in this

case the effective temperature of the carriers is well above the lattice temperature, assuming that the energy is distributed throughout the total distribution of carriers. To illustrate this, let's consider the case of two isolated energy levels in the conduction band separated by an energy Δ, so that a coupling phonon with Δ eV must be absorbed or emitted for a transition of electrons between the bands. The occupation numbers of electrons is represented by distribution functions $f(1)$ and $f(2)$, so that the latter corresponds to the upper energy band. The net rate of change in energy of electrons sitting in the upper band is given by 2.5 [29]:

$$\frac{\partial f(2)}{\partial t} = \Omega_{12} N_\Delta f(1) - \Omega_{21} N_\Delta f(2), \tag{2.5}$$

where by detailed balance $\Omega_{12} = \Omega_{21}$ and correspond to a set of constants, and N_Δ is the phonon Bose-Einstein distribution function. The rate given by 2.5 is essentially a transport equation of $f(2)$, simplified from the Boltzmaan Transport Equation (BTE) given by equation 1.1 presented in chapter 1. The first term on the right-hand-side (RHS) of 2.5 corresponds to an energy gain term, whereas the second is an energy loss, that in equilibrium balance each other leading to the Boltzmann factor for the ratio of occupancies.

When there is an applied electric field, the gain and loss terms in 2.5 are no longer equal raising the average energy of carriers. This in turn increases the density of electrons in level 2, leading to an out of equilibrium distribution of electrons $f(2)$ above its equilibrium value. Nevertheless, the system still reaches steady-state when the dacay term is greater than the gain term, so that

$$\frac{f(2)}{f(1)} > \frac{N_\Delta}{N_\Delta + 1} = exp\left(-\frac{\hbar\omega_0}{\kappa_B T}\right), \tag{2.6}$$

where T is the lattice temperature and the RHS in 2.6 is the equilibrium Boltzmann

factor. Evaluating a simple distribution function for $f(E)$ with electron temperature T_e, from 2.6 it can be seen that the only way to reach steady-state condition is that $T_e > T$, hence the term hot electrons. This analysis also suggests that the electron energy distribution function under high electric field is far from its equilibrium form, typically assumed to be Maxwellian [29].

Concerning the development of GaN HEMTs, the effect of hot electrons has a significant impact in device reliability. In simple terms, hot electrons are carriers accelerated by the electric field gaining high kinetic energy. As these carriers travel through the device the energy can be transferred to the lattice by means of collisions, producing undesired effects such as heat generation by electron-phonon interactions [57]. Since the peak electric field in devices is located in the drain-side of the gate, most of the hot-electron generation takes place on the gate-drain access region in channel and in the buffer, however the carriers do not remain confined there. If the energy is high enough i.e. greater than an activation value, hot-carriers can overcome or tunnel through potential barriers and be injected from the 2DEG channel into the AlGaN barrier or they can be injected from the gate contact into the passivation layer getting trapped there [58].

Furthermore, hot-carriers can even break atomic bonds and create interface states or activate traps through dehydrogenation [59]. The presence of electrically active traps is of high relevance because they enable mechanisms of trapping and de-trapping of electrons producing different effects in the device performance, such as gradual, permanent or recoverable degradation of parameters and increased noise due to charge fluctuations. The degree of damage depends on aspects such as material quality, fabrication process and electrical stress used to induce hot-electron generation [23].

The type of traps generated by hot electrons is correlated to the fabrication of the device, since this process requires high densities of precursor imperfections associated

to growth conditions, such as passivated point defects. By imperfections it is meant atoms other than those of the crystal structure, such as hydrogen bonds at the surface formed during passivation stages. Also, density functional theory studies have characterized the typical precursors for trap generation in the buffer and AlGaN barrier of devices fabricated by Metal-Organic Chemical Vapor Deposition (MOCVD) or Molecular Beam Epitaxy (MBE) in ammonia-rich environments [60]. While substitutional hydrogen in N antisites lead to acceptor-like traps, substitutional hydrogen of Ga vacancies leads to donor-like traps [59].

Hot-electrons that are injected from the channel into the AlGaN barrier or into the SiN/Barrier surface, are associated with generation of point defects acting as deep level acceptor-like traps, which when occupied have an effective negative charge. These traps can be located at the surface interface between the passivation layer and the barrier, or in the $AlGaN/GaN$ interface close to the channel. DC characterization of hot-electron induced trapping/de-trapping phenomena results in dispersion of parameters known as drain or gate lag. The negative charge of the occupied traps depletes the 2DEG channel resulting in higher resistance, lower drain current and lower transconductance G_m [26; 57]. Figure 2.8 reproduced here from the original research by Meneghini et al. [26], shows the signature hot-electron degradation attributed to acceptor-like trap generation in the AlGaN barrier in GaN HEMTs. In particular, left and right panels identified as (a) and (b) in the figure correspond to the I_{DS}-V_{DS} and the transconductance G_m-V_{GS} curves in that order, of an experimental device measured before and after electrical stress inducing hot electron generation.

Effects of hot electrons are also associated to generation of acceptor-like traps in the buffer and channel of the device. In this case, the negatively charge trap is close to the channel which significantly depletes the 2DEG, increasing the on-resistance or reducing the slope of the linear region of the I_{DS}-V_{DS} curve. It also increases

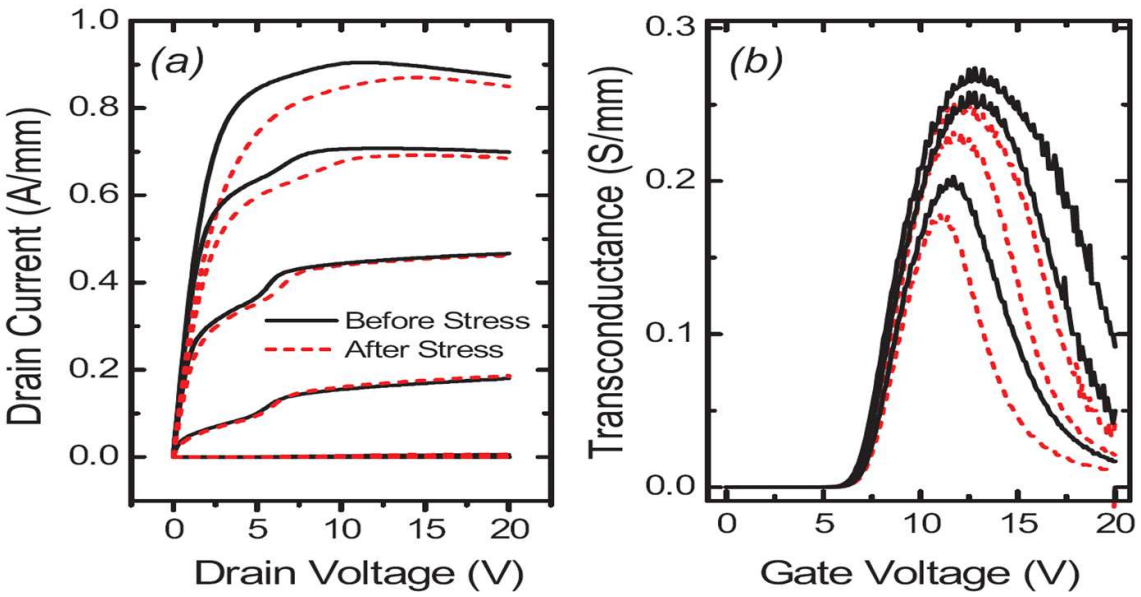

Figure 2.8: Signature degradation of (a) drain current and (b) transconductance caused by hot-electron effects in the AlGaN barrier of an experimental GaN HEMT. The figure is reproduced from original research presented by Meneghini et al. [26].

the parasitic channel resistance resulting in reduced drain current and transconductance, whereas the threshold voltage is positively shifted. In addition, the excess energy of the hot-carriers in the channel can be transferred to the lattice through electron-phonon scattering raising the temperature of the device enhancing the current collapse [22].

Another degradation mechanism corresponds to the injection of electrons from the gate contact into the barrier, which in the path of crossing the $AlGaN/GaN$ heterojunction become hot and generate donor-like deep level traps at the interface close to the channel. In this case, the gate leakage current is significantly increased due to enhanced trap-assisted tunneling. Since the unoccupied traps yield positive charge, an additional gate capacitance is added to the device that is detectable under CV measurements. Some of the deleterious effects of the excess positive charge at the $AlGaN/GaN$ interface are threshold voltage shifting towards negative values due to an increased 2DEG in the channel. It also enhances short channel effects such as

drain-induce barrier lowering (DIBL) and punch through and due to higher electric fields in the AlGaN barrier it may lead to lower breakdown [23].

The effects of hot electrons have been extensively studied under DC stress conditions, with very few experiments performed on devices driven with RF signals. Since the CMC framework is well suited to study non-equilibrium highly non-linear transport, the research work presented in the following chapters focuses in the study of hot electron generation under RF power amplifier operation of GaN HEMTs.

2.3.2 Self-heating Current Collapse

As it was mentioned in the previous section, the interaction of hot electrons with the lattice through phonon scattering leads to the significant increment of the temperature in the device. This effect, known as self-heating, is particularly strong in GaN HEMTs because one of the dominant scattering mechanisms corresponds to polar optical phonons [29], and also because non-native substrates create thermal interfaces and provide thermal conductivities that affect the thermal dissipation of the device [61]. Even though sapphire (Al_2O_3) was originally considered as a good substrate, GaN on SiC became the dominant technology due to the higher thermal conductivity and better thermal interface achieved 4H-SiC substrates, with cost as a trade-off. The signature degradation of self-heating corresponds to the monotonic reduction of drain current as the drain voltage increases, so that I_{DS} goes below the saturation current value, which is commonly referred to as current collapse. This is shown in figure 2.9, reproduced from the original work presented by Zanoni [23], which corresponds to experimental measurements of the I_{DS}-V_{DS} characteristics of a GaN HEMT with strong self-heating current collapse.

In terms of reliability studies, self-heating effects are particularly challenging because experimental techniques currently available only provide direct information of

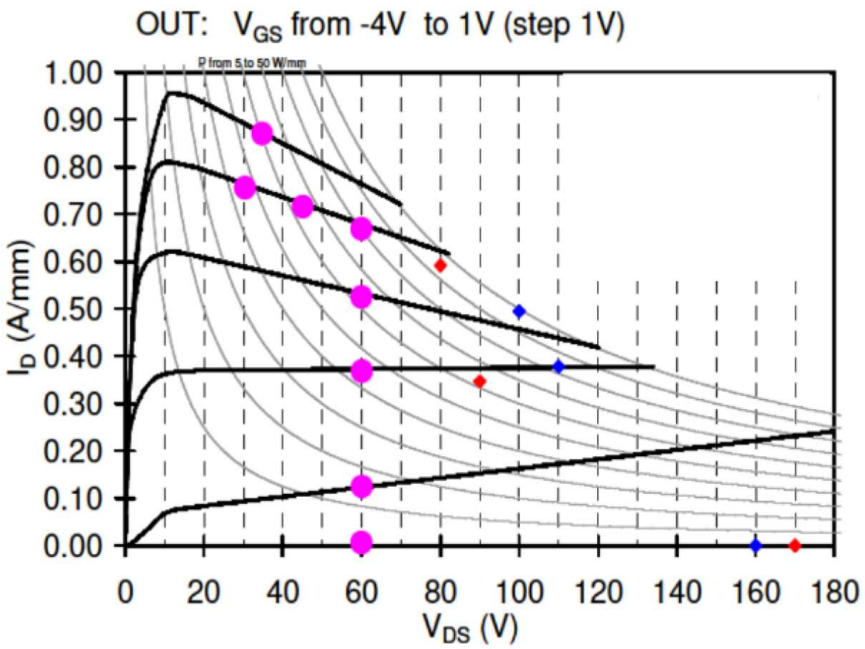

Figure 2.9: Signature degradation of self-heating current collapse observed in the I_{DS}-V_{DS} curves of an experimental GaN HEMT. The figure is reproduced from the white paper presented by Zanoni [23].

the surface or backside of the device, while knowledge of the channel temperature and spatial distribution is indirect. In fact, as of this writing there is no experimental technique with the capability of providing sub-surface device temperature mapping with the necessary lateral and depth resolution [23].

Among the experimental techniques for temperature mapping are infrared thermography, photolomuniscence spectroscopy, scanning thermal microscopy, thermoreflectance and microRaman spectroscopy. Recently, an advanced experimental set-up combining thermoreflectance measurements with Raman spectroscopy carried out from both the top on the gate metallization of the devices and from the backside of the wafers, provided a complete thermal mapping of fabricated devices with high lateral and depth resolution [62]. Despite the significant impact of these results, the validity of the sub-surface mapping of the measured channel temperature in fully processed devices is still under discussion. Furthermore, the set-up required to carry

out such experiments is significantly more complicated than other techniques that require calibration.

For these reasons, the accurate evaluation of the thermal effects in GaN HEMTs relies heavily in advanced numerical and simulation techniques, which are compared and calibrated to the available experimental data with the aim of obtaining information concerning the channel temperature, and spatial resolution of the location of hot-spots in the device [23]. For this reason, the available CMC particle-based device simulator has been expanded to include thermal effects self-consistently coupled to the highly accurate electrical simulations. This is presented in detail in the last chapter of this work, where a GaN on Si experimental HEMT is characterized in terms of electro-thermal simulations, and the modeled device is used to asses the effect of lateral scaling.

2.3.3 Threading Dislocations

During the development of GaN technology, material quality has been a major roadblock due to the lack of a native or lattice-matched substrate. This issue was overcome by developing growth techniques on substrates such as Al_2O_3, SiC and Si, which also allows to exploit the advantages of those materials such as high thermal conductivity of the former, and high integration capability of the latter.

However, GaN layers grown on foreign substrates are subjected to biaxial strain due to the large lattice mismatch of 2.5 % and 3.6 % for Al_2O_3 and SiC respectively, reaching 11.947 % for Si, in addition to large differences in the thermal expansion coefficient [63]. This in turn results in the formation of defects in the crystal structure, particularly of threading edge and screw dislocations (TDs), typically oriented parallel to the c-axis of the material due to growth conditions, with densities ranging between 10^8 to 10^{12} cm^{-2} [64].

In terms of $AlGaN/GaN$ HEMTs, the presence of TDs in the buffer or channel of the device results in degradation of electrical performance and affects long-term reliability [65], because TDs behave as electrically active traps with positive or negative charge acting as centers of Coulomb scattering. This effect can be modeled as a 2D scattering rate, that affects carriers moving perpendicular to a vertical line of charges due to the presence of TDs [66]. This model has been implemented withing the CMC framework as a thread dislocation scattering rate shown in figure 1.6 in chapter 1, calculated using the approach developed by Weinmann et al. [55; 67].

The effect of TDs on device performance is illustrated in figures 2.10 (a) and (b), corresponding to the transfer and output characteristics of the simulated $AlGaN/AlN/GaN$ HEMT device of section 2.2. The simulations were performed at three different dislocation densities N_{TDs}, distributed in the buffer and channel of the device with a filling factor of 0.5. From the I_{DS}-V_{GS} curves calculated at $V_{DS} = 4\,\mathrm{V}$, it can be seen that as the concentration of dislocations increases the drain current is significantly degraded going from $0.8\,\mathrm{A/mm}$ to $0.2\,\mathrm{A/mm}$ when N_{TDs} goes from $10^9\,\mathrm{cm}^{-2}$ to $10^{12}\,\mathrm{cm}^{-2}$, which are typical values for SiC and Si substrates respectively. The transconductance peak, not shown in the figure but calculated as $\partial I_{DS}/\partial V_{GS}$, is also degraded from $0.2\,\mathrm{S/mm}$ to $0.05\,\mathrm{S/mm}$, in turn degrading the frequency response of the device. The output characteristic I_{DS}-V_{DS} at $V_{GS} = 0\,\mathrm{V}$, shown in figure 2.10 (b), also exhibits the same behavior, where not only the saturation current degrades from $1.75\,\mathrm{A/mm}$ to $0.6\,\mathrm{A/mm}$, but also the turn-on slope and knee voltage are degraded as N_{TDs} increases. Throughout this work, the effect of TDs is included in all the simulations, with defect densities calibrated to available experimental information. This is particularly important for HEMTs fabricated on Si substrates.

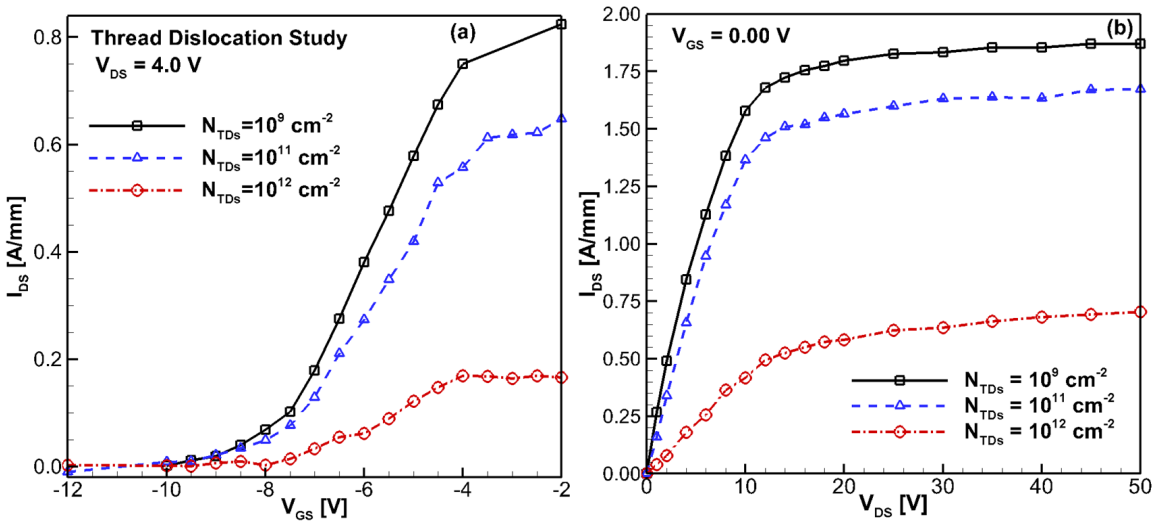

Figure 2.10: Effect of thread dislocations (TDs) in the electrical performance of GaN HEMTs, shown with simulations of (a) I_{DS}-V_{GS} and (b) I_{DS}-V_{DS} curves at three different TDs density N_{TDs}.

2.3.4 Inverse Piezoelectric effect

HEMTs based on the $AlGaN/GaN$ system, with or without AlN interlayers, are intrinsically subjected to high electric fields. As it was discussed in section 2.1, the lack of symmetry of the crystal structure in GaN and $AlGaN$ produce high built-in electric fields associated to a strong spontaneous polarization. Furthermore, the lattice mismatch between $AlGaN$ and GaN produce heterostructures with layers under tensile stress, inducing a strong piezoelectric effect responsible for the conductive 2DEG at the heterojunction's interface. Under high reverse bias voltage operation, so that the potential $V_{DG} = V_{DS} - V_{GS}$ is large, the potential is dropped in a very short region on the drain-side of the gate producing high electric fields across the barrier. Since the materials are strongly piezoelectric, the external field induces large levels of mechanical stress in the crystal structure, which adds to the intrinsic elastic energy stored due the tensile strain of the heterojunction. If V_{DG} exceeds a critical value, the mechanical stress induced by the inverse piezoelectric effect produce crystallographic defects, which are electrically active and degrade the performance of the device.

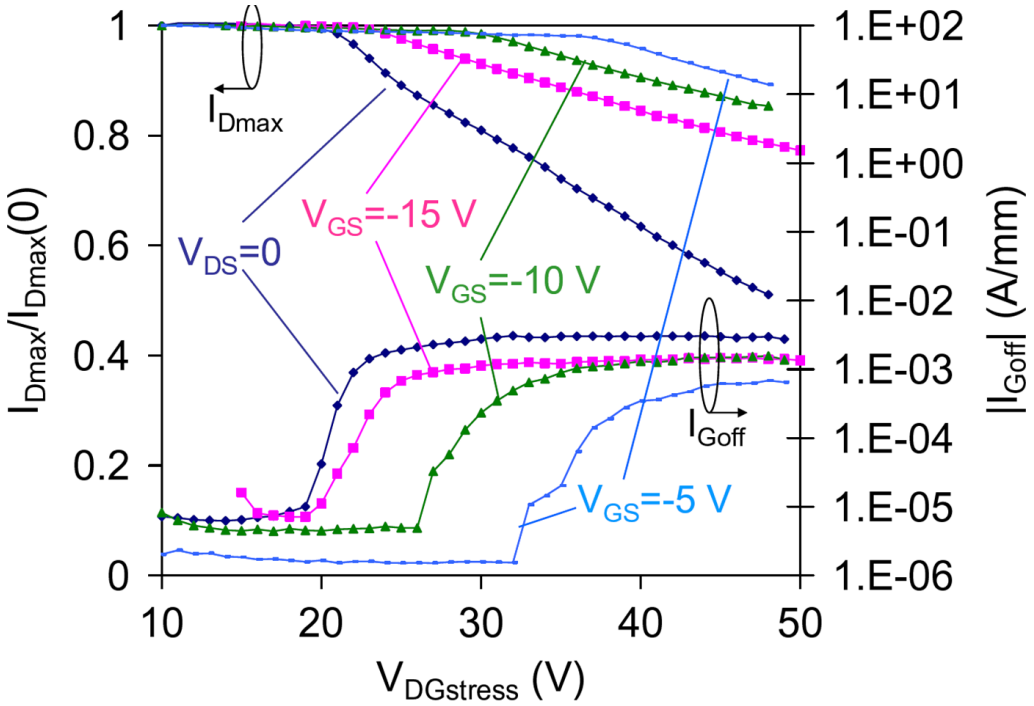

Figure 2.11: Signature degradation of drain current (left axis) and gate leakage current (right axis) in an experimental GaN HEMT caused by the inverse piezoelectric effect. Figure reproduced from original research presented by del Alamo and Joh [69].

This mechanism was first proposed by del Alamo and Joh [68], as a theory to explain the signature degradation of GaN HEMTs observed under electrical stress consisting of high electric fields and no current, i.e. applying either high V_{DG} or V_{GS} keeping the device off. The degradation of the device consists in the abrupt increase of the gate leakage current I_{Goff}, and the monotonic reduction of the maximum drain current I_{Dmax}, as shown in figure 2.11 reproduced from the work presented by del Alamo and Joh [69]. Their experiment consists in keeping the device off by applying a high negative gate voltage V_{GS}, and then stepping the voltage applied to the drain in turn increasing V_{DG}. It can be seen that for all values of V_{GS} there is a critical V_{DG} from which the signature degradation is observed. These results could not be explained in terms of hot electron induced traps, since they were independent of current and horizontal electric fields. The inverse piezoelectric effect is essentially an

electro-mechanical degradation mechanism, whose effects are well understood. For this reason, it will not be given further consideration in this work.

2.3.5 Electrical-Stress Testing Conditions

The semiconductor industry relies on the ability to manufacture highly reliable devices rated in terms of their useful life or mean time to failure (MTTF), which for industry and commercial standards is in the order of 25 years or $2\text{x}10^5$ h [70]. Since testing devices for several years is impractical, thorough protocols known as long-term accelerated testing have been developed by the Si industry, which consist of operating the devices under increased temperature conditions for a few thousand hours while measuring a given set of parameters to determine performance. Then, the data is used to extrapolate life time and extract the MTTF. Among these tests the High Temperature Revers Bias (HTRB) is one widely used for power and RF systems, where devices are operated under DC reverse bias while the environment temperature is increased to achieve a junction temperature typically of 125 °C [51].

Even though the accelerated testing protocols are common for Si devices, the validity of its results in GaN technology is a matter of discussion. Following industry standard protocols such as HTRB yields unrealistically high MTTF estimates [51]. The reason for this is that many of the degradation mechanisms for GaN devices are not accelerated by temperature, and in order to trigger them it is necessary to apply specific bias conditions that can be either DC or AC [22]. This is why new protocols focused on electrical stress instead of or in addition to increased temperature have been developed, and are mainly designed as short-term life-time testing limited to less than 100 h [23].

The electrical stress conditions of interest for GaN HEMTs are shown in figure 2.12 (a) where 4 bias points have been highlighted on the I_{DS}-V_{DS} plane of the

simulated $AlGaN/AlN/GaN$ HEMT introduced in section 2.2. The four basic conditions are known as On-State, High-Power State, Semi-On State and Off-State each of which enhances or minimizes a particular degradation mechanism. Figures 2.12 (b) to (d) show the 3D band diagrams of the simulated HEMT with the x-component of the electric field as a contour, under biasing conditions corresponding to the Off, On and High-Power states respectively. Exploiting the capabilities of the CMC simulator, the figures also include a snap-shot of the electron distribution in real space represented as spheres, whose colors correspond to the carriers energy. It must be noticed that the 0 eV energy in the simulations is set to the top of the valence band, hence thermalized electrons located at the bottom of the conduction band have an energy equal to the band gap of the material, in this case 3.41 eV for GaN.

The Off-state condition, shown in figure 2.12 (b) for V_{GS} =-20 V and V_{DS} =50 V, is characterized for presenting high electric field both in the channel and the barrier while keeping low drain current, typically within leakage range. The degradation in this condition is mainly due to the creation of traps induced by the inverse piezoelectric effect. Performance degradation is typically smaller as compared to the high-power case, and it has been observed to be reversible [69]. Even though the channel is depleted of electrons and the current is very small, from the figure it can be seen that the few electrons that reach the high-electric field region gain high kinetic energy becoming hot, which allows the study of hot-electron effects and impact ionization. However, there are other bias points where the concentration of hot electrons is higher.

The On-State corresponds to a point where the current is high, therefore a high gate voltage is required, but the lateral electric field is kept low by applying low drain voltage [22]. Figure 2.12 (c) shows the band diagram for this point simulated at V_{GS} =0 V and V_{DS} =12 V. It can be seen that the field throughout the device is low with values around -3.5x10^4 V/cm with a hot-spot of -3.1x10^5 V/cm located on the

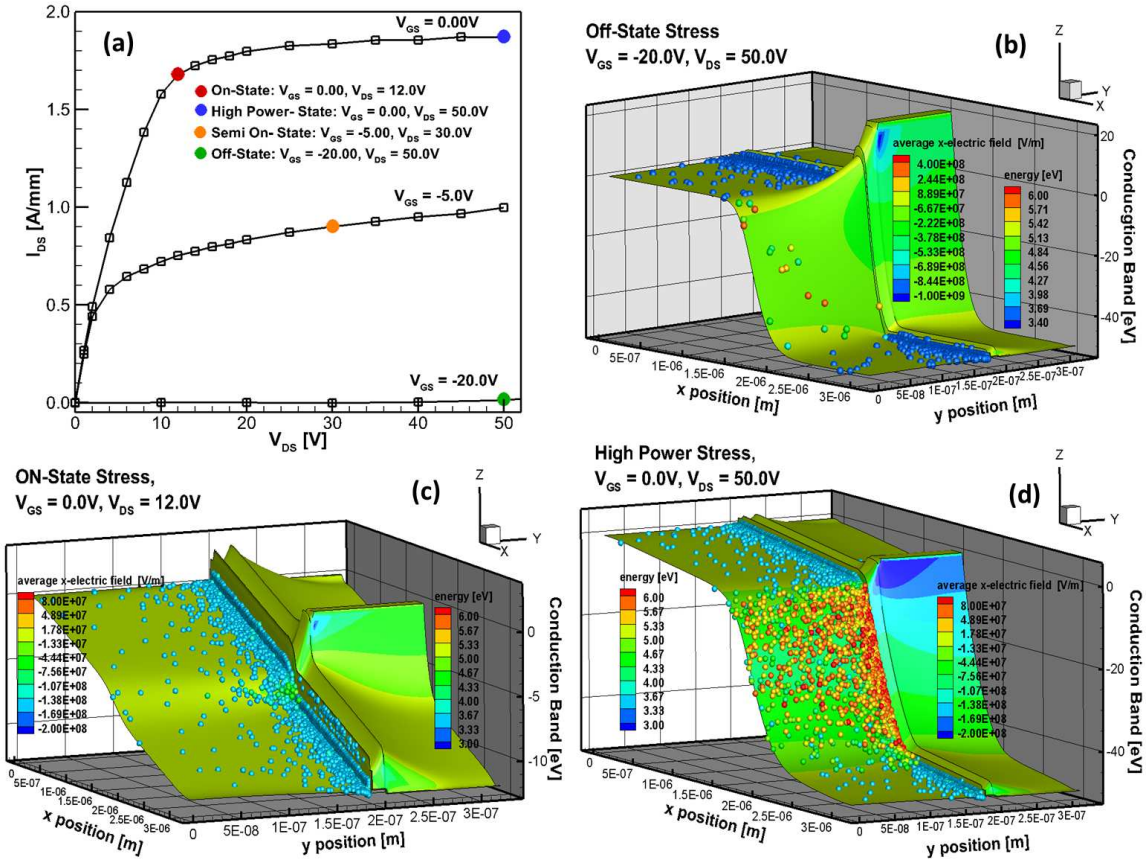

Figure 2.12: Electrical stress characterization of GaN HEMTs, showing (a) I_{DS}-V_{DS} curves indicating stress bias points. Also, in (b) to (d) the electric field mapping is presented on 3D band diagrams of the device for 3 stress bias points. The spheres represent a snap-shot of real-space distribution of carriers with its corresponding energy as a contour.

drain-side of the gate. In addition, most of the electrons are thermalized and mainly confined to the channel with a small concentration in the buffer. Nevertheless, there is a small population of carriers in the channel with energies up to 5 eV (1.6 eV above the conduction band) located on the drain-side of the gate, which corresponds to the peak electric field. For this reasons, the On-state stress has been used to study hot-electron degradation with electroluminescence experiments [26], however other operating conditions with higher electric fields have proven to be more appropriate. In addition, self-heating and current crowding effects are also studied in the On-state stress point.

Figure 2.12 (d) corresponds to the High-Power state simulated at V_{GS} =0 V and V_{DS} =50 V, so that both drain current and lateral electric field are high. Since hot-electron generation depends exponentially on the electric field and linearly on the current [69], in this case it can be seen that there is a high concentration of hot carriers whose energy reach 2.4eV above the conduction band (6 eV in the figure). These electrons are located both in the channel and the buffer in the drain-side of the gate, where the electric field reaches its peak value of -1.2x10^6 V/cm. Experimantal studies have shown that permanent device degradation is induced after stress in these operating condition, and it has been attributed to the presence of hot-electrons and self-heating effects [22].

It must be noticed that under High-Power State the dissipated DC power reaches high values, up to 95 W/mm in the simulation, which is not desirable or even feasible experimentally because it can cause thermal damage before any short or long term degradation phenomena is observed. This limitation is overcome by stressing devices in the Semi-On State, indicated in figure 2.12 (a) for V_{GS} =-5 V and V_{DS} =30 V, which can be understood as an intermediate case between the High-Power and Off-State cases. The main advantage of this set-up is that high currents and high electric fields are simultaneously present while the total DC power dissipated is non-prohibitive. Furthermore, several studies dedicated to hot-electron effects have demonstrated that this state corresponds to the peak generation of hot carriers [26; 59].

Concerning test protocols, short-term testing is typically done on-wafer and the acceleration is achieved by implementing stepped biasing strategies to electrically stress the device, in addition to constant DC bias. There are 3 main testing protocols where the voltage or current can be increased taking the device into the desired state. These are shown in figure 2.13, and correspond to (a) step-stress, (b) step-stress recovery and (c) stress-recovery cyclic stress loading. These techniques are

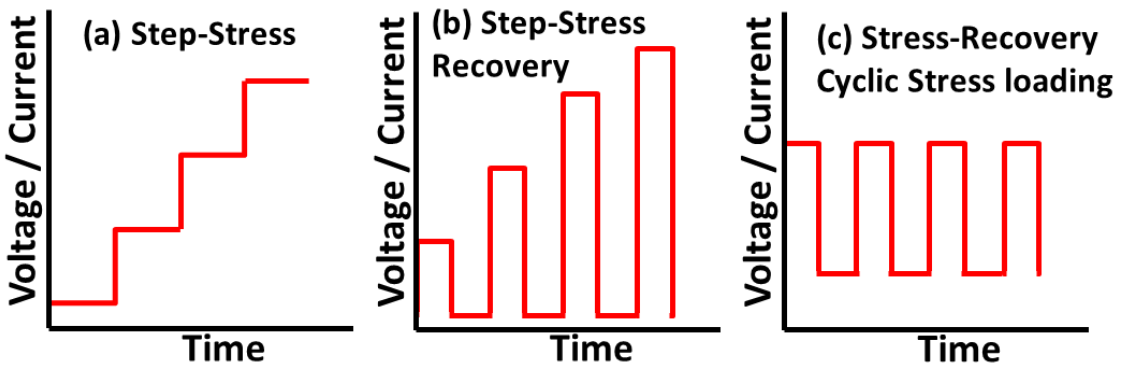

Figure 2.13: Voltage or Current signal waveforms used in short-term life-time reliability testing, identifying the 3 main protocols (a) Stepped-Stress, (b) Step-Stress Recovery and (c) Stress-Recovery Cyclic Stress loading.

used instead of increasing the temperature and have proven to yield more realistic estimates of reliability under relevant operating conditions. In particular, recovery cycle protocols are relevant since they lead to current and voltage transients enabling the study of deep traps, already present in the device or created during the test, which produce characteristic waveforms such as exponential recovery, increased noise or gate leakage [22].

The stepped-stress protocols are crucial during all stages of technology development of GaN HEMTs, since they provide in-depth information regarding the physical mechanisms of degradation in a short time, without necessarily inducing catastrophic failure. In terms of industry and commercially available products, short-term testing is typically followed by conventional long-term life-time testing [70]. Finally, it must be highlighted that nearly all the effort for reliability testing has been developed around DC biasing, and more protocols are required in order to include relevant operating conditions such as RF-stress, particularly important for mm-wave power amplifier operation. In this regard, a few studies have focused on hot-electron generation under RF operation by means of electroluminescence measurements [71; 72], however a deeper understanding is still lacking and this is why a major effort of this

work is precisely directed to further explain the transport and physical mechanisms of hot-carrier generation of GaN HEMTs operating under RF condition.

2.4 Radio Frequency Power Amplifiers

A Power Amplifier (PA) is a circuit that transforms a low energy radio frequency (RF) signal into a large signal with significantly higher power, by converting DC input power into RF/microwave output power. In contrast to typical small-signal amplifiers, PAs are intrinsically non-linear and they operate under large voltage and current excursions over much of the device's $I_{DS}(V_{GS}\text{-}V_{DS})$ space. This requires a non-linear analysis approach as well as the full consideration of non-linear effects in simulations for its proper circuit design. The range of applications for PAs is varied going from wireless communications to RF heating and imaging, translating into a wide variety of frequencies from very low frequency (VLF) to millimeter-wave (mm-wave) operation, with output powers varying from 10 mW in short-range wireless systems to 1 MW in long-range broadcast transmitters [73].

The basic circuit typically used in wireless transmitter applications is with a single active device in a common source topology [74] as shown in figure 2.14. The active device can be a FET or a BJT, but for all the analysis presented in this work only FETs will be considered. For correct operation a PA requires an Input Matching Network stage (IMN) which transforms the impedance of the RF power source into the input impedance of the transistor. IMNs are designed to provide maximum power transfer into the active device and it is done by conjugate matching, which can be implemented with lumped circuit elements (inductors and capacitors) or with transmission line stubs for high frequency mm-wave operation. The input network also plays an important role affecting the PA's stability.

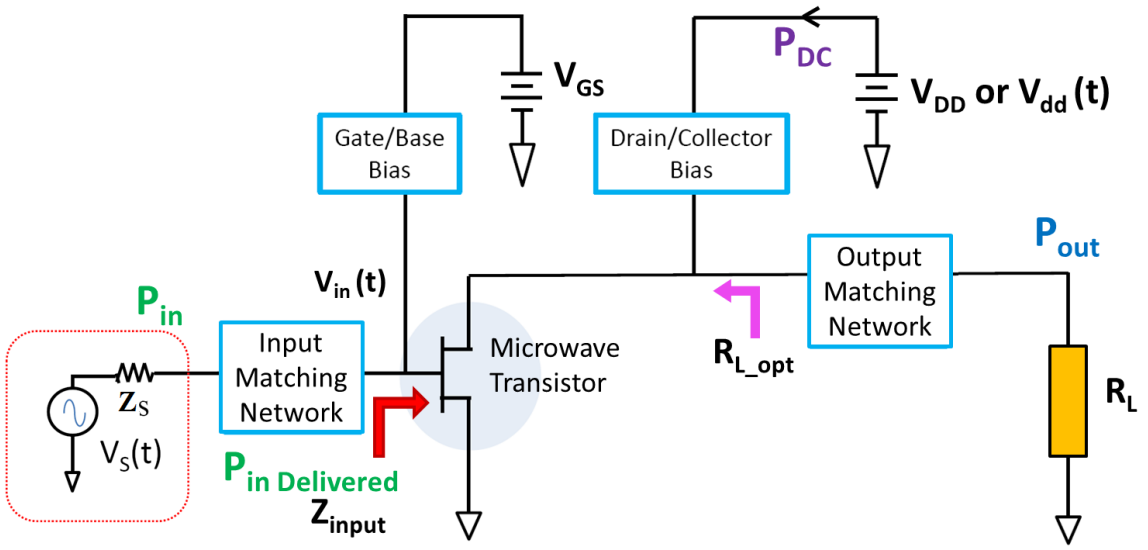

Figure 2.14: Schematic of Power Amplifier (PAs) circuits.

The quality of the matching provided by IMN determines the input reflection coefficient Γ given by 2.7 as:

$$\Gamma = \frac{Z_{input} - Z_S}{Z_{input} + Z_S} = \sqrt{\frac{P_{in\_ref}}{P_{in}}}, \tag{2.7}$$

where Z_{input} is the input impedance of the transistor and Z_S is the impedance of the AC source. The reflection coefficient is sometimes given in dB as the input Return Loss $RL = 20 \cdot log(\Gamma)$. The value of Γ is a complex number which provides information on the amplitude and phase of the reflected wave, hence it can also be defined as the ratio of the power of the reflected wave P_{in_ref} to the input power at the source P_{in}, as indicated in the RHS of 2.7 [75]. Furthermore, the power delivered to the device is then given by 2.8:

$$P_{in\_Del} = \Gamma^2 \cdot P_{in}. \tag{2.8}$$

The gate bias circuit is essential to achieve the desired PA class of operation since it establishes the DC gate voltage V_{GS} used as DC operating point for the device, also referred to as the quiescent operating point. The effect of the lumped or distributed components used in this stage must be included when studying amplifier stability. In a similar way, the drain bias circuit plays a crucial role in providing DC power P_{DC} to the active device. This network must behave as a stiff source so that there are no variations of drain voltage V_{DD} in time or with input signal level. Also, it should present a high impedance at RF frequencies and a short circuit at DC. For these reasons, the drain bias circuit is typically implemented with a large choke inductor, which can also be used to facilitate the output matching of the device. This rules are true except for the case of supply modulated PA architectures where the DC bias varies over time with signal $V_{dd}(t)$ which will not be considered in this study.

Concerning the Output Matching Network (OMN) its goal is to transform the load impedance R_L, generally $50\,\Omega$ or an antenna, to the optimum impedance to the transistor output R_{Lopt}. It must be highlighted that for PAs the value of R_{Lopt} is not designed for a conjugate match as would be done for a small-signal linear amplifier [74]. The OMN in most cases also includes a high-Q network to provide the desired impedance to the active device at harmonic frequencies, which can be done by shunt or series LC resonators or transmission lines. This design is known as tuned-load and it is crucial for correct PA operation under a given topology. The OMN also affects the stability of the amplifier.

The key attributes of a power amplifier are power, gain, conversion efficiency and linearity, and in every design or topology there is a compromise between these attributes. In order to characterize the performance of a given topology, certain metrics or figures of merit (FOM) are defined for each of the attributes.

Of high interest is the power delivered to the load at the fundamental frequency P_{out} given by:

$$P_{out} = \frac{V_{out(rms)}^2}{R_L}, \tag{2.9}$$

where $V_{out(rms)}$ is the rms value of the output voltage and R_L is the load impedance, typically $50\,\Omega$. In the case of modulated RF signals, the average output power must be used, defined as:

$$P_{out\_Avg} = \frac{1}{T}\int_0^T p(t)dt, \tag{2.10}$$

where $p(t)$ is the instantaneous power. Concerning the DC power it is define as $P_{DC} = I_{DD} \cdot V_{DD}$, where I_{DD} and V_{DD} correspond to the DC current and voltage, determine by the quiescent bias point although for some topologies I_{DD} can vary with the input signal. If the architecture of the PA requires supply modulation then the average DC power must be considered.

With respect to gain in PAs, the term gain in general refers to transducer gain G_T, which corresponds to the the small-signal gain S_{21}, and it is defined as:

$$G_T\,[dB] = 20 \cdot log\left(\frac{P_{out}\ [W]}{P_{in}\ [W]}\right) = P_{out}\ [dBm] - P_{in}\ [dBm], \tag{2.11}$$

where P_{in} corresponds to the total available input power at the source. If input mismatch effects are to be included, then the power gain G_p is used, given by:

$$G_P\,[dB] = 20 \cdot log\left(\frac{P_{out}\ [W]}{P_{in\_Del}\ [W]}\right) = P_{out}\ [dBm] - P_{in\_Del}\ [dBm], \tag{2.12}$$

where P_{in_Del} is the input power delivered to the active device, defined by equation 2.8. This figure of merit is common in loadpull measurements as it accounts for mismatch effects.

The efficiency is a metric that is crucial for PA characterization, because it relates the conversion of DC power to RF power. A widely used FOM is the drain efficiency η which does not include the effect of bias currents and is defined as:

$$\eta\,[\%] = \frac{P_{out}\;[W]}{P_{DC}\;[W]}\,100. \tag{2.13}$$

Another common metric is the Power Added Efficiency (PAE), which unlike the drain efficiency considers the amplifier's gain, and is given by:

$$PAE\,[\%] = \frac{P_{out}\;[W] - P_{in}\;[W]}{P_{DC}\;[W]}\,100. \tag{2.14}$$

An important aspect about the efficiency of PAs is the trade-off inherently associated with linearity. For high linearity, the simplest approach consists in using a high quiescent bias point so that V_{DD} and I_{DD} are high enough to allow the full excursion of the output voltage and current signals without clipping or distortion. However, this approach leads to high P_{DC} which in turn reduces both FOM for efficiency, particularly for low values of P_{out}. In some topologies this issue is overcome by either allowing the dynamic change of P_{DC} or by choosing a bias point with low P_{DC}, at the expense of lower linearity.

The sketch of the figures of merit for typical PA operation are shown in figure 2.15, where the P_{out} in dBm, G_T in dB and PAE in % are shown as a function of P_{in} also in dBm. It can be seen that whereas P_{out} increases linearly with P_{in} up to a certain point, the gain remains constant throughout the same range. However, as the non-linear distortion becomes dominant there is a reduction in the slope of P_{out} at the fundamental frequency and a monotonic reduction of gain, because at this point part of the power is transferred to higher order harmonics. This is known as gain or power compression, and a relevant metric for it is the P_{1dB} defined as the point where P_{out}

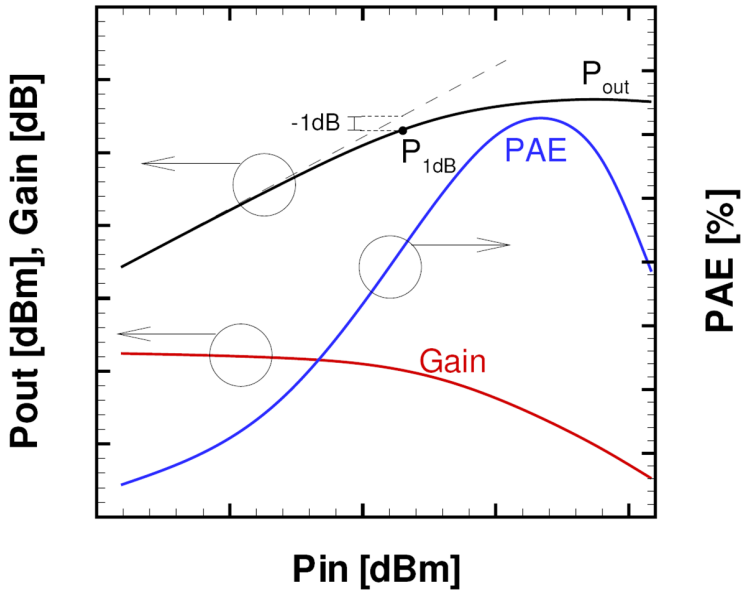

Figure 2.15: Sketch of figures of Merit for a typical power amplifier for single-tone input, showing output power P_{out} and transducer gain G_T on the left-axis, and power added efficiency PAE in the right-axis as a function of the input power P_{in}. Also indicated the 1dB compression point P_{1dB}.

is 1 dB below the value it would have extrapolating the linear response. The PAE increases monotonically in a non-linear fashion as P_{in} becomes larger, because the ratio of RF power to DC power increases, until it reaches a peak efficiency point after the device has reached compression, to then drastically drop. Even though this are general trends, each PA topology will have a signature behavior for the FOMs.

The effect of non-linearities give rise to high order harmonics, i.e. signal components located at multiples of the fundamental frequency, each one with specific consequences in the performance of the PA. Even order terms produce self-biasing effects through harmonics located close to DC or 0 frequency in the spectrum of the output signal. Odd orders such as 3^{rd}, 5^{th} and higher create amplitude/amplitude or amplitude/phase effects such as non-linear and non-monotonic gain. Traditionally, a distortion analysis used for linearity is the Inter Modulation Distortion (IMD) which is characterized by the surge of spectral content at undesired frequencies close to the

fundamental that did not exist originally, when the input is a two-tone signal (two sinusoids) at frequencies $\omega 1$ and $\omega 2$, non multiples of each other, closely spaced by a frequency $\delta = \omega 2 - \omega 1$. The cause for the spurious spectrum under two-tone input testing is the non-linear mixing of the input signals.

In particular, for a two-tone signal with fundamentals located at frequencies $\omega 1$ and $\omega 2$ the undesired spectral components that appear too close to the desired signals that cannot be filtered are termed 3^{rd} order and 5^{th} order IM products or IM3 and IM5 respectively. These components are located at higher and lower frequencies than the fundamentals. A full analytical derivation can show that the exact location is given by $f_{3rd\_h} = 2\omega 1 - \omega 2$, $f_{3rd\_l} = 2\omega 2 - \omega 1$, $f_{5th\_h} = 3\omega 1 - 2\omega 2$ and $f_{5th\_l} = 3\omega 2 - 2\omega 1$ corresponding to high and low terms of IM3 and IM5 products in that order [74]. The typical spectrum of the output signal of a PA under IMD analysis is sketched in figure 2.16 (b).

From IMD tests several figures of merit can be derived. One of them is the suppression S_3 or S_5 of IM3 and IM5 spectral components respectively relative to the desired signal, defined as the ratio of their power in watts $P_{(\omega 1,\omega 2)}/P_{(f\_IM3,f\_IM5)}$ or their difference in dBm. Another FOM widely used is the intercept point where the IM spectral components would theoretically be equal in power to the fundamental components. The concept behind it relies on the fact that as P_{in} increases, the power transferred to the spurious components IM3 and IM5 increases more rapidly than that at the fundamentals $\omega 1$ and $\omega 2$. Therefore, with increasing P_{in} an intercept point is reached and the corresponding input and output power at which this condition is achieved constitutes the FOM. The formal definition is as follows: the intercept point of IM3 products is given by coordinates IIP3 and OIP3 for input and ouput power respectively, and in a similar way IIP5 and OIP5 for IM5 products.

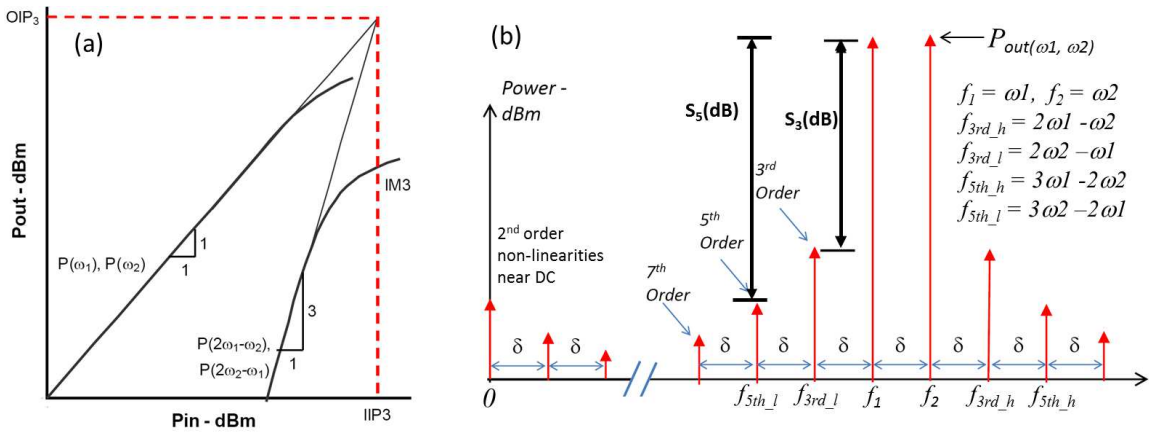

Figure 2.16: Sketch of Inter Modulation Distortion (IMD) measurements for a power amplifier (PA) with a two-tone input signal separated by a small frequency $\delta = f_{\omega 2} - f_{\omega 1}$. Determining IIP3 and OIP3 can be done from measuring (a) output power delivered to fundamental and harmonic frequencies vs input power sweep, and (b) Output power spectrum of the PA for a given input.

Determining the exact coordinates (IIP3,OIP3) and (IIP5,OIP5) can be done in a couple different ways. The first consists in sweeping the input power of the PA and plotting P_{out} as a function of P_{in} measured at the fundamentals and also at the corresponding IM3 or IM5 components, which results in in two curves whose slopes will have a 3:1 or 5:1 ratio for IM3 and IM5 respectively. Then, by drawing the linear extrapolation of both curves the intercept point can be found and the direct reading of the value of P_{in} and P_{out} yields the (IIP3,OIP3) or (IIP5,OIP5) coordinates. This is demonstrated in figure 2.16 (a) for the IM3 components.

Another way which is more convenient, both for experiments and simulations, consists in calculating the spectrum of the output signal for a single high value of P_{in} as shown in figure 2.16 (b). Then, identifying the IM3 and IM5 products allows to calculate the IM suppression relative to carrier as:

$$S_3\,[dB] = P_{out}|_{\omega 1,\omega 2}\,[dBm] - P_{out}|_{(2\omega 1-\omega 2),(2\omega 2-\omega 1)}\,[dBm], \tag{2.15}$$

$$S_5\,[dB] = P_{out}|_{\omega1,\omega2}\,[dBm] - P_{out}|_{(3\omega1-2\omega2),(3\omega2-\omega1)}\,[dBm], \tag{2.16}$$

It must be noticed that the measurement of the suppression can be done between either of the input signals $\omega1$ or $\omega2$ and either of the high or low IM3, IM5 components. With the values determined with equations 2.15 and 2.16, then it is possible to calculate the values of OIP3 and OIP5 with the relations:

$$OIP3\,[dBm] = P_{out}|_{\omega1,\omega2}\,[dBm] + \frac{S_3\ [dB]}{2}, \tag{2.17}$$

$$OIP5\,[dBm] = P_{out}|_{\omega1,\omega2}\,[dBm] + \frac{S_5\ [dB]}{4}, \tag{2.18}$$

then the values of the input coordinates can be determined as $IIP3 = OIP3 - G$ and $IIP5 = OIP5 - G$ for IM3 and IM5 respectively where G is the gain.

Even though IMD studies are convenient and widely used, they present important limitations for characterization of linearity in modern systems where the modulation of the input signals is done simultaneously in amplitude, frequency and phase typical with digital modulation schemes. In this case, the state-of-the-art technique is the Adjacent Channel Power Ratio ACPR or Adjacent Channel Leakage Ratio ACLR [73].

In this test the input of the PA is a digitally modulated signal upconverted to a desired RF carrier frequency, producing a band limited input. The PA distortions create new spectral content within adjacent bands called regrowth. Then by determining the spectrum of the output signal, a ratio between the power in the adjacent bands to the power in the main channel can be calculated. Furthermore, the spectrum is usually compared to a mask defined by a standard body such as the FCC.

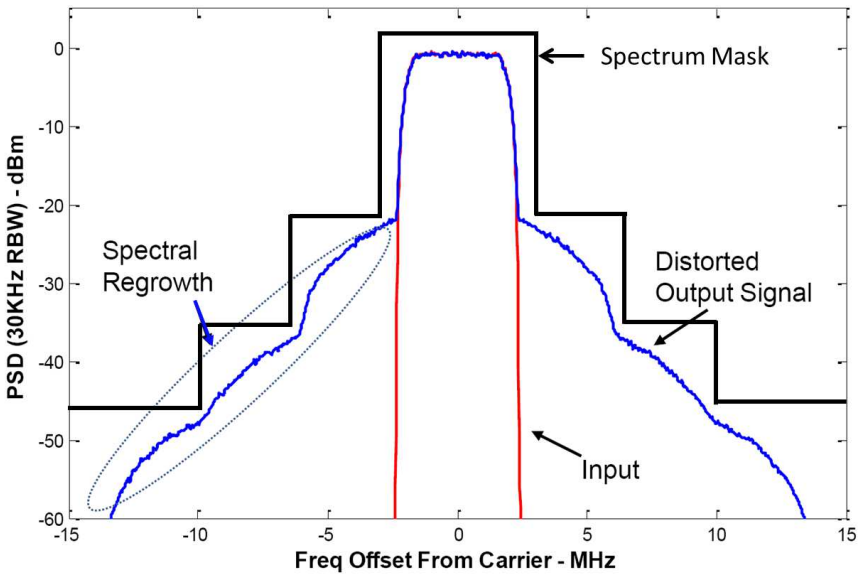

Figure 2.17: Example of the output spectrum of a PA with a single carrier WCDMA input signal band limited to 5 MHz, evidencing the spectral regrowth used in Adjacent Channel Power Ratio (ACPR) measurments. A typical spectrum mask is also included.

An example of this method is shown in figure 2.17 where the output spectrum corresponding to a single carrier WCDMA input signal band limited to 5 MHz is plotted along with a protocol mask. It can be seen that non-linearities of the PA cause spectral regrowth which translates into distortions of the the signal, nevertheless for this particular case the output spectrum fits below the mask, implying that the PA meets protocol specifications.

Concerning the PA topologies, these are defined as Classes named from A to F each of which have different advantages and characteristic behavior in terms of method of operation, efficiency and output power capability [76]. The definition of the topology can be given in terms of the conduction angle θ, which is defined as the period of the waveform where the device is conducting, and it can take values between $0 < 2\theta < 2\pi$. Besides the conduction angle, a well designed PA will present signature shapes of load-lines for each Class, which are nothing but the representation of the output waveforms in the I_{DS}-V_{DS} space, e.g. the ac output current $i_{ds}(\omega)$ is

plotted as a function of the ac output voltage $v_{ds}(\omega)$. The load-lines play a crucial role in the correct operation of each Class of PA, being also very useful to identify the Class as well as the presence of mismatched parasitic components. Since all the relevant definitions for power amplifier characterization have been discussed, in the next subsections a brief description of two of the main PA topologies is presented. The chosen topologies are Class A and Class AB because these are amongst the most used in telecommunications, particularly with $AlGaN/GaN$ HEMT devices, and they will be used for characterization of reliability in the next chapters of this work.

2.4.1 Class A Operation

The main characteristic of a PA in Class A configuration is that the transistor operates in the active region at all times, or in other words $\theta = 2\pi$. In this sense, the device is acting as a linear current source controlled by the gate bias and it is not operated as a switch [73; 74]. This is achieved by selecting a quiescent or large-signal operating point (LSOP) which corresponds to a DC value of V_{GS} well above the threshold voltage, near the center of the most linear portion of the transconductance, resulting in a drain current I_{DD} roughly half of the maximum saturation current of the device. On the other hand, the DC drain voltage V_{DD} is chosen to be half way between the knee voltage V_{Knee} and the maximum voltage usually taken to be the device breakdown V_{BR}, allowing for the full excursion of the output signals without clipping even at peak power.

A typical implementation of a Class A power amplifier is shown in figure 2.18, which corresponds to the schematic circuit used for DC and large-signal simulations performed with the commercial software Advanced Design System (ADS) by Keysight. From the figure it can be seen that the OMN is a simple LC circuit which transforms the load impedance into the optimum impedance presented at the output

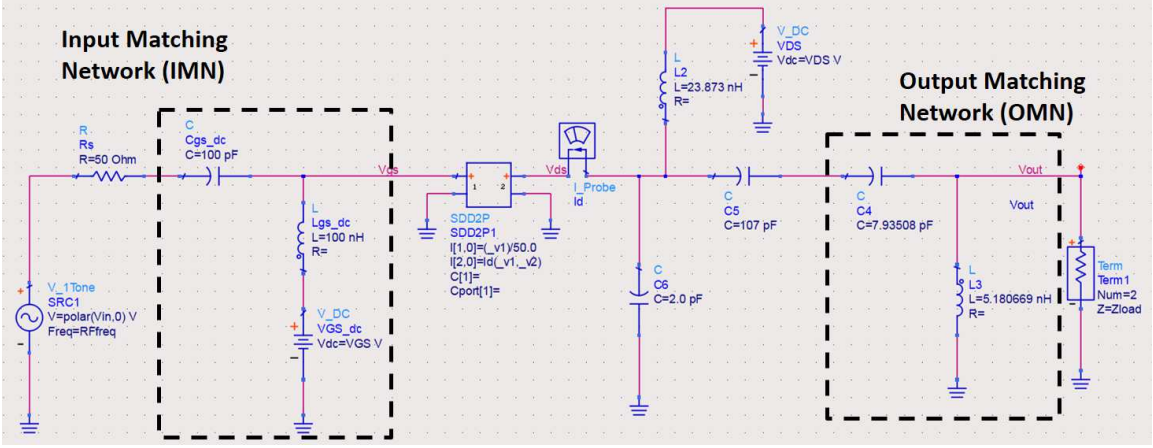

Figure 2.18: Schematic of a Class A power amplifier circuit topology, simulated with commercial software Advanced Design System (ADS) by Keysight.The active device is simulated by using the analytical model by Statz, and the input and output matching networks are identified.

of the device, which in this case for correct PA operation at $f_{req} = 1\,\text{GHz}$ corresponds to $R_{Lopt} = 15\,\Omega$. The active component was described by the analytical compact model of microwave transistors proposed by Statz [77], and the output capacitance C_{DS} was included as lumped shunt capacitor $C_6 = 2\,\text{pF}$ shown in the figure. An IMN was designed to minimize the reflected power of the input signal, also serving the purpose of gate bias circuit. A choke inductor and a series output capacitor were used as drain bias circuit.

The DC simulation of the transistor is shown in figure 2.19 (a), presenting the typical I_{DS}-V_{DS} space for a device with constant transconductance G_m evidenced by the fixed step in increments of drain current for increments of V_{GS}. The large-signal characterization was performed by means of a harmonic-balance simulation, and the RF output drain current $i_{ds}(\omega)$ and drain voltage $v_{ds}(\omega)$ waveforms are plotted in figures 2.19 (b) and (d) in that order, whereas the typical figures of merit are shown in 2.19 (c). First, it can be seen that the OMN effectively transformed the load into a purely resistive $15\,\Omega$ impedance, evidenced by the closed, straight shape of the load-lines also included in figure 2.19 (a).

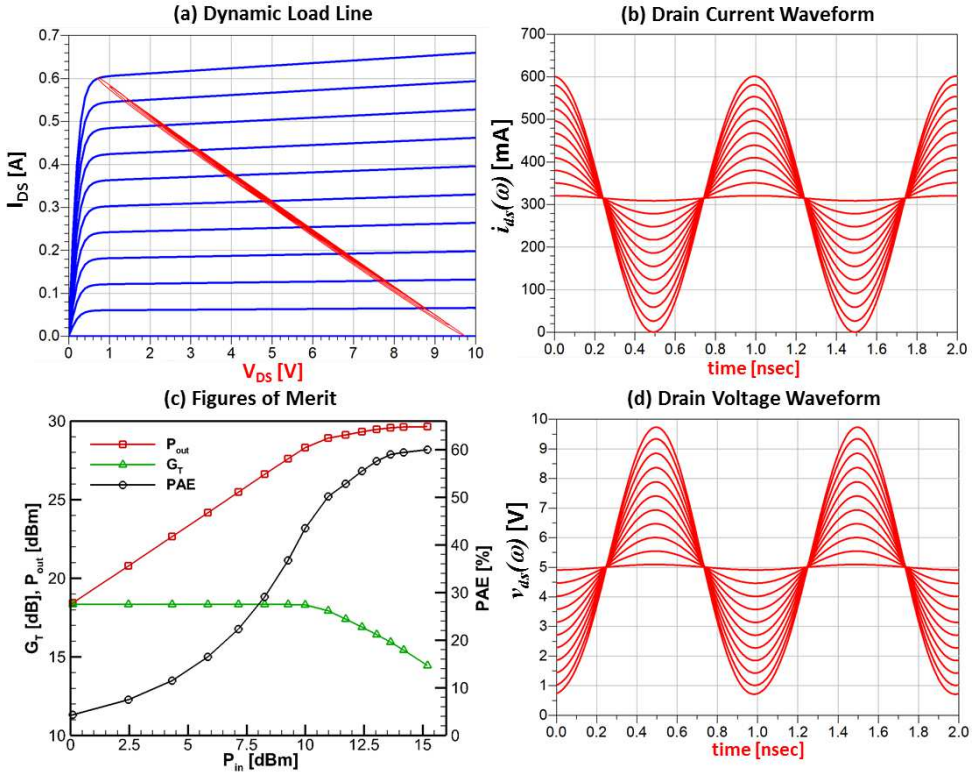

Figure 2.19: Characterization of a Class A power amplifier, obtained with simulations of the schematic shown in figure 2.18. In the panels is presented (a) I_{DS}-V_{DS} space along with load-lines, (b) ac current waveform $i_{ds}(\omega)$, (d) ac drain voltage waveform $v_{ds}(\omega)$ and (c) typical figures of merit.

Concerning the waveforms, both $i_{ds}(\omega)$ and $v_{ds}(\omega)$ are full sinusoids throughout all the input powers considered. The FOMs show the signature behavior of Class A operation: a linear P_{out} vs P_{in} relation at constant G_T from $P_{in} = 0\,\text{dBm}$ to 11 dBm which corresponds to the full power excursion before the PA is pushed into compression. The PAE begins at very low values due to the high DC power required to bias the device, reaching its maximum theoretical value of 50 % before compression, with a maximum peak output power of 29 dBm at 11 dBm of input power. For $P_{in} > 11\,\text{dBm}$, the PA is operated in overdrive, where clipping occurs in both current and voltage waveforms with the advantage of increasing the P_{out} delivered to the fundamental, also allowing for higher theoretical PAE of up to 60 % [73; 74].

In summary, Class A PA operation offers some inherent advantages. First and foremost it is perhaps the most simple topology to design in terms of IMN and OMN requirements. Also, it offers high linearity because the device is operated near the center of the most linear portion of G_m and the output waveforms $i_{ds}(\omega)$ and $v_{ds}(\omega)$ have full excursion without clipping for a large range of $P_{out} - P_{in}$. However, the main disadvantages are high thermal dissipation and low conversion efficiency due to the high DC bias point. For this reasons, Class A operation is rarely used in mobile applications but is still a basic building block for applications such as output stage in base stations transmitters or high fidelity amplifiers where portability is not a requirement [73; 74].

2.4.2 Class AB Operation

In order to improve the efficiency of a PA, an effective strategy consists in reducing the conduction angle so that the transistor is off for some portion of the load-line excursion which reduces P_{DC}, at the expense of increasing the higher order harmonics resulting in lower linearity of the circuit. This is the principle of operation of PA topologies such as Class B achieved with $\theta = \pi$ and Class C with $0 < \theta < \pi$ for which the maximum theoretical efficiency can be up to 78.5 % and 100 % respectively, with the trade-off of having non-linear gain and significantly lower output power than Class A topologies. Nevertheless, there is a sweet spot for maximum efficiency while preserving linearity and output power, and it corresponds to deep Class AB operation for which $\pi < \theta < 2\pi$ which is enabled by the subthreshold soft turn-on of real transistors [74].

Under Class AB the quiescent bias point of the device is selected so that V_{GS} is close to but still higher than the threshold voltage, which sets the drain current between 0 and 1/2 of the maximum saturation value, while V_{DD} is chosen to be half

way between V_{Knee} and V_{BR} as it is done fore Class A. By choosing this bias point, P_{DC} is kept low making this topology inherently more efficient than Class A PAs. In addition, the efficiency is further improved because the transistor only operates in the active region for less than half of the total excursion. Figure 2.20 presents the circuit of a typical Class AB PA topology implemented in ADS for DC and large-signal simulations, where the active device is the same used for Class A characterization of section 2.4.1.

In terms of circuit complexity, the IMN and gate bias circuits remain constant with respect to the ones used for Class A operation. The drain bias circuit is again implemented with a choke inductor and a series capacitor with updated values. However, the main difference and added complexity is found in the OMN which can be understood as a 2-stage circuit. First, the OMN includes a simple LC circuit for impedance transformation, followed by a high-Q shunt LC resonator which is tuned to the frequency of operation of 1 GHZ. The shunt resonator is necessary to provide a purely resistive load at the fundamental frequency, while providing a short-circuit impedance (0 Ω) for higher order harmonics [75], which is known as a tuned-load configuration and its performance is directly proportional to the Q of the circuit. This OMN is implemented to minimize the output power deviated from the fundamental to higher harmonics resulting in higher linearity. The main difficulty for this design is to implement the shunt LC resonator with high-Q because the lumped elements become harder to fabricate as the Q of the circuit increases, which is necessary for correct tuning of the harmonics. If the PA was to be operated at higher frequencies, then quarter-wave transmission line stubs would be more appropriate to implement the OMN.

The load-lines obtained with a harmonic-balance simulation are shown in figure 2.21 (a), on top of the corresponding I_{DS}-V_{DS} space, whereas the ac waveforms

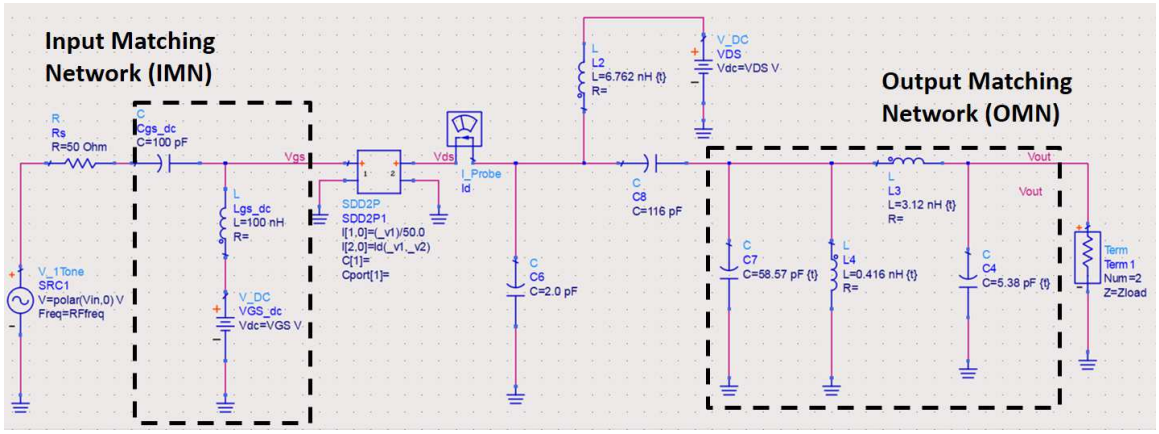

Figure 2.20: Schematic of a Class AB power amplifier circuit topology, simulated with commercial software Advanced Design System (ADS) by Keysight.The active device is simulated by using the analytical model by Statz, and the input and output matching networks are identified.

are shown in figures 2.21 (b) and (d) corresponding to $i_{ds}(\omega)$ and $v_{ds}(\omega)$ in that order. The output current waveform is a half-rectified sinusoid with a duty cycle slightly larger than half the period, whereas the output voltage exhibits the full sinusoid excursion, highlighting the fact the there is no clipping for the P_{in} range. Furthermore, neither of the waveforms present harmonic distortion, indicating excellent tuning of the higher order harmonics to $0\,\Omega$ impedance provided by the shunt LC resonator. In addition, the load-lines are nearly a perfect straight line indicating good matching to a purely resistive load. These results corroborate that the device is in fact operating in the active region for slightly more than half of the full excursion.

From the FOMs shown in 2.21 (c), the main characteristics of Class AB operation can be discussed. From $P_{in} = 0\,\text{dBm}$ to $11\,\text{dBm}$ G_T decreases monotonically and it is lower than that of Class A operation. Also, a reduction P_{out} is observed for the same range as compared to Class A. However, the PAE is higher under Class AB than under Class A throughout the entire range due to a lower LSOP and lower P_{DC}. When the device is driven into compression for $P_{in} > 11\,\text{dBm}$, P_{out} reaches $29\,\text{dBm}$ as it did under Class A, with the advantage of higher PAE larger than $60\,\%$.

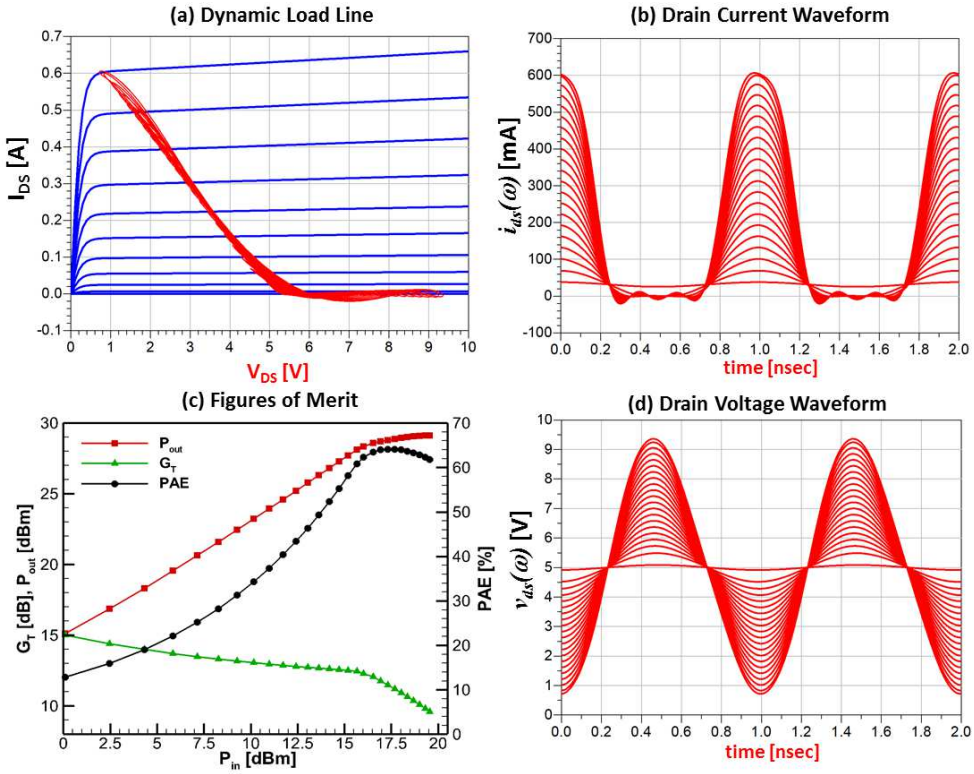

Figure 2.21: Characterization of a power amplifier in Class AB topology with tuned-load. The Results correspond to simulations of the schematic from figure 2.20, showing (a) I_{DS}-V_{DS} space along with load-lines, (b) ac current waveform $i_{ds}(\omega)$, (d) ac drain voltage waveform $v_{ds}(\omega)$ and (c) typical figures of merit.

To summarize, Class AB power amplifiers constitute an excellent topology for high efficiency operation while preserving high output power and linearity. The higher efficiency, with respect to Class A topology, arises from the inherently lower P_{DC} required to bias the device, which is controlled by the conduction angle. For these reasons, Class AB is by far the more extensively utilized PA configuration in battery powered and mobile communications, and also as a fundamental building block in other PA architectures such as supply modulated PAs, envelope tracking and Doherty PAs [74]. The main disadvantages are the relatively lower linearity with respect to Class A operation, and the more stringent requirements for the OMN circuitry [73].

2.5 Conclusions

A brief review of the main concepts concerning GaN technology and devices was presented in this chapter. The objective was to provide the necessary context and definitions necessary to fully comprehend the scope and contributions of this work, as well as ease the understanding of the analysis presented in the next chapters. First, the polarization effect that gives rise to the high current densities and high mobilities of GaN devices was studied. Then the principles of design and operation of $AlGaN/AlN/GaN$ HEMTs were discussed by means of CMC simulations of a typical Ga-polar HEMT, while also providing details of the CMC simulation framework. Then a short description of HET devices, necessary to understand the contribution of chapter 3 was introduced. This was followed by a review of the main aspects of HEMTs reliability, focusing on those relevant to this work such as hot electron effects, self-heating and electrical-stress operating conditions. Finally, the principles of operation and design of PAs were presented, making emphasis in the Class A and Class AB operation used in the subsequent chapters.

Chapter 3

STUDY OF NON-IDEALITIES IN III-N ISOTYPE HETERJUNCTION DIODES

In this chapter, a simulation study of the non-ideal effects in polarization-dipole induced isotype heterojunction diodes, proposed as the emitter-base junction for Hot Electron Transistors (HETs), is presented. First, simulations of the IV characteristic of the diode under forward bias are performed with the Cellular Monte Carlo particle-based device simulator (CMC), using the experimental layout of the ideal structure, showing that the simulated IV curves significantly differ from the reported experimental data. The CMC results are then validated by means of hydrodynamic simulations performed with the commercial software Sentarus by Synopsys (TCAD), showing excellent agreement with the CMC curves, suggesting that non-ideal effects must be included. A further analysis is then carried out with hydrodynamic simulations only, where physically meaningful non-ideal effects are proposed and studied. The results show that when variations in the mole fraction of the $Al_xGa_{1-x}N/GaN$ barrier and the effect of Schottky contacts are included in the model, the experimental data is successfully reproduced. In the context of the development of HETs technology, this work provides valuable information about the factors limiting the device's performance.

3.1 Introduction

Isotype heterojunctions in the $AlGaN/GaN$ system are a basic building block for several applications, such as High Electron Mobility Transistors (HEMTs), ultraviolet photodetectors, gas sensors, high-voltage rectifiers, varactors [78] and, more recently, HETs which are being considered for high power, high frequency applications [20;

21]. HETs were originally proposed a few decades ago and, due to difficulties in the fabrication of high quality crystals of wide band-gap materials in those early stages, very few successful attempts of HETs have been reported. Nevertheless, the $AlGaAs/GaAs$ system enabled the fabrication of hot-electron launching structures traditionally used in the study of transport properties of high-energy carriers.

Shur [79] was the first to propose the $AlGaN/GaN$ heterostructure to overcome issues found in the $AlGaAs/GaAs$ system which hamper the development of HETs such as a limited range of injection energies, small barrier height difference between emiter-base and base-collector junctions and increased impurity scattering. The large band gap of AlN and GaN, as well as the large ΔE_c=1.8eV between them, allows engineering adequate emitter-base (EB) and base-collector (BC) barriers. Additionally, the strong polarization effect in this system provides a highly conductive base without doping which in theory enables ballistic transport of carriers and therefore high-frequency operation. Furthermore, the injection energy range is also increased due to the higher intervalley separation (1.3eV for GaN as compared to 0.3eV for $GaAs$), allowing the injection of electrons with higher energies before the transfer to satellite valleys becomes relevant and limits the current gain. The first common base characteristic of HETs fabricated in the $AlGaN/GaN$ sytem has been recently demonstrated by Dasgupta et al. [20].

Since Hot electron Transistors seem to be a promising technology for high-frequency devices, efforts have been directed to engineer and fabricate heterojunctions based on $AlGaN/GaN$. Among these, Gupta et al. [80] reported optimized $AlGaN/GaN$, and $InGaN/GaN$ diodes to be used as the emitter-base and base-collector structures respectively for HET applications. Within this context, this work presents a study of the EB heterojunction diodes, where non-ideal effects are proposed and discussed in order to explain the experimental device performance.

The simulated structure is shown in figure 3.1 and corresponds to the ideal layout of the experimental device fabricated at the University of California Santa Barbara [80]. The dipole-induced barrier is achieved with a thin, 4 nm layer of $Al_{0.45}Ga_{0.55}N$ unintentionally doped (UID), on top of a n^+ (N_D $= 5\text{x}10^{18}\,\text{cm}^{-3}$), 200 nm thick GaN substrate. An UID 35 nm region of $Al_{0.15}Ga_{0.85}N$ is placed on top of the barrier to reduce the reverse leakage current. A smooth transition between the $Al_{0.15}Ga_{0.85}N$ layer and the 50 nm n^+ GaN layer is achieved by means of a n^+ 15 nm graded region, i.e. a region where the mole fraction is gradually increased. All layers were grown by metal organic chemical vapor deposition (MOCVD) [50; 81] on Ga-Polar Bulk GaN substrate. Finally, a step of reactive ion etching followed by Al/Au deposition was done with the aim of fabricating $n-$type ohmic contacts.

It must be noticed that this design was optimized to be used as the emitter-base launching structure (EB diode) required in HET devices. In this regard, the $Al_{0.45}Ga_{0.55}N$ layer was intended to be AlN, since $AlGaN/AlN/GaN$ structures show higher mobility as well as higher sheet charge concentration in the 2 Dimensional Electron Gas (2DEG) [82]. However, it has been reported that pure AlN layers cannot be fabricated with Metal-Organic Chemical Vapor Deposition (MOCVD) due to Ga incorporation during the growth process [83]. Nevertheless, high Al mole fraction in the $AlGaN$ barrier still produces a conduction band discontinuity and a polarization-dipole between the $Al_{0.45}Ga_{0.55}N/GaN$ layers which lead to a high barrier in the conduction band [43], suitable for the injection of electrons from the emitter into the base. Furthermore, a high concentration 2DEG is induced in the GaN base, close to the interface, which increases the base charge reducing its sheet resistance.

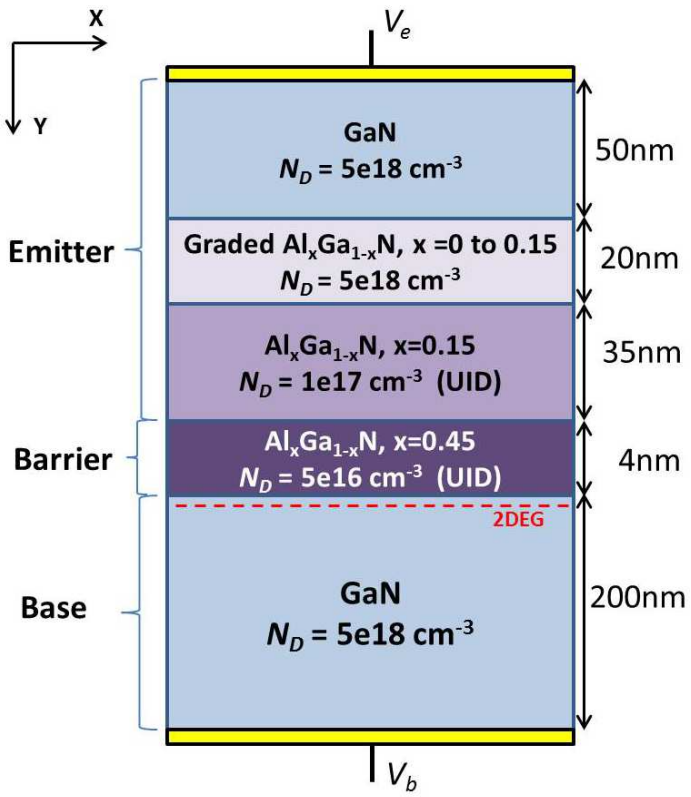

Figure 3.1: Experimental device layout corresponding to the ideal structure of the emitter-base diode. The emitter and base contacts are included.

3.2 Characterization of Experimental Device with Ideal Layout

The experimental device was first studied by means of CMC simulations, using the ideal structure shown in figure 3.1, with the emitter and base contacts placed at the top and bottom of the stack, respectively. The polarization dipole charge σ_p present at the $AlGaN/GaN$ interfaces, as well as the bands discontinuities ΔE_C and ΔE_V where calculated following Ambacher's formalism [45; 84].

Figure 3.2 shows the equilibrium conduction band profile (V_{be}= 0.0V) and the concentration of electrons along the vertical (y) direction. The band shows a barrier height ϕ_b = 1.7 eV, followed by an accumulation region where a 2DEG is formed right at the $Al_{0.45}Ga_{0.55}N/GaN$ interface, with a charge sheet concentration of $8x10^{13}$ cm^{-2}. It should be noticed that the $Al_{0.15}Ga_{0.85}N/Al_{0.45}Ga_{0.55}N$ interface produces a thin

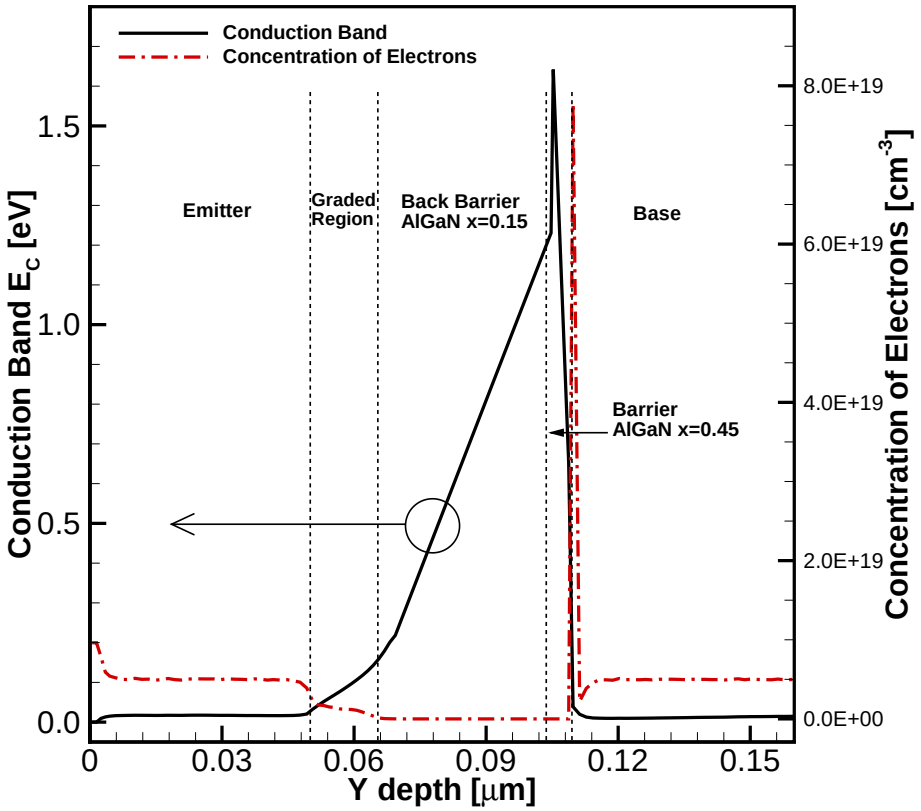

Figure 3.2: Profile of the conduction band in thermal equilibrium (V_{be}=0.00V), along with the concentration of electrons (right-axis) as obtained with CMC simulations.

barrier 4nm thick and approximate height of 0.5 eV, suggesting tunneling as one of the charge transport mechanisms besides thermionic emission.

Figure 3.3 shows the evolution of the band profile as V_{be} is increased (positive voltage applied to the base contact while the emitter is grounded). At low bias ($qV_{be} << \phi_b$), all the applied voltage is dropped across the barrier, modulating the barrier height seen by electrons injected from the emitter into the base. This leads to the rectifying IV characteristic shown in figure 3.5. As the bias is increased, some of the voltage drops across the $Al_{0.15}Ga_{0.85}N$ barrier and the current begins to saturate. Finally, at high bias ($qV_{be} > \phi_b$) the current saturates, and its value is high enough to make the base resistance non negligible, which produces band bending in the base.

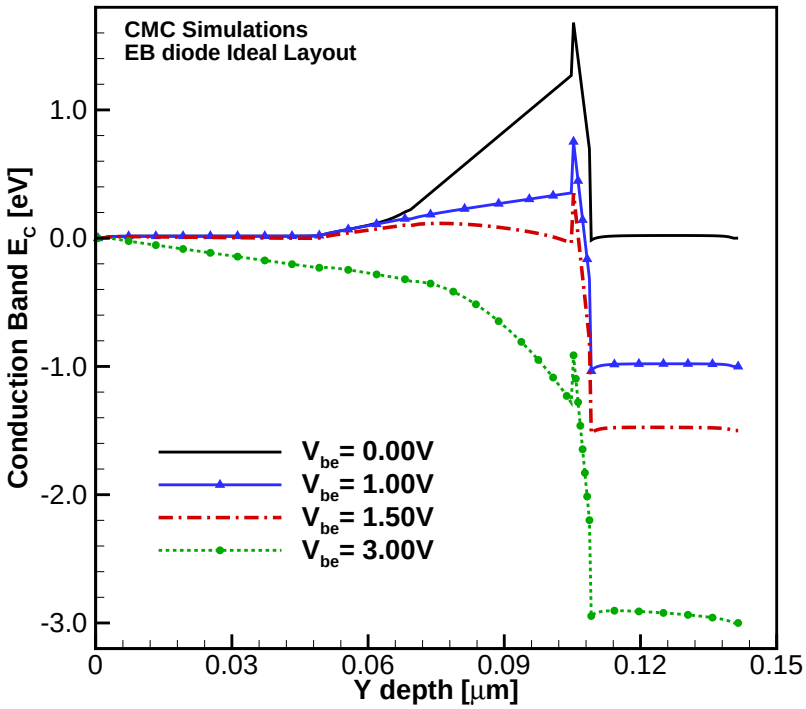

Figure 3.3: Evolution with bias of the conduction band profile along the y-direction of the EB diode as obtained with CMC simulations.

Recalling that this diode has been designed to work as a "launcher" of hot-electrons or EB diode for application in HETs, it is important to point out the potential of particle-based simulations to study such devices. As an example, figure 3.4 shows the 3D conduction band at $V_{be} = 1.25\,\mathrm{V}$ along with a snap-shot of the electron distribution in real space in the device, where spheres correspond to the simulated super-particles and their colors indicated the carrier's energy.

From figure 3.4, it can be seen that at the given bias, electrons are injected from the emitter into the base of the device by means of direct tunneling and thermionic emission, creating a population of hot-electrons that is different to that of the electrons at the bottom of the conduction band at equilibrium. As the injected electrons transit across the base, they lose energy through inelastic scattering events, so that some of them thermalize while others maintain their high kinetic energy. In the case of a HET, a second barrier (base-collector) would be placed at some distance after the

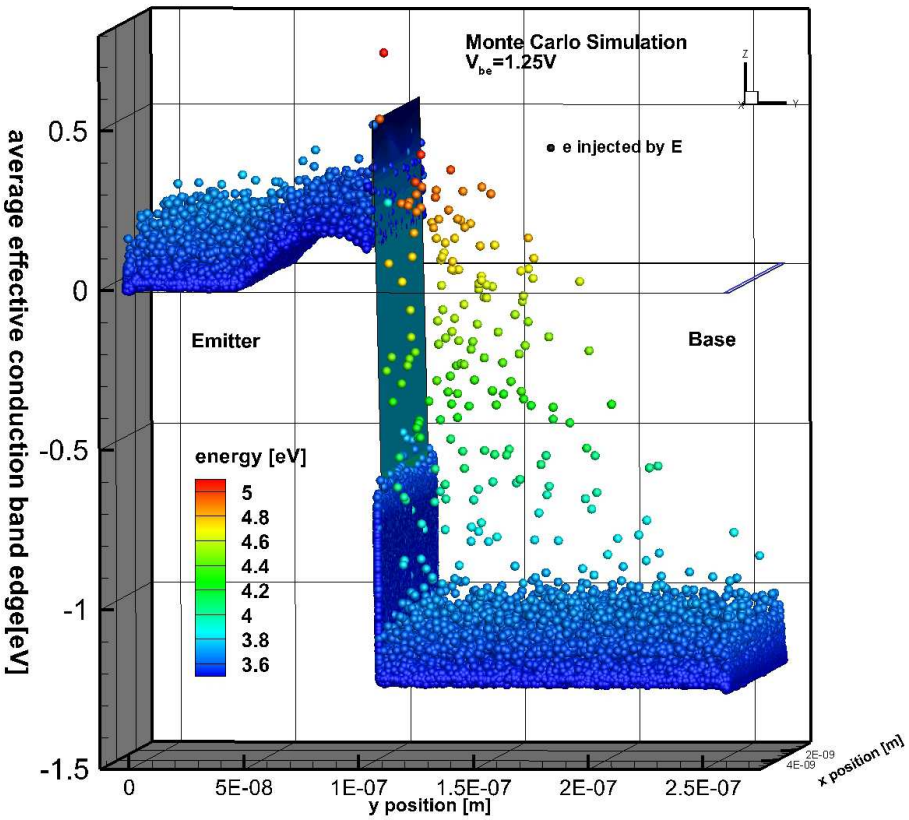

Figure 3.4: Conduction Band in 3D and a snap-shot of the electron distribution in real space at V_{be}=1.25V. The color of each super-particle corresponds to its energy.

EB barrier, and only the electrons with energy higher than the barrier would be collected. The CMC capability of handling two populations of electrons allows the study of charge transport and electrical characteristics relevant for the design of HET devices such as the mean free path of electrons, the critical base thickness for high-frequency operation and the effect of the BC barrier height on transfer coefficient, among others [21].

Figure 3.5 shows the IV characteristic of the EB diode obtained with CMC simulations along with the experimental data. One limitation of the CMC is that current densities typically under $10\,\mathrm{A/cm^2}$ are in the order of magnitude of the numerical statistical noise and cannot be resolved for this geometry. As an example of this, the current point at $V_{be} = 0.75\,\mathrm{V}$ is shown. Additionally, the simulation results show a

notorious discrepancy with respect to the experimental data. The simulated IV curve has a lower turn-on voltage, a steeper slope at low bias (or lower ideality factor n) and a saturation current several orders of magnitude higher.

In order to validate the CMC results, hydrodynamic simulations were performed with software TCAD, showing an excellent agreement with those obtained with CMC. The observed discrepancies between simulations and experimental data can be explained by considering several non-idealities of the experimental device. It must be highlighted that since hydrodynamic simulations are based on the analytical solution of the drift-diffusion equations, it is possible to resolve small currents accurately which is a characteristic necessary to reproduce the full range of the IV curve of the EB diode. Furthermore, the effect of Schottky barriers can also be included within the hydrodynamic approach. For these reasons, the study of non-idealities presented in the following sections is performed with TCAD simulations only.

3.3 Mole Fraction Variations in the $Al_xGa_{1-x}N/GaN$ Barrier

As it was mentioned before, the heterojunction diode under study was designed to be the EB junction in a HET. The optimal structure for such applications consists of an $AlGaN/AlN/GaN$ stack instead of the typical $AlGaN/GaN$ structure. It has been reported by Dasgupta [43] that a high mole fraction of Al in $AlGaN/GaN$ barriers induces a higher sheet charge at the interface. However, the centroid of the 2DEG wave function is drawn closer to the heterointerface where alloy scattering and interface roughness scattering are higher [85], leading to lower mobilities in the 2DEG [82]. The mobility degradation issue can be overcome by adding a thin AlN interlayer that reduces the effect of the scattering mechanisms by drawing the 2DEG away from the $AlGaN$ layer, in addition to providing higher sheet charge densities and lower sheet resistance [82; 43].

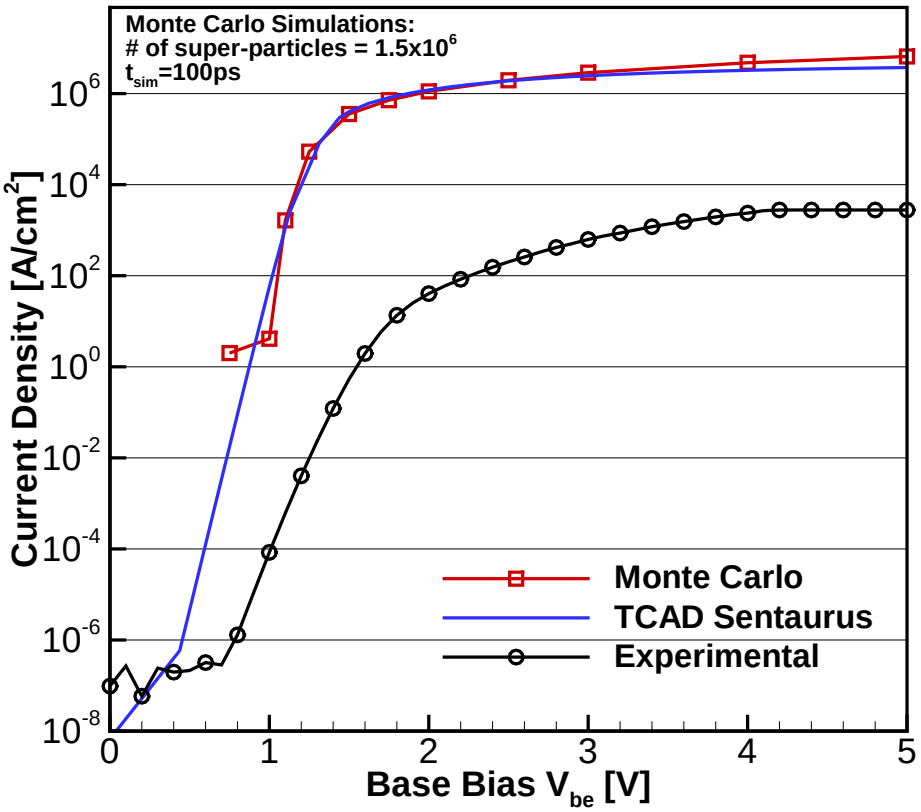

Figure 3.5: IV characteristic in semi-log scale of emitter-base diode, showing experimental data along with simulation results obtained by CMC, and TCAD Sentaurus simulations.

Mazumder et al. [83] have recently reported that growing pure AlN interlayers in $AlGaN/AlN/GaN$ heterostructures by MOCVD is not possible due to the unintentional incorporation of Ga during the growth process. The high temperature required for MOCVD ($T > 1100$C) leads to evaporation of Ga species from both the GaN substrate and unintentional deposits of GaN on the reactor. Besides, as the thickness of the AlN interlayer is increased, the structural quality of the AlN/GaN junction tends to degrade due to the presence of threading dislocations [86]. In their experiments, Mazumder et al. [83] used Atom Probe Tomography (APT) to characterize $AlGaN/AlN/GaN$ heterostructures grown by MOCVD, showing that for an AlN interlayer of 2 nm, the incorporation of Ga resulted in a mole fraction of approximately $x = 0.47$. These results suggest that there is a significant uncertainty regarding the

exact mole fraction in the $Al_xGa_{1-x}N/GaN$ barrier obtained during the fabrication of the EB diode. Therefore, variations in the mole fraction constitute a physically meaningful non-ideal effect and its impact is studied with TCAD simulations.

Figure 3.6 shows the equilibrium band profile in the vertical direction and the electron density for two values of mole fraction x=0.45 and 0.50, as obtained with hydrodynamic simulations. It can be seen that for the $Al_{0.50}Ga_{0.50}N/GaN$ junction with higher mole fraction, the barrier height results in 2 eV, compared to 1.7 eV obtained with $x = 0.45$. As the Al mole fraction increases, the band gap and the polarization charge increase as well, inducing a larger barrier and higher electron density of 2.4x10^{19} cm^{-3} for $x = 0.50$, compared to 2.0x10^{19} cm^{-3} for $x = 0.45$. These results are in agreement with those previously reported [83], where a barrier height of 2.1eV was obtained by Hall measurements on MOCVD grown $AlGaN/AlN/GaN$ heterostructures.

The IV characteristic computed for both mole fraction values is shown in figure 3.7 along with the experimental data. The turn-on voltage for the structure with $x = 0.50$ is larger by approximately 0.4 V compared to that obtained with $x = 0.45$, which is in good agreement with the experimental curve. For low bias ($qV_{be} < \phi_b$), the applied voltage drops entirely across the barrier. Since the barrier seen by electrons on the emitter side is higher for $x = 0.50$ than for $x = 0.45$, the structure with larger mole fraction yields smaller currents when the same bias is applied. At high bias where the saturation current is reached, the barriers no longer control the current flow and both structures yield the same current.

Whereas increasing the mole fraction of the barrier shifts the turn-on voltage, it has no effect in changing neither the slope of the curves at low bias nor the saturation current at high voltages, and the discrepancy with the experimental data remains the same in this regard. One possible explanation for the difference of 4 orders of

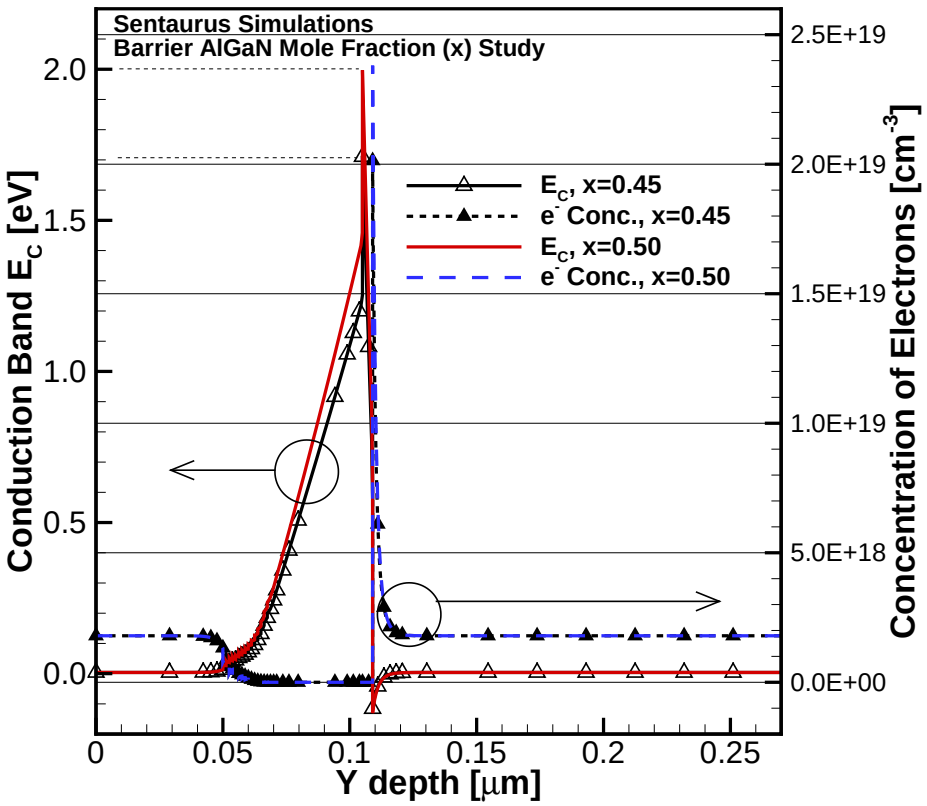

Figure 3.6: Vertical profile of the conduction band in equilibrium (left-axis) for $AlGaN$ with mole fractions $x = 0.45$ and 0.50. The corresponding concentration of electrons is also shown in the right-axis.

magnitude in the current at high voltage, is the effect of a parasitic series resistance R_s. Hence, the contact resistance parameter R_s was then included in the model.

This attempt is shown in figure 3.8, where the simulated IV curves of the structure with stack $Al_{0.50}Ga_{0.50}N/GaN$ and values of $R_S = 0\,\Omega$, $10^4\,\Omega$, $10^6\,\Omega$ and $10^7\Omega$ are plotted along with the experimental data. It can be seen that the current density is significantly reduced so that with the highest resistance value, the simulation and experimental results show the same saturation current density.

However, this simplistic explanation neglects other effects of the physics of this heterojunction stack. First of all, a parasitic series resistance of $10\,\mathrm{M}\Omega$ is unrealistic. Furthermore, with high R_s the simulated IV curve shows a nearly constant value

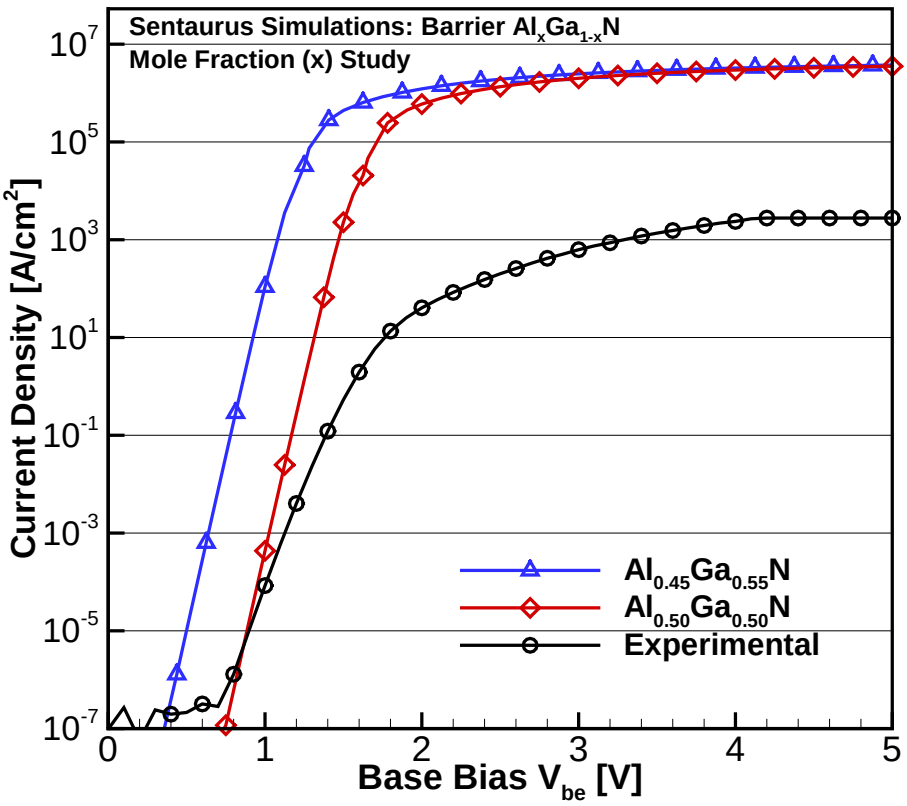

Figure 3.7: IV characteristic in semi-log scale, showing experimental data, along with simulation results for mole fraction values $x = 0.45$ and 0.50, at the $Al_xGa_{1-x}N/GaN$ barrier.

from $V_{be} = 1.5\,\mathrm{V}$ to $4\,\mathrm{V}$, whereas the experimental curve shows two clearly different slopes, one from $V_{be} = 0.8\,\mathrm{V}$ to $1.6\,\mathrm{V}$ and another one from $1.6\,\mathrm{V}$ to $4\,\mathrm{V}$. The current in the range from $4\,\mathrm{V}$ to $5\,\mathrm{V}$ is not treated as a third slope because its value is artificially limited by the experimental setup. These observations suggest the presence of additional barriers that affect the current control of the main structure, and non-intentional Schottky contacts both at the emitter and the base provide a physically sensible explanation for the experimental data [87].

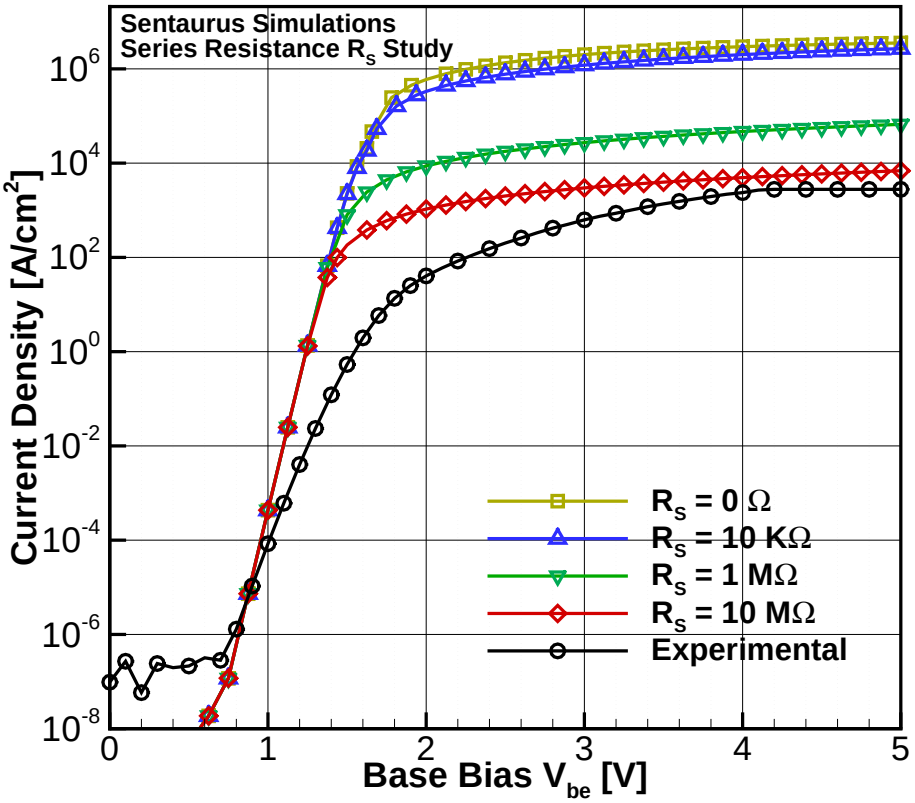

Figure 3.8: IV characteristic in semi-log scale of the EB diode showing experimental data, along with TCAD simulation results for different values of parasitic series resistance R_S in the contacts.

3.4 Effect of Emitter and Base Schottky Contacts

As it has been modeled by Chen et al. [87] and Lv et al. [88], the double slope observed in forward-biased IV characteristics of heterojunction diodes can be explained by considering the effect of all the barriers present in the device, usually a Schottky Metal-Semiconductor and an $AlGaN/GaN$ barrier. In the case under study, the contacts of the experimental device were fabricated by reactive ion etching to reach the bottom contact layer followed by Al/Au deposition, and ideally both emitter and base contacts should have an ohmic behavior. However, for this device structure annealed contacts cannot be used [80], which might lead to undesired Schottky behavior both in the emitter and the base. In order to understand the independent effect that

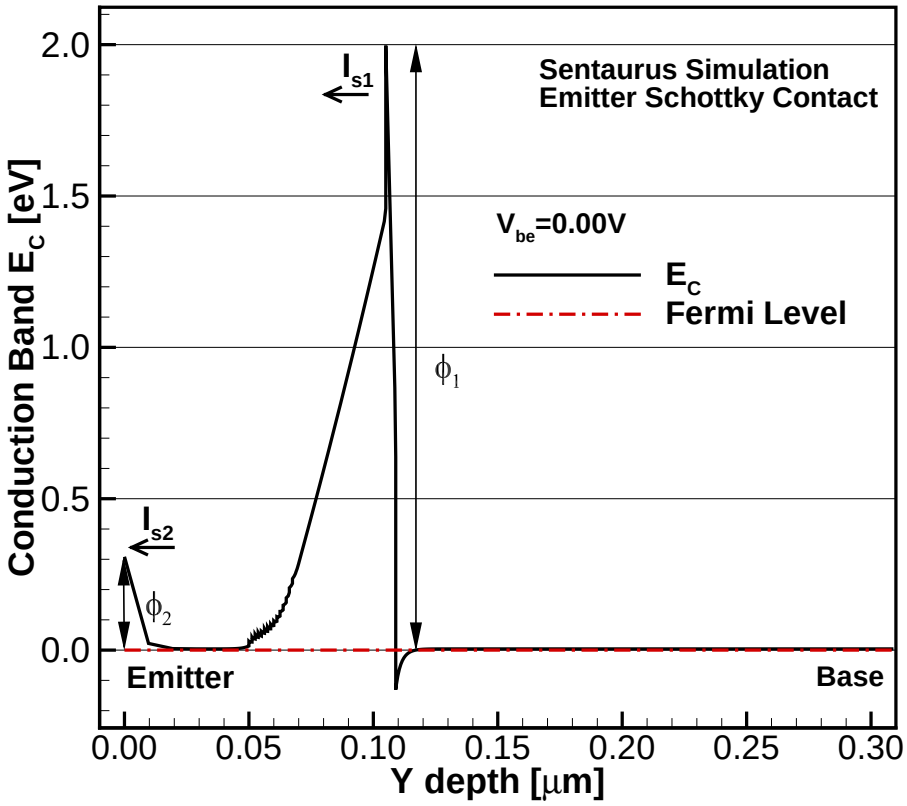

Figure 3.9: Vertical conduction band profile at equilibrium, where a Schottky contact is placed at the emitter with barrier height ϕ_2=0.45eV.

Schottky contacts have on the device performance, the emitter and base Schottky contacts were included in the simulated device.

Figure 3.9 shows the equilibrium band profile along the vertical direction, including a Schottky contact in the emitter and an ideal ohmic base contact. The band diagram suggests that the structure can be understood as two diodes in series, back to back: the $Al_{0.50}Ga_{0.50}N/GaN$ with barrier height ϕ_1 and saturation current I_{s1}, and the Schottky contact with barrier height ϕ_2 and saturation current I_{s2}.

When a small bias V_{be} is applied so that $qV_{be} < \phi_1 - \phi_2$, the $AlGaN/GaN$ diode is forward biased whereas the Schottky diode is reverse biased. Figure 3.10 (a) depicts the band diagram for this case at $V_{be} = 1.0$ V with and without Schottky contact for comparison purposes. At this point, since $\phi_1 >> \phi_2$, the current through the device

is comparable to I_{s1} which is several orders of magnitude smaller than I_{s2}, making the resistance across the $AlGaN/GaN$ diode several orders of magnitude higher than that of the Schottky contact. This results in the applied voltage being entirely dropped across the $AlGaN/GaN$ interface.

Note that both band profiles depicted in figure 3.10 (a), with emitter contact ohmic and Schottky, are identical except for the presence of the Schottky barrier at the emitter indicated as ϕ_{2a}, yielding the same current density for both devices at the given bias point, as it can be seen in figure 3.11, where the IV characteristic of the EB-diode is shown for different Schottky barrier heights ϕ_2. When the condition $qV_{be} > \phi_1 - \phi_2$ is reached, the voltage dropped across the barrier ϕ_1 pushes the Schottky barrier further into reverse bias regime, then the current through the device increases and becomes comparable to I_{s2}, and the Schottky barrier at the emitter begins to dominate. This situation is portrayed in figure 3.10 (b), for $qV_{be} = 1.5\,\mathrm{V}$, showing both the ohmic and Schottky emitter contact conduction band profile.

It can be seen in figure 3.10 (b) that electrons injected from the Schottky contact into the emitter see an effective barrier towards the base higher by a factor of $q\Delta V$, than the one seen by electrons in the ohmic case for the same bias. This is due to the reverse bias voltage dropped across the Schottky contact, which pulls down the conduction band in the $Al_{0.15}Ga_{0.85}N$ region on the emitter side. In turn the current is reduced by nearly 3 orders of magnitude as shown in figure 3.11, comparing the cases $\phi_b = 0.00$ (ohmic), and $\phi_b = 0.45\,\mathrm{eV}$ at $V_{be} = 1.5\,\mathrm{V}$.

As a higher bias is applied, the current is no longer dominated by the $AlGaN/GaN$ barrier and the Schottky barrier dominates suppressing the saturation current. Since the Schottky diode is reverse biased, the barrier height lowering effect was included in the simulation setup, in order to capture the image force effect and the thermionic-field emission current [7; 78], as depicted in figure 3.10 (b) with $\phi_{2b} < \phi_{2a} < 0.45\,\mathrm{eV}$.

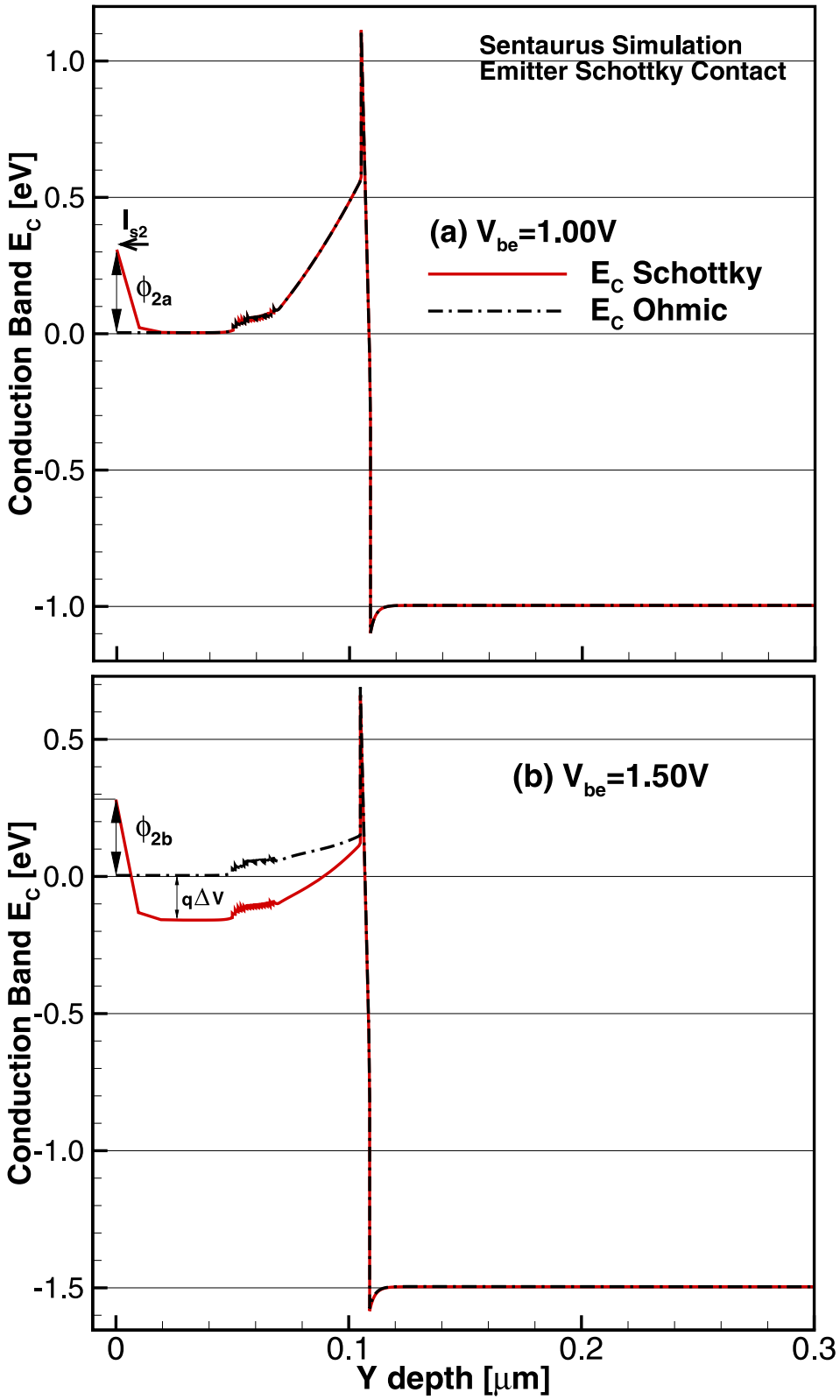

Figure 3.10: Vertical conduction band profile with and without emitter Schottky contact for a) V_{be} = 1.0 V and b) V_{be} = 1.5 V, with Schottky barrier height ϕ_2 = 0.45 eV, while keeping an ohmic base contact.

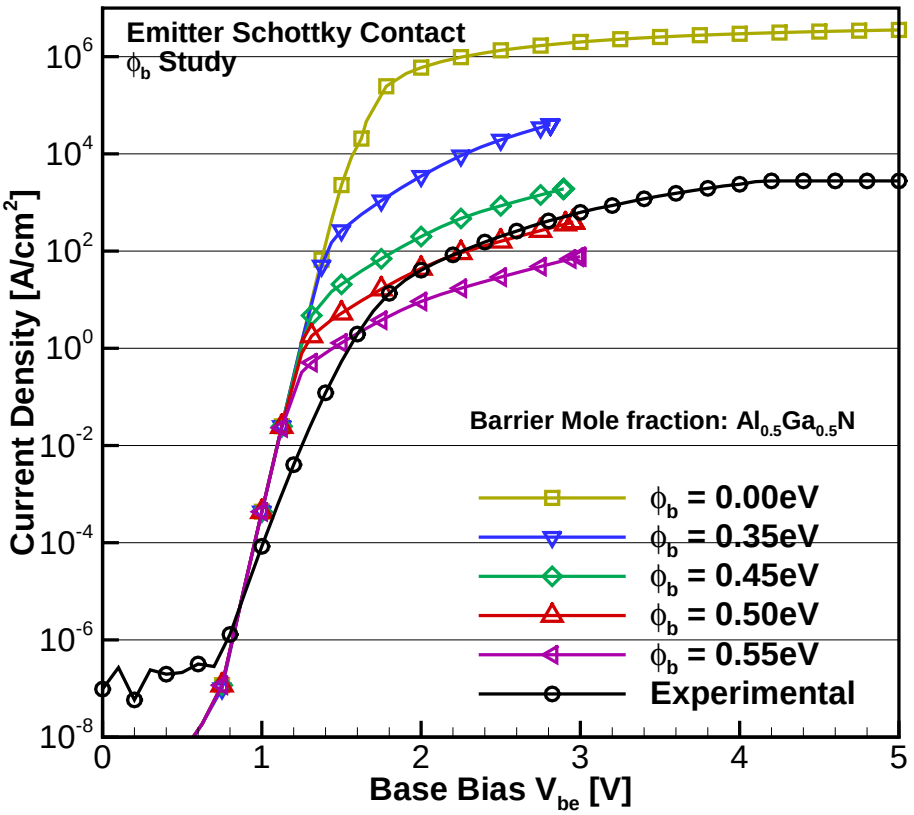

Figure 3.11: IV characteristic in semi-log scale, showing experimental data along with simulation results including the effect of a Schottky contact at the emitter, for different barrier heights.

The IV curves in figure 3.11 show that this approach based on a non-ideal Schottky contact at the emitter is in excellent agreement with the experimental data at high positive bias from $V_{be} = 1.8\,\mathrm{V}$, not only in terms of the intensity of the current density, but also in its slope or ideality factor n, which is modulated by the dependence of the barrier height on voltage introduced with the barrier height lowering effect. It must be highlighted that this result could not be achieved by adding a parasitic series resistance, since in that case the current would not be modulated by a barrier.

In the case of a Schottky contact in the base, the behavior of the device is different to that observed for the emitter Schottky contact. Figure 3.12 corresponds to the equilibrium conduction band profile of the diode under study with a base Schottky

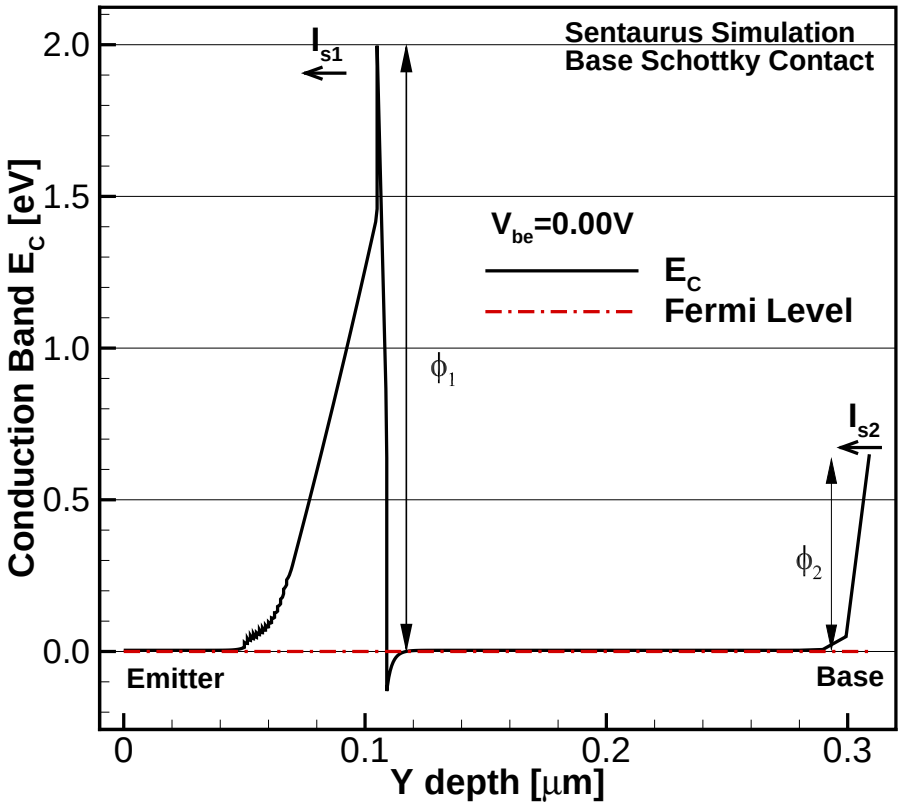

Figure 3.12: Vertical conduction band profile at equilibrium, with a Schottky contact placed at the base with barrier height $\phi_2 = 0.65\,\text{eV}$, while keeping an ohmic emitter contact.

contact with barrier $\phi_2 = 0.65\,\text{eV}$. Once more, the band profile suggests the presence of two diodes in series, one given by the $Al_{0.50}Ga_{0.50}N/GaN$ structure with barrier height ϕ_1 and saturation current I_{s1}, and the other by the base Schottky contact with barrier height ϕ_{b2} and saturation current I_{s2}.

In contrast to the emitter case, when a small bias V_{be} is applied so that $qV_{be} < \phi_1 - \phi_2$, both the $AlGaN/GaN$ diode and the Schottky diode at the contact are forward biased. Figure 3.13 (a) shows the band diagram for $V_{be} = 0.5\,\text{V}$ with and without the base Schottky contact. At this bias point, the current through the device is comparable to I_{s1} which is smaller than I_{s2} because $\phi_1 > \phi_2$, making the resistance across the $AlGaN/GaN$ diode larger than that of the Schottky contact and all the voltage is dropped across the $AlGaN/GaN$ interface.

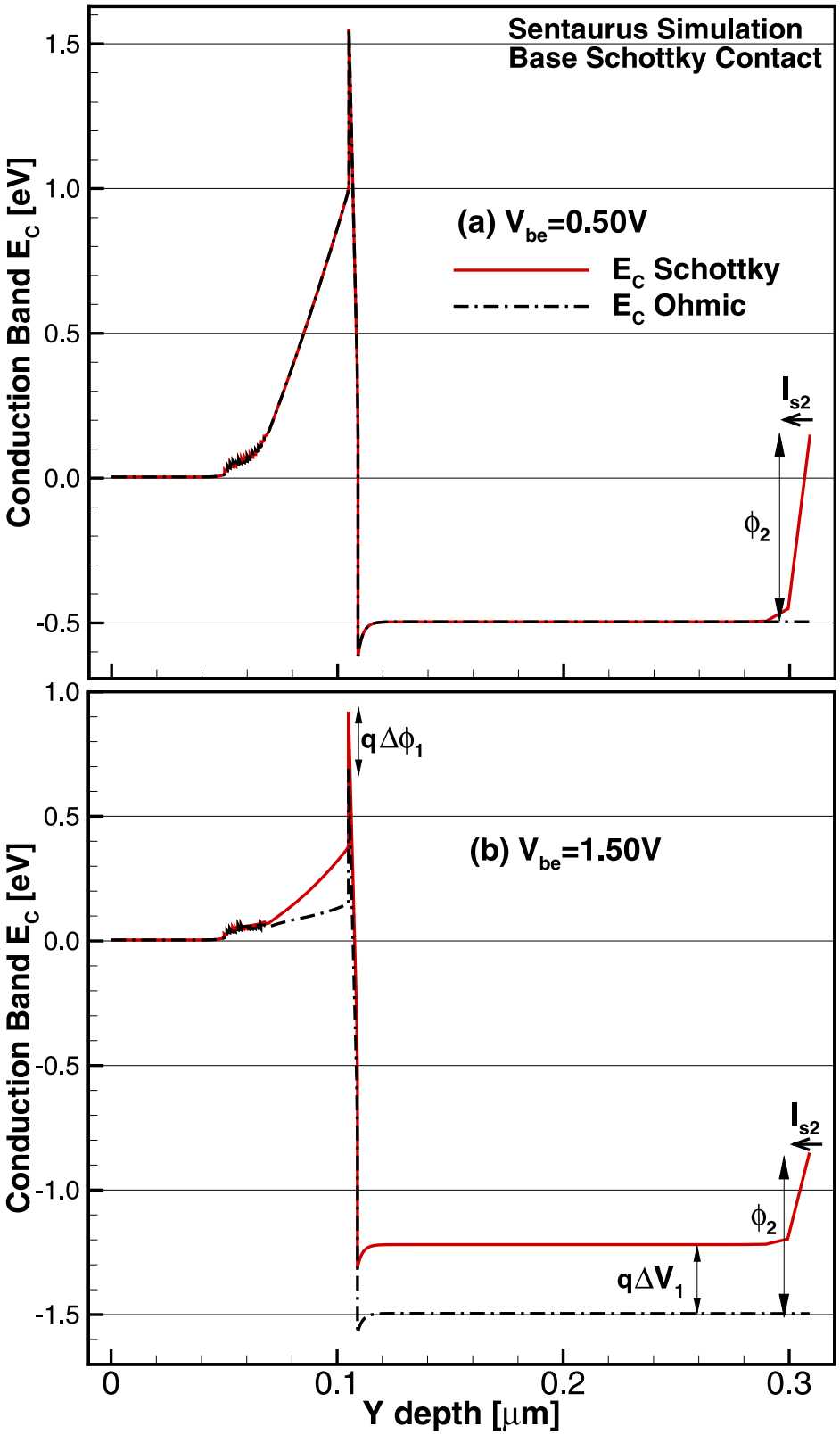

Figure 3.13: Vertical conduction band profile with and without base Schottky contact for a) $V_{be} = 0.50\,\mathrm{V}$ and b) $V_{be} = 1.5\,\mathrm{V}$, with $\phi_2 = 0.65\,\mathrm{eV}$.

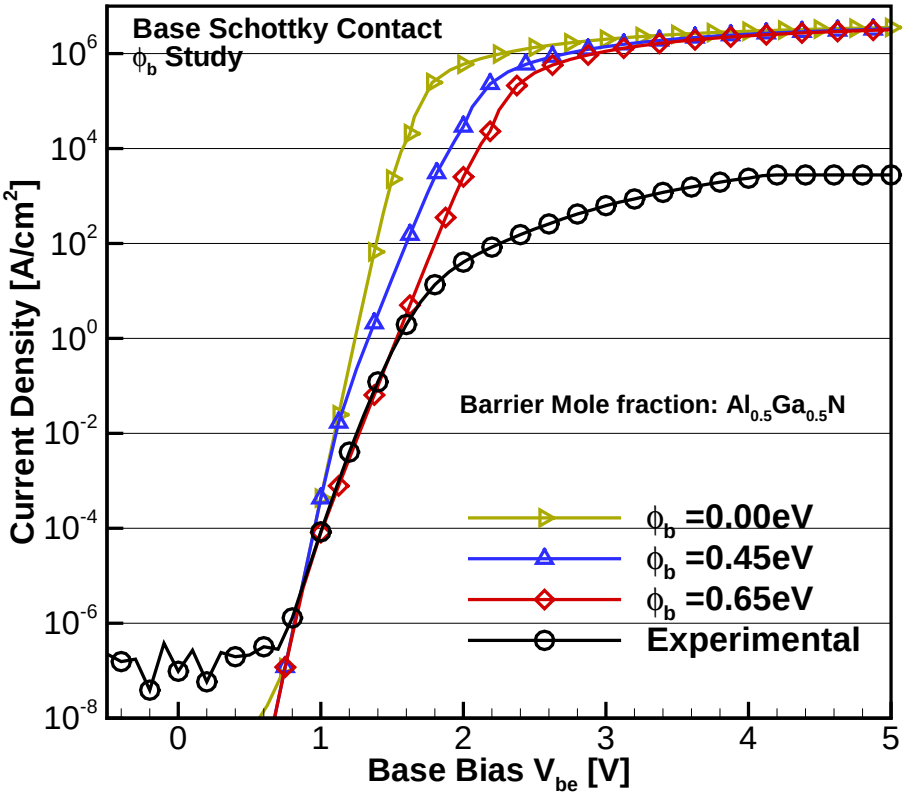

Figure 3.14: IV characteristic in semi-log scale showing experimental data, along with simulation results including the effect of a Schottky contact at the base, with different barrier heights.

Hence, both band profiles with ohmic and Schottky base contact are the same, as well as the simulated currents at low bias as shown in figure 3.14. When the applied voltage increases, the current through the device becomes comparable to I_{s2}. From this point, the Schottky diode resistance is no longer negligible and some voltage will drop across the contact. This situation is presented in figure 3.13 (b) for $V_{be} = 1.5\,\mathrm{V}$, showing both the ohmic and Schottky base contact conduction band profiles.

Unlike the emitter case, from figure 3.13 (b) it can be seen that the voltage drop at the base, indicated as ΔV_1, reduces the voltage across the $Al_{0.50}Ga_{0.50}N/GaN$ barrier. As a consequence, the barrier seen by electrons in the emitter towards the base is higher by a factor $\Delta\phi_1$ in this case than with an ideal ohmic base contact.

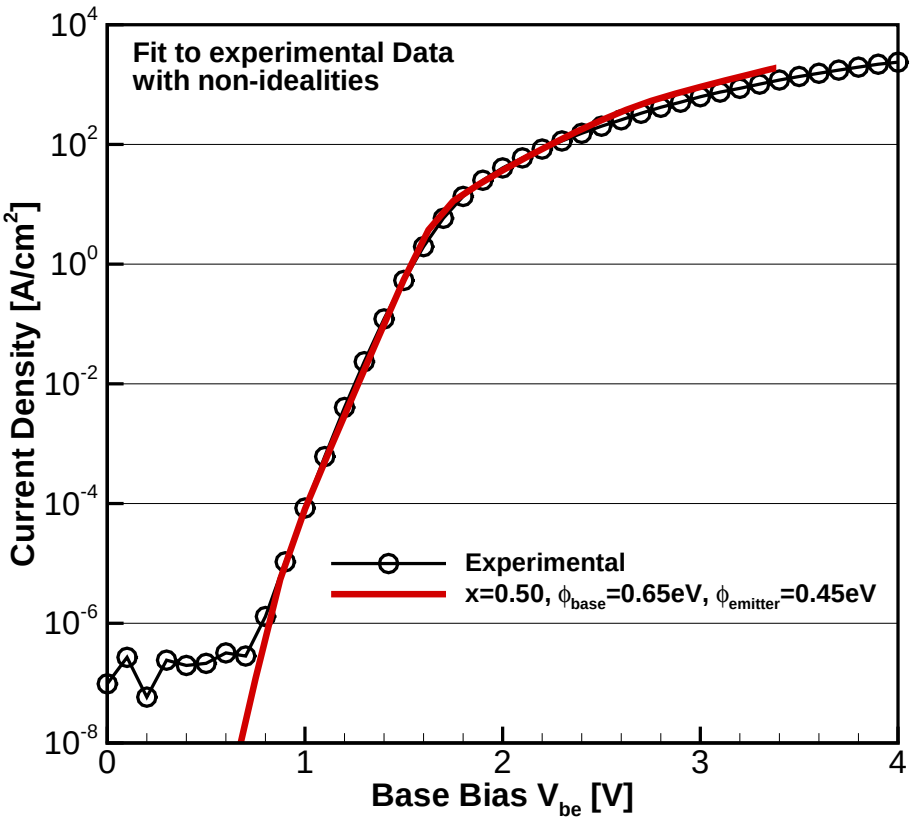

Figure 3.15: Semi-log IV characteristic, showing experimental data along with the simulated curve corresponding to a structure with $Al_{0.50}Ga_{0.50}N$, $\phi_{base} = 0.65\,\text{eV}$, $\phi_{emitter} = 0.45\,\text{eV}$.

The factor $\Delta\phi_1$ depends only on the voltage drop ΔV_1, indicating that the $Al_{0.50}Ga_{0.50}N/GaN$ barrier is being modulated by the voltage drop in the Schottky contact. As further positive bias is applied, the Schottky contact goes further into forward bias and the $AlGaN/GaN$ barrier will reach saturation at higher V_{be} than in the ohmic case, but the saturation current density will remain in the same order of magnitude for both cases.

This can be observed in the IV characteristics shown in figure 3.14 where the experimental data is plotted along with simulation results for three different Schottky barrier heights of the base contact. As explained before, the slope of the IV curve at low bias depends on the base Schottky barrier height. In particular, the IV curve for

$\phi_b = 0.65\,\text{eV}$ is in excellent agreement with the experimental data from $V_{be} = 0.8\,\text{V}$ to 1.6 V, whereas the saturation current value remains several orders of magnitude higher than the experimental one, independently of the Schottky barrier height at the base. Finally, when all the non-idealities described are put together in the model, an excellent agreement between simulation and experimental data is reached for the full extension of the IV curve. Figure 3.15 shows the fitted IV characteristic with the following parameters: $Al_{0.50}Ga_{0.50}N$, $\phi_{base} = 0.65\,\text{eV}$, $\phi_{emitter} = 0.45\,\text{eV}$.

3.5 Conclusions

In this work, a simulation study of isotype heterojunction diodes was presented. CMC simulations of the forward biased IV characteristic of the diode using the ideal layout of the device were presented and validated with hydrodynamic simulations performed with the commercial software TCAD Sentaurus by Synopsys, showing excellent agreement. However, as the simulated IV curves differed significantly from the reported experimental data, a further analysis including non-idealities was done with hydrodynamic simulations only. When variations in the mole fraction of the $Al_xGa_{1-x}N/GaN$ barrier and the effect of Schottky contacts were included in the model, the experimental data was successfully reproduced by simulations. In the context of the development of HETs technology, this work provides valuable information about non ideal effects that might be deleterious to the performance of HETs.

Chapter 4

HOT ELECTRON GENERATION UNDER LARGE-SIGNAL RF OPERATION OF GAN HEMTS

In order to assess the underlying physical mechanisms of hot-carrier related degradation such as defect generation in millimeter-wave (mm-wave) GaN power amplifiers (PAs), a study of the electron energy distribution function under large-signal radio frequency (RF) conditions in $AlGaN/GaN$ high-electron-mobility transistors is presented here. The study is performed with a full band Cellular Monte Carlo particle-based device simulator (CMC) self-consistently coupled to a harmonic balance circuit solver [89].

Simulations of a Class AB power amplifier at 10 GHz show that the peak of hot electron generation is up to 43% lower under RF drive than it is under DC conditions, regardless of the input power or temperature of operation. However, at mm-wave operation up to 40 GHz, RF hot carrier generation reaches that from DC regime and even exceeds it up to 75% as the amplifier is driven into compression. Increasing the temperature of operation also shows that degradation of DC and RF characteristics are tightly correlated and mainly caused by increased phonon scattering. In addition, due to the popularity of Class A PAs, a comparative study between amplifier topologies is performed, indicating that devices operating in Class A circuit configuration are more vulnerable to hot electron degradation due to higher hot-carrier concentration as compared to Class AB. The accurate determination of the electron energy mapping is demonstrated to be a powerful tool for the extraction of compact models used in lifetime and reliability analysis.

4.1 Introduction

The remarkable performance of GaN high-electron-mobility transistors (HEMTs) have made them a prominent technology in the field of integrated RF power amplifiers (PAs), limited only by cost, thermal management, and reliability concerns. Many studies have been done on HEMT reliability [22], and hot electron degradation is consistently cited as one of the most relevant mechanisms [26]. Under electrical stress, hot carriers with energy higher than an activation threshold can create electrically active traps by dehydrogenation of passivated point defects [90]. Recently [59; 91], it was shown that determining the electron energy distribution function (EDF) using a multi-band Ensemble Monte Carlo device simulator, can be used to explain the degradation observed during DC lifetime testing in terms of the threshold voltage shift and reduced transconductance. However, most of the studies are limited to DC operation only, and have been proven to fail in predicting RF reliability [92].

While in some experimental reports [93; 94] degradation under RF stress was found to be more severe than in DC conditions, in other works [95] the degradation was comparable or even higher in DC regime, with hot-electrons as the proposed mechanism in both cases. To the best of the author's knowledge, the only study of hot electron concentration under RF conditions was done by means of electroluminescence measurements [71; 72], with the limitation that an accurate quantitative estimate of the carrier energy distribution has not been obtained to date. As of this writing, the correlation between DC and large-signal RF reliability testing is still a topic of active research, and a deeper understanding of the underlying physical mechanisms of hot-carrier generation under large-signal RF is necessary.

In this chapter, the study of hot electron generation under large-signal RF operation is performed by means of the accurate simulation of the EDF, which enables

a qualitative description of hot-carrier induced defect generation associated to the degradation of device performance. In this approach, the CMC simulation model of a GaN HEMT is first calibrated to experimental DC curves of a fabricated device [96]. Then, RF simulations are performed by means of CMC simulations [37], self-consistently coupled with with a harmonic balance circuit solver based on the active load-line technique [89]. This technique exploits the accuracy of full-band Monte Carlo simulations in the description of highly non-linear, out of equilibrium carrier-dynamics, while allowing for realistic in-circuit operating conditions.

This study demonstrates that at low frequencies of operation in Class AB configuration, the peak hot electron generation under RF is lower than that in DC conditions for all input powers and temperatures of operation, suggesting a lower degradation under RF drive. However, as the frequency of operation is increased reaching the millimeter-wave band, the electron EDF reaches that of DC, and even exceeds it at high energies when the amplifier is driven into compression, which implies a more severe degradation under RF than in DC conditions. In addition, it is quantitatively shown that DC performance degradation at high temperatures of operation is tightly correlated to the degradation of the RF figures of merit (FOM), and mainly caused by increased phonon scattering in the channel of the device. Finally, a comparative study between power amplifier classes is presented, showing that hot electron generation under Class A is higher than under Class AB configuration suggesting that devices are more susceptible to degradation under Class A topology.

4.2 GaN HEMT Model Validation

The experimental device modeled in this study is shown in figure 4.1. It consists of an $Al_{0.29}Ga_{0.71}N/AlN/GaN$ epitaxial structure on SiC, with thicknesses of 15.4 nm and 1 nm for the $AlGaN$ and AlN layers respectively. The device is capped with a

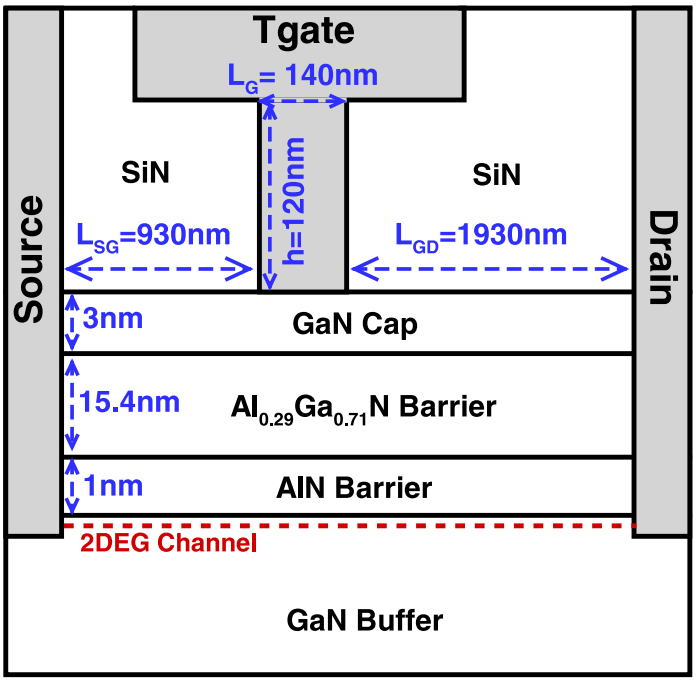

Figure 4.1: Device layout of the fabricated GaN HEMT used in the CMC model.

3 nm GaN layer followed by 120 nm of SiN passivation. The gate contact is T-shaped with 140 nm of metallurgical length, and the source-gate and gate-drain access region lengths are 930 nm and 1930 nm respectively. In all simulations, the carrier dynamics is modeled within the CMC framework [37] by tracking carriers in the first Brillouin zone, considering scattering processes due to deformation potential phonons, polar optical phonons, piezoelectric (polar acoustic) phonons, ionized impurities, impact ionization, and thread dislocations.

The electronic band structure is obtained by an empirical nonlocal pseudopotential method, with material parameters previously calibrated for wurtzite GaN [97], whereas the phonon effects are computed with the volume-shell approach [39]. The device domain in real space is simulated by a particle-based dynamic kernel self-consistently coupled with a 2D multigrid Poisson solver [98]. The polarization-induced charge due to the high polarization discontinuity across the heterojunction interfaces has been calculated following Ambacher's formalism [45], and placed as charge sheet

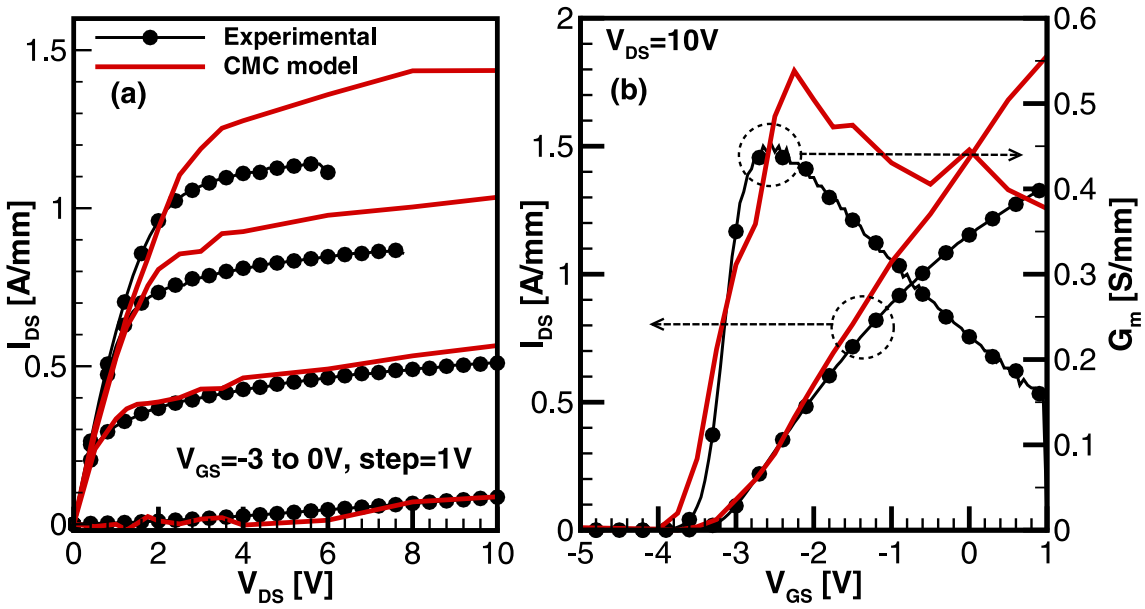

Figure 4.2: Experimental and simulated (a) output and (b) transfer (along with Transconductance G_m) characteristics.

layers at the appropriate heterointerfaces, which effectively reproduces the charge distribution inside the device, creating a high density 2D electron gas (2DEG) at the AlN/GaN interface [55].

The validation of the CMC model with the DC experimental data of the output and transfer characteristics is presented in figures 4.2 (a) and (b), respectively, showing excellent agreement at low gate voltage. However, since self-heating effects are not included in this study, the drain current is overestimated for values above 0.8 A/mm as reported previously [99]. By means of small-signal AC simulations, the values of cut-off frequency f_t = 104 GHz and maximum oscillation frequency f_{MAX} = 234 GHz were determined.

4.3 Large Signal RF Characterization

Conventional AC small-signal modeling fails to predict the performance of power amplifiers, because under large-signal operation the active device presents a highly

non-linear response which is affected by the external load and the output matching network (OMN). Nevertheless, it is possible to accurately characterize the device as a power amplifier by self-consistently coupling the time-domain solution of the active non-linear device obtained from the CMC simulator with a harmonic balance solver that provides a frequency domain solution of the external circuitry formed by the load connected to the drain through a high-Q OMN. The extrinsic network is implemented by using the active load-line technique [89], such that the equivalent impedance of the circuit is emulated with sinusoidal voltage generators connected in series to the drain, one per each desired frequency component (fundamental and n-harmonics).

The simulation begins with an initial guess for the amplitude and phase of the load-line generators to emulate a target load impedance $Z_{Ltarget}$. After a CMC solution is obtained, the actual synthesized load impedance Z_L is determined in post-processing for each frequency component by

$$\begin{aligned} Z_L(\omega_n) &= |Z_L(\omega_n)|\, e^{j(Z_L(\omega_n))} \\ &= \frac{v_D(\omega_n)}{i_D(\omega_n)} = \frac{|v_D(\omega_n)|}{|i_D(\omega_n)|} e^{j(\angle v_D(\omega_n) - \angle i_D(\omega_n))}, \end{aligned} \tag{4.1}$$

where $v_D(\omega_n)$ and $i_D(\omega_n)$ are the complex phasors of the drain current and voltage sinusoids at frequency ω_n. The difference between the synthesized and target impedance is compared to a predefined tolerance, and if the target is not achieved, another CMC simulation is run with adjusted values for the amplitude and phase of the voltage generators, calculated by

$$|v_D'(\omega_n)| = |v_D(\omega_n)| \left(\frac{|Z_{Ltarget}(\omega_n)|}{|Z_L(\omega_n)|} \right), \tag{4.2}$$

$$\angle v_D'(\omega_n) = \angle v_D(\omega_n) + (\angle Z_{Ltarget}(\omega_n) - \angle Z_L(\omega_n)). \tag{4.3}$$

This process is iterated until convergence to the target impedance is achieved for all the frequencies. The RF characterization of the experimental device was performed

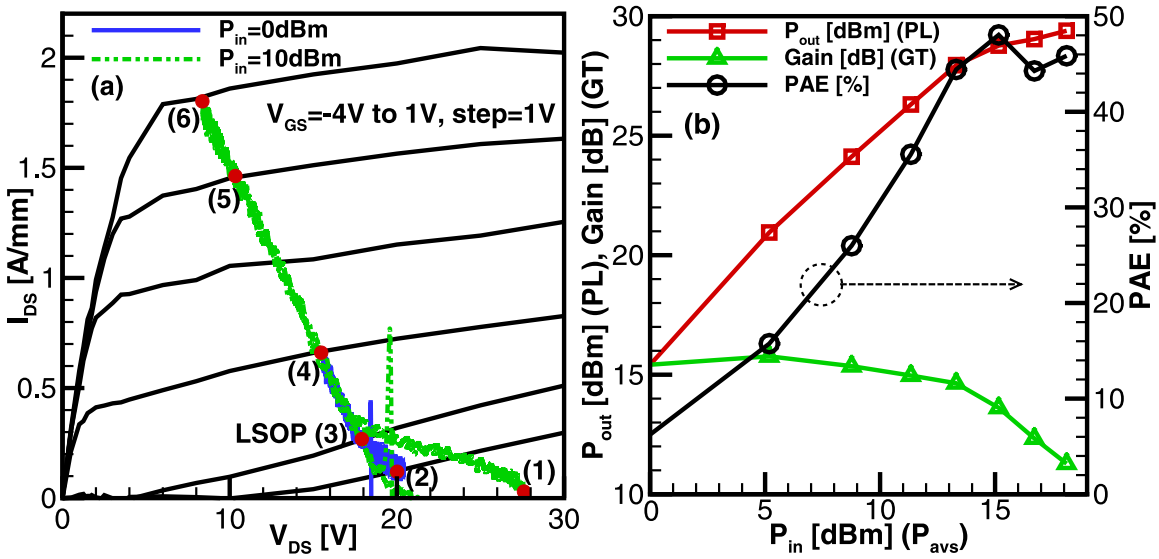

Figure 4.3: Power amplifier characterization under Class AB (RF) operation, showing (a) dynamic load-lines and (b) FOM, for $f_{req} = 10\,\text{GHz}$ and tuned-load $Z_{Load} = 100\,\Omega$.

for a Class AB configuration with a large signal operating point (LSOP) of V_{GS} = -3 V, V_{DS} = 18 V and I_{DS} = 250 mA/mm, at f_{req} = 10 GHz with a tuned-load impedance Z_{Load} = 100 Ω for fundamental and 0 Ω for higher harmonics. The dynamic load lines are shown in figure 4.3 (a) along with the typical figures of merit (FOM) in figure 4.3 (b). Class AB with tuned-load operation was chosen for the study because it is one of the preferred PA topologies due to the high output power and high power conversion efficiency, evidenced by the simulated values of P_{out} = 29 dBm and power added efficiency PAE = 49 % for an input power of P_{in} = 15 dBm.

4.4 Hot Electron Generation

In order to establish a comparative baseline for the hot electron generation under RF, we first study the EDFs for DC bias points indicated with solid dots (1) to (6) in figure 4.3 (a), which describe the full excursion of the dynamic load line for input powers from 0 dBm to 10 dBm. The resulting EDFs, calculated in a region on the drain side of the gate over 100 ps, are shown on a semi-log scale in figure 4.4, where

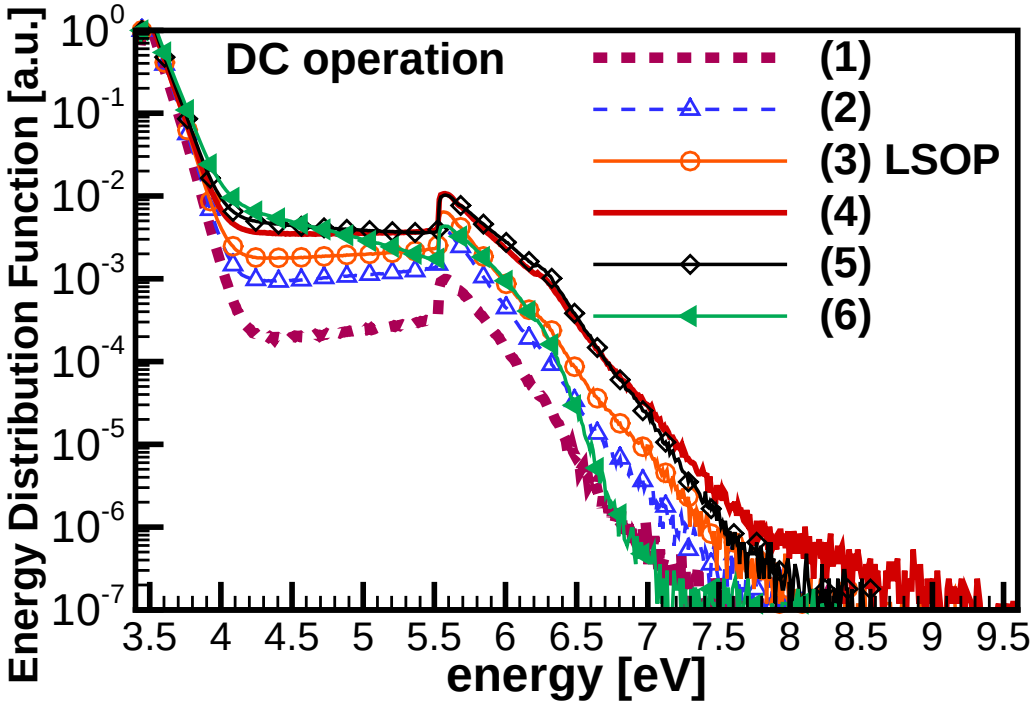

Figure 4.4: Electron energy distribution function for the DC bias points (1) to (6) indicated in figure 4.3 (a) with solid dots.

the 0 eV energy corresponds to the top of the valence band, and carriers at the bottom of the conduction band have an energy of 3.41 eV equal to the band gap of GaN.

For all bias points the EDF is a nearly Maxwellian distribution only for low energies ($E < 4\,eV$), then a significant tail of high energy electrons is observed, including the transfer from the Γ point to satellite valleys at approximately 5.5 eV, evidencing the highly non-linear transport effects of high-energy carriers. The tail of hot-carriers reaches a minimum in the off-state point (1) for V_{GS} = -7 V, and it increases with V_{GS}, reaching its peak at V_{GS} = -2 V, V_{DS} = 16 V shown with curve (4), which is known as the semi-on condition, where high electric fields and moderate carrier densities are simultaneously present [59]. As V_{GS} continues to increase until it becomes positive as in curve (6), known as the on-state, the drain current increases but the lateral electric field is lowered, making the tail of hot electrons decrease substantially, even below the value corresponding to the LSOP (3).

The resulting energy distribution functions under DC conditions can be explained in terms of the electron motion within the device. In particular, figures 4.5 (a), (b) and (c) show a close-up of the profile along the channel of the x-component of the electric field E_x, the concentration of electrons and the total scattering, respectively. From the figures it can be seen that bias point (4), which corresponds to the peak of hot electron generation, presents a high electric field peak in the channel, in addition to a high concentration of electrons which are accelerated as they transit under the gate gaining significant amounts of kinetic energy that is not transferred to the lattice evidenced by the low scattering hence becoming hot carriers.

When the device operates under bias point (6), which corresponds to the peak of drain current of figure 4.3, a high concentration of electrons is present in the channel. However, the low electric field and the high occurrence of scattering under the gate prevents the carriers from gaining much energy which translates into a lower concentration of hot carriers as compared to bias point (4) as indicated in figure 4.4. Finally, under bias point (1) the device presents the highest peak electric field under the gate, but since the channel is depleted of electrons (hence the discontinuity in plots (b) and (c) which are in semi-log scale), there are no carriers to be accelerated. It most be noticed that even though the channel is fully depleted of carriers, the buffer has a low concentration of electrons which drift from source to drain due to the high electric field, producing a leakage current. In the process they gain kinetic energy becoming hot, which explains the non-zero EDF for bias point (1) shown in figure 4.4. This analysis suggests that points (4) and (3) constitute the appropriate upper and lower EDF comparative values, respectively, between DC and RF modes of operation.

Aiming to obtain a statistically significant EDF under RF operation at 10 GHz, simulations were run for 200 ps or 2 full cycles of RF signal with $8\text{x}10^4$ super-particles. The EDF was built during the last simulated cycle by updating the field every 1.5 fs

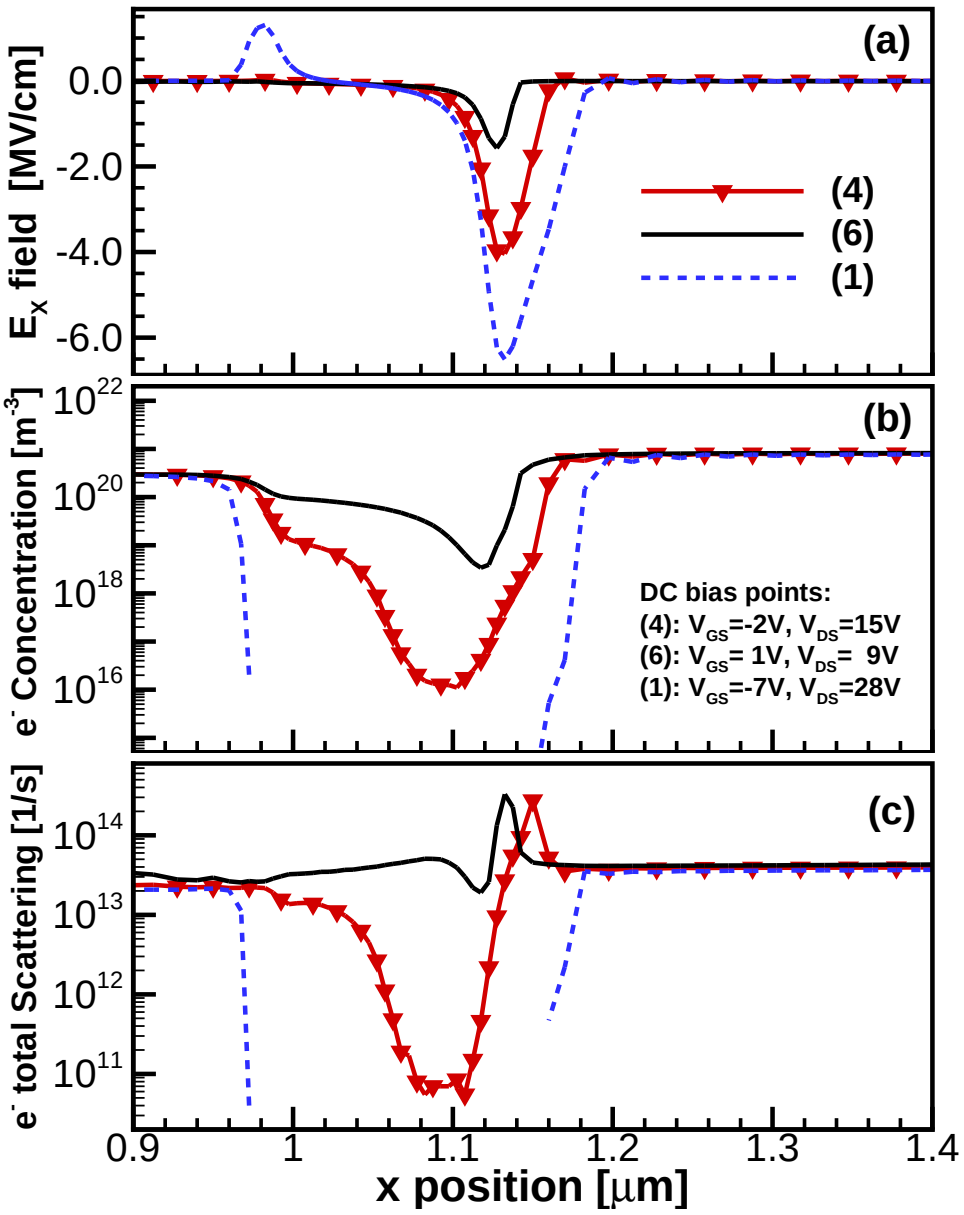

Figure 4.5: Close up under the gate of the profile along the channel of (a) x- component of the Electric field E_x, (b) electron concentration and (c) total scattering at DC bias points (4), (6) and (1) from figure 4.4, which trace the load-line in figure 4.3 (a).

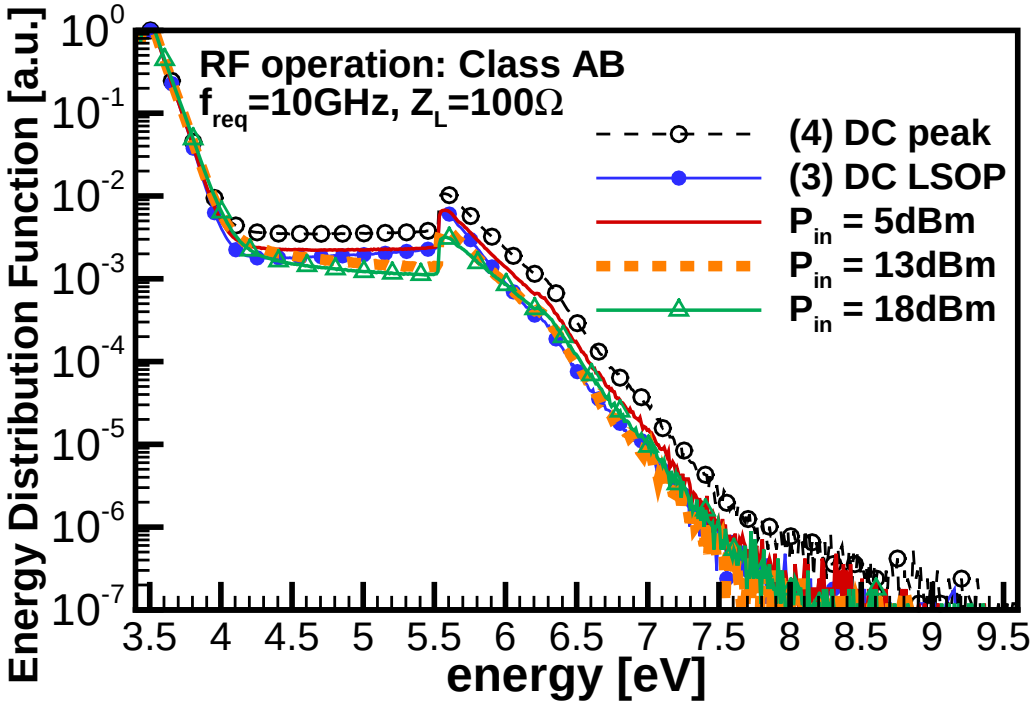

Figure 4.6: Electron energy distribution function under RF Class AB operation for different input powers. Curves for DC peak generation (4) and LSOP (3) from figure 4.4 also shown for comparison.

(Poisson solver time step) and normalized to the number of iterations. Figure 4.6 presents the resulting EDFs at three different input powers of 5 dBm, 13 dBm and 18 dBm, along with the DC comparative curves. It can be seen that the peak RF distribution is reached at 5 dBm, which corresponds to a load-line excursion around the semi-on DC condition. However, the hot electron generation is 43 % lower than the DC peak for energies higher than 5.5 eV. At lower energies the DC LSOP curve sets the lower limit for the peak RF distribution.

As the input power increases, the swing of the dynamic load line takes the device away from the semi-on state, into the on-state and off-state represented by the DC operating points (6) and (1) respectively, which in turn results in a hot electron concentration 39 % and 49 % lower than the DC LSOP case for P_{in} = 13 dBm and 18 dBm respectively, for $E < 6.3$ eV. At higher energies, the RF distribution matches

the LSOP case at 13 dBm and it goes above it by 19 % in deep compression. In terms of reliability, these results suggest that devices operating under DC electrical stress are potentially more susceptible to hot-electron degradation effects than when operating under Class AB RF power amplifier at low frequencies, since a higher concentration of energetic carriers can create electrically active traps.

4.5 High Frequency RF Operation

Since the studied device has a cut-off frequency of 104 GHz, the analysis was extended by performing simulations under Class AB operation up to 40 GHz, keeping the LSOP and tuned-load Z_{Load} constant. The characterization of the device at the two frequencies of $f_{req} = 10$ GHz and 40 GHz is shown in figures 4.7 (a) and (b), corresponding to the load-lines and typical figures of merit, respectively. From figure 4.7 (a) it can be seen that at the high frequency of 40 GHz the load-line deviates from the purely resistive behavior (straight line) observed at 10 GHz, due to the higher output capacitance C_{DS} which in turn presents a low-impedance difficult to perfectly match for tuned-load conditions. Since the load impedance is not perfectly matched, some of the power is delivered to higher order frequency harmonics rather than to the fundamental one. The lost of power towards harmonics reduces the PAE and gain at high frequency as shown in figure 4.7 (b), in comparison with the performance observed at 10 GHz. Nevertheless, the power amplifier still presents high output power which makes it viable for mm-wave operation.

In terms of hot-electron generation, the energy distribution functions at three different frequencies are shown in figures 4.8 (a) and (b) for two values of output power P_{out} = 26 dBm and 30 dBm, respectively. For low output power 26 dBm below the compression point, figure 4.8 (a) shows that the hot carrier distribution increases with the frequency of operation, approaching the DC peak generation for all energies.

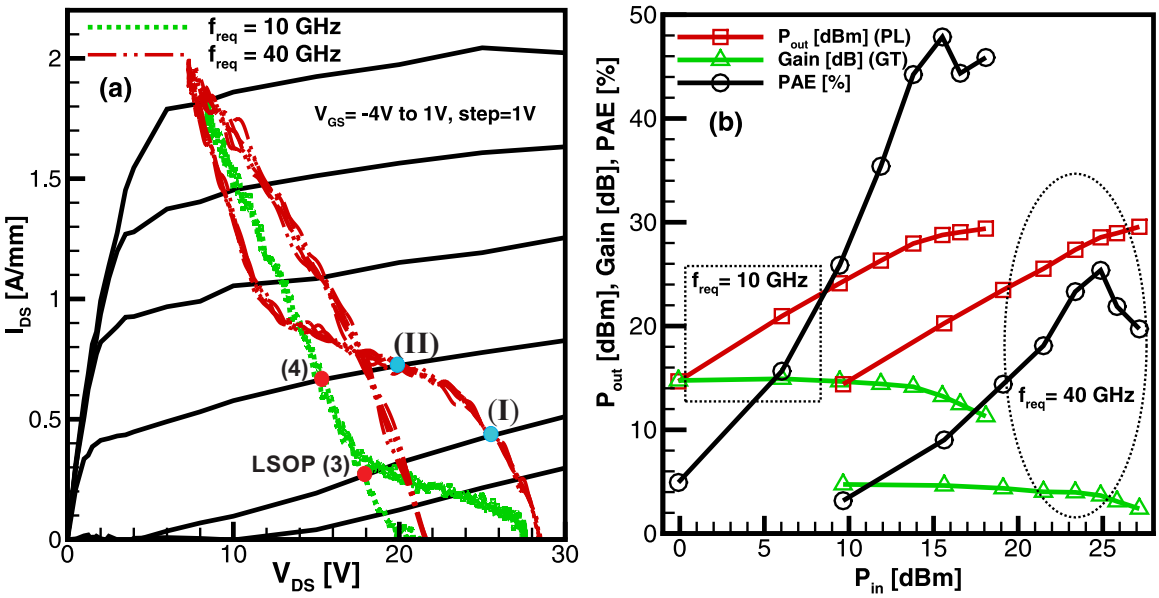

Figure 4.7: Frequency study of power amplifier characterization under Class AB (RF) operation, showing (a) dynamic load-lines and (b) FOM, at frequencies of $f_{req} = 10\,\text{GHz}$ and 40 GHz and tuned-load $Z_{Load} = 100\,\Omega$.

In particular, at 25 GHz the EDF is 45 % lower than the DC peak as compared to 64 % at 10 GHz for $4 < E < 7\,\text{eV}$, whereas at 40 GHz the RF distribution matches the DC counterpart, even surpassing it by 25 % for energies higher than 6.5 eV. As the PA is driven into compression as shown in figure 4.8 (b), it is observed that unlike the X-band case, hot carrier generation increases with the input power (or output power) at higher frequencies, particularly for energies higher than 6 eV, where operation at 25 GHz matches the DC peak generation, meanwhile at 40 GHz it is 75 % higher than in DC.

These results imply that hot electron degradation under RF power amplifier operation depends not only on the input power but also in the frequency of operation. Furthermore, the results suggest that at high frequency and high power the degradation of devices due to hot-electron effects is stronger than that under DC conditions, which is in good agreement with RF degradation experiments [94]. From a physics point of view, the increment in hot-carrier concentration is due to the high output

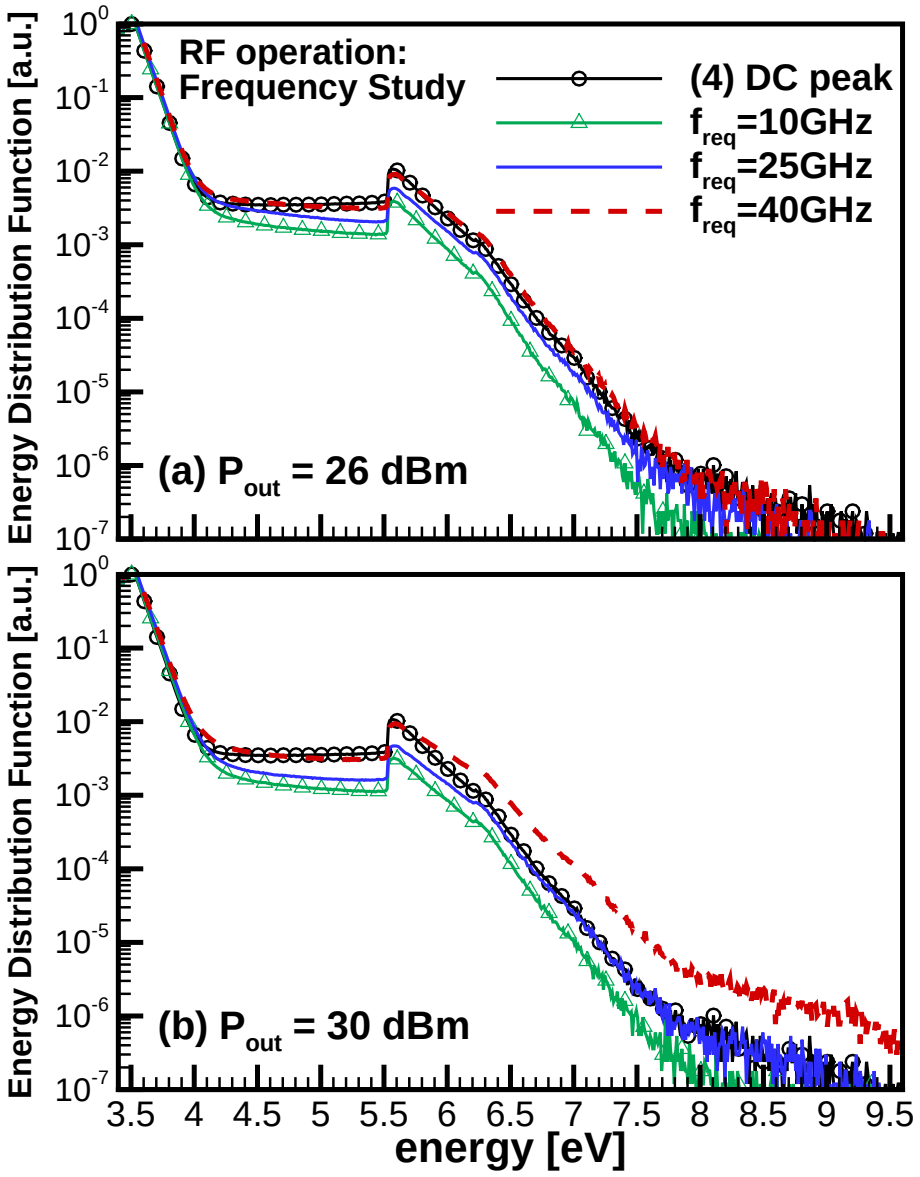

Figure 4.8: Frequency study of the energy distribution function for Class AB operation at (a) P_{out} = 26 dBm and (b) P_{out} = 30 dBm.

capacitance which causes the broadening of the load-lines as the frequency and power increase. The dynamic load-line swing under these conditions broadens, deviating from a straight line into an ellipse as shown in figure 4.7 (a), pushing the device into a high-power state where high current and high electric field are present simultaneously. This condition results in higher generation of hot electrons even in DC, as demonstrated in figure 4.9 showing the EDFs computed for DC bias points indicated in figure 4.7 (a) as (I) and (II) located on the 40 GHz load-line. Also shown in figure 4.9 for comparison the EDFs for DC bias points (3) and (4) corresponding to the load-line at 10 GHz. It can be seen that the points located on the 40 GHz load-line result in higher generation of hot electrons with respect to the points located on the 10 GHz load-line. Experimentally, this state is often studied under RF pulsed conditions only, because in DC operation the high power dissipation can lead to device destruction.

4.6 High Temperature Operation

The effects of temperature on the performance of the device under RF conditions were also studied for Class AB operation at $f_{req} = 10$ GHz by means of isothermal simulations. In this approach, the corresponding scattering rates are calculated at a constant lattice temperature. The DC and large-signal power amplifier characterization results for $T = 300$ K and 500 K are shown in figure 4.10, depicting the dynamic load-lines along with the I_{DS}-V_{DS} curves in (a) and the RF figures of merit in (b). Concerning the DC performance, it can be seen that as the temperature increases from 300 K to 500 K, the DC saturation current I_{Dsat} at $V_{GS} = 1$ V goes from 1.8 A/mm to 1.4 A/mm, representing a 23 % reduction, whereas R_{on} increases by the same amount. On the other hand, at low V_{GS}, the degradation is only significant for high V_{DS}.

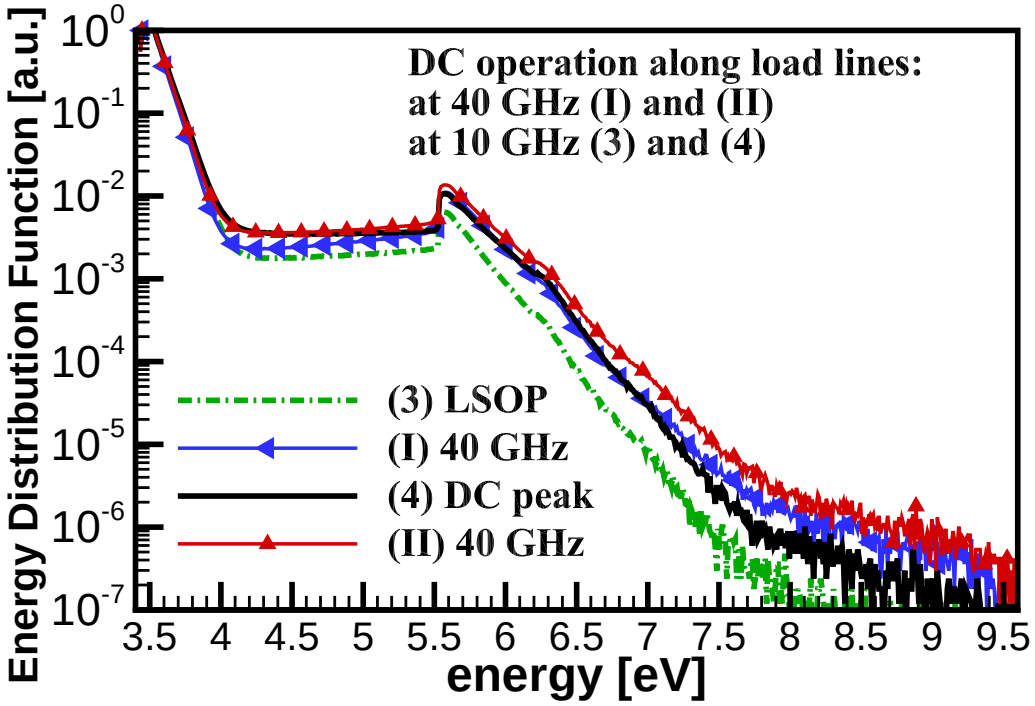

Figure 4.9: Electron energy distribution function for the DC bias points (I), (II) and (3), (4) indicated in figure 4.7 (a) with solid dots, corresponding to DC points along the load-lines at 40 GHz and 10 GHz respectively.

The DC performance degradation can be explained in terms of changes in transport quantities such as the carrier velocity and scattering rates due to increased phonon scattering in the 2DEG channel at high temperatures [100]. This is presented in figures 4.11 (a) and (b), showing the total electron scattering rate and the x-component of the carrier velocity, respectively, obtained at $T = 300$ K and 500 K in a region under the gate along the channel. As the temperature increases, the total scattering increases due to the thermal energy gained by the crystal atoms. This in turn reduces the electron velocity and hence the current.

In terms of RF performance, at 500 K the power gain and the output power drop by 4.31 dB (27 %) and 4.3 dBm (21 %), respectively, significantly degrading the PAE by 40 % from $P_{in} = 0$ dBm to 15 dBm, whereas in compression (or high input power) the degradation is less than 5 %. It should be noted that the degradation of DC and

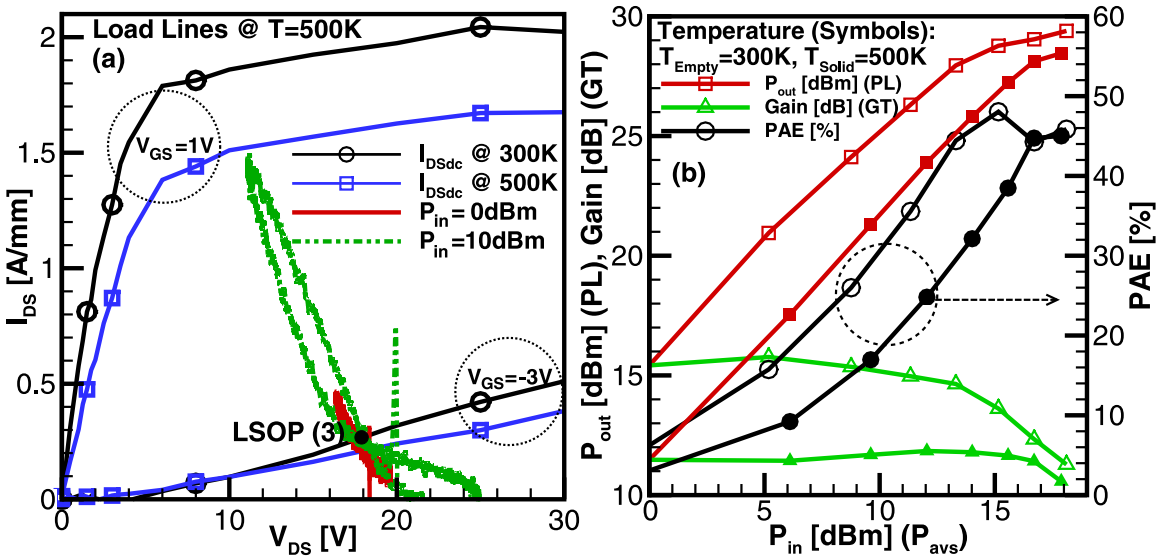

Figure 4.10: Effect of temperature in Class AB operation for f_{req} = 10 GHz and Z_{Load} = 100 Ω showing (a) I_{DS}-V_{DS} curves along with load-lines, and (b) figures of merit.

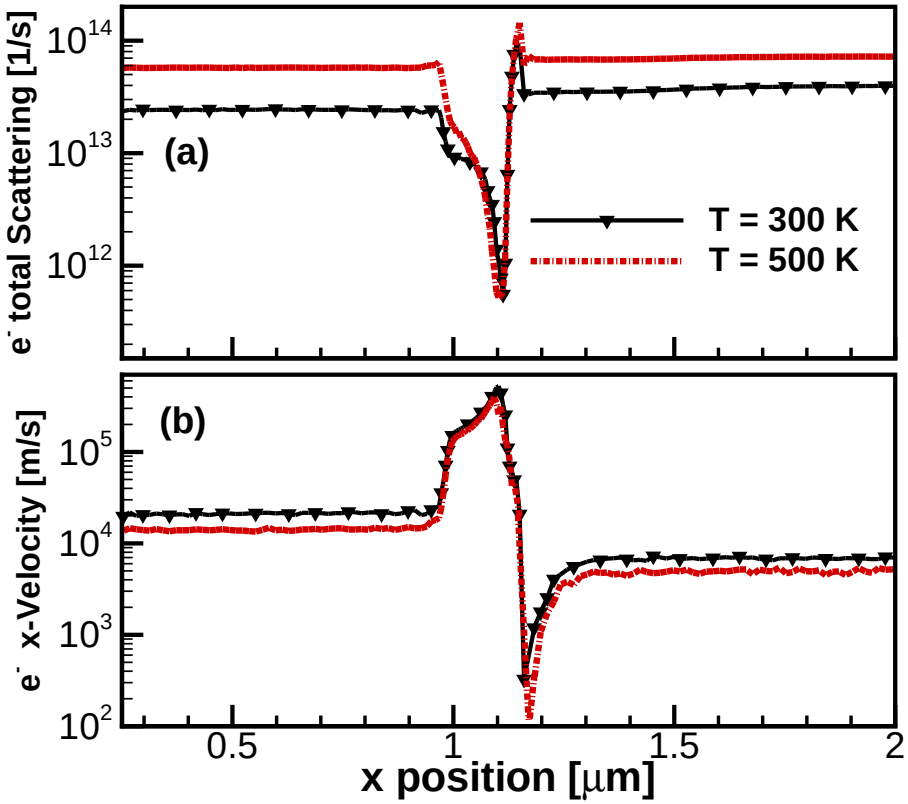

Figure 4.11: Effect of temperature on (a) the total scattering rate and (b) the x-component of the electron velocity, shown in a region under the gate along the channel.

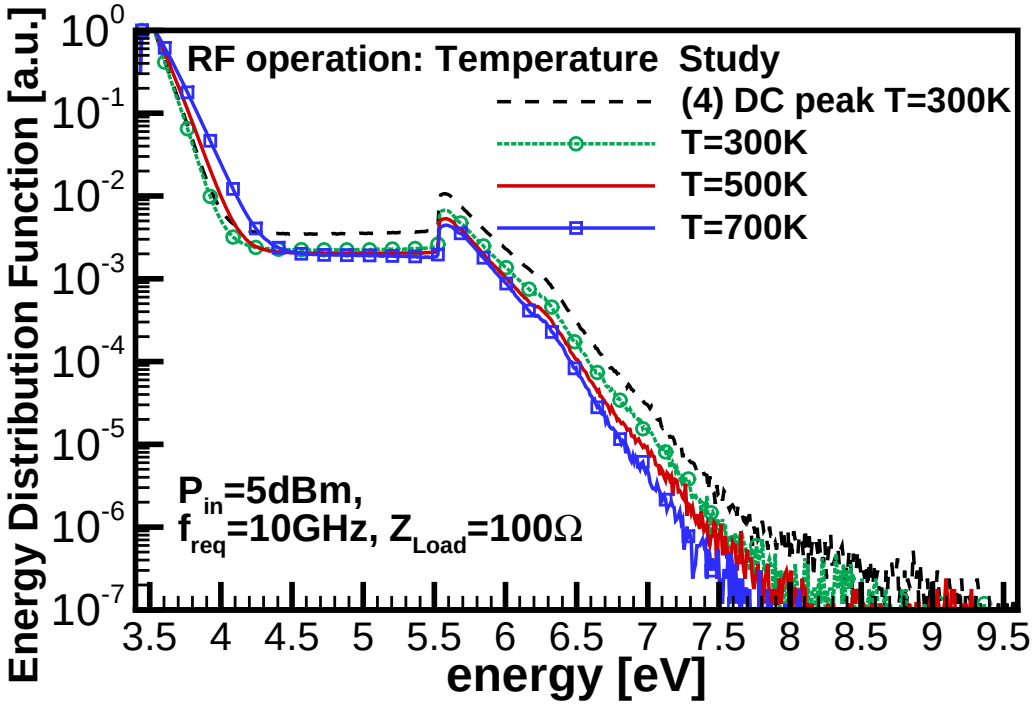

Figure 4.12: Temperature dependence of the distribution function for Class AB operation at $P_{in} = 5\,\text{dBm}$, for $f_{req} = 10\,\text{GHz}$ and $Z_{Load} = 100\,\Omega$.

RF performance are tightly correlated. As DC parameters I_{Dsat} and R_{on} degrade, the swing of the dynamic load-line is reduced, as shown in figure 4.10(a), where the 10 dBm load-line at 500 K (dotted line) shows a smaller excursion than that at 300 K shown in figure 4.3 (a). This in turn reduces P_{out} and the PAE of the amplifier.

Finally, the RF electron EDF up to 700 K is shown in figure 4.12 for $P_{in} = 5\,\text{dBm}$. In RF regime, the distribution function at low energies ($E < 4.5\,\text{eV}$) increases with the temperature, being 30 % and 72 % higher than the DC peak at 300 K, for $T = 500\,\text{K}$ and 700 K, respectively. This implies higher trap generation for defects with low activation energies as suggested by Mukherjee [91]. On the other hand, the tail of high energy electrons ($E > 5.5\,\text{eV}$) consistently decreases with temperature, due to increased phonon scattering at high temperatures [100]. For all cases in the X-band, the RF hot electron generation is lower than the DC peak at 300 K, regardless of the input power or the temperature variations.

4.7 Comparison between Class A and Class AB Operation

In terms of power amplifier design, the Class AB topology is by far the most widely used design both as a stand-alone PA and as a fundamental building block of other PA topologies for both wireless and base station applications due to its high efficiency and low thermal dissipation. Nevertheless, because of its high harmonic content which can only be reduced to a certain extent through high-Q output matching networks, it has the disadvantage of relatively lower linearity when compared to other topologies. In particular, Class A power amplifiers constitute not only the most simple of all PA circuits, but also offer high linearity, high gain and high output power, along with inherently low harmonic distortion and no cross-over [73].

The reason for the higher linearity relies on the fact that for Class A operation, the DC bias point of the device is chosen so that it is always on for all the load-line excursion, which also allows to bias the device near the center of the most linear portion of the transconductance curve. This in turn produces inherent disadvantages such as low conversion efficiency and high thermal dissipation due to the high DC power consumption. For these reasons, the hot-carrier generation study is extended in this section to compare the reliability of devices operating in Class A and Class AB power amplifier topologies.

The analysis begins with a comparison of the PA characterization in terms of the load-lines excursion and the typical figures of merit as shown in figures 4.13 (a) and (b) respectively at $f_{req} = 10\,\text{GHz}$ and $Z_{Load} = 100\,\Omega$ (tuned with a high-Q OMN for the Class AB). From figure 4.13 (a) it can be seen that for Class A operation the LSOP corresponds to a high DC power of 27 W/mm whereas for Class AB the DC power is only 4.5 W/mm. On the other hand, the load-lines excursion for Class A show a purely resistive behavior even at high input power evidencing a negligible

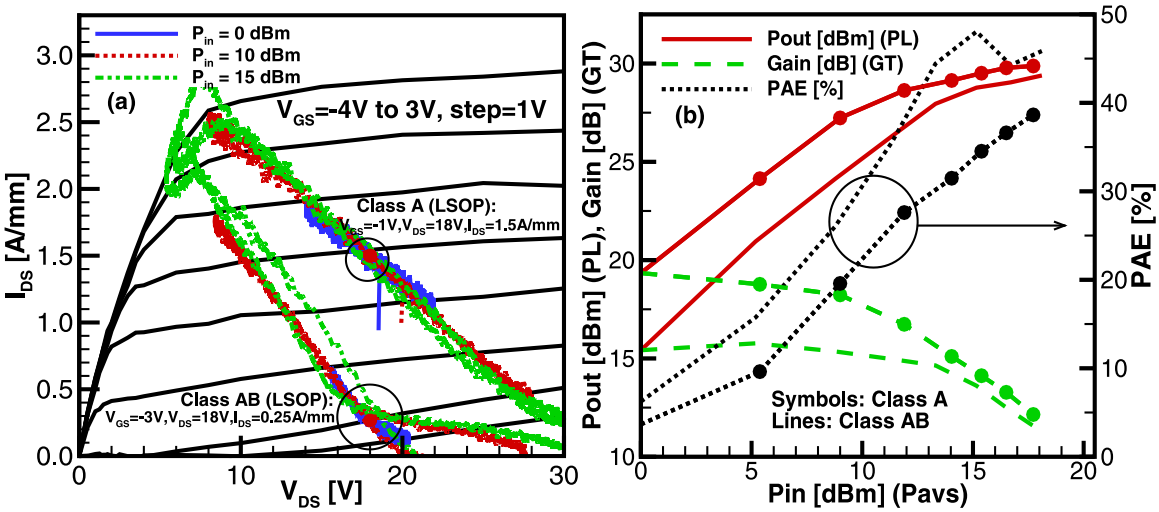

Figure 4.13: Power amplifier characterization under Class A and Class AB operation for $f_{req} = 10\,\text{GHz}$ and tuned-load $Z_{Load} = 100\,\Omega$, showing (a) dynamic load-lines along with I_{DS}-V_{DS} curves. Also, the typical figures of merit are shown in (b) where lines with and without symbols correspond to Class A and Class AB respectively.

harmonic content, which in turn allows for higher linearity, as compared to the Class AB topology. In terms of the figures of merit shown in figure 4.13 (b), the main disadvantage of Class A operation is evidenced as a low PAE as compared to that of the Class AB amplifier, due to the difference in DC power. Nevertheless, higher output power and gain are obtained in Class A PAs.

Concerning the hot-carrier concentration, the electron energy distribution functions under Class A at 10 GHz were obtained running large-signal simulations for 200 ps or 2 full cycles of RF signal with 80 K super-particles, as it was indicated before for Class AB operation. Figure 4.14 presents the resulting EDFs at 3 different input powers $P_{in} = 0\,\text{dBm}$, 10 dBm and 15 dBm for Class A (lines with symbols) and Class AB (lines). It can be seen that for energies lower than 4 eV all curves show a nearly Maxwellian distribution, however, a large tail of high energy electrons is observed including the transfer to satellite valleys at 5.5 eV. Under Class A operation, the peak EDF is reached at the lowest input power $P_{in} = 0\,\text{dBm}$ and then it decreases monotonically as P_{in} increases, in contrast to what was previously described for Class

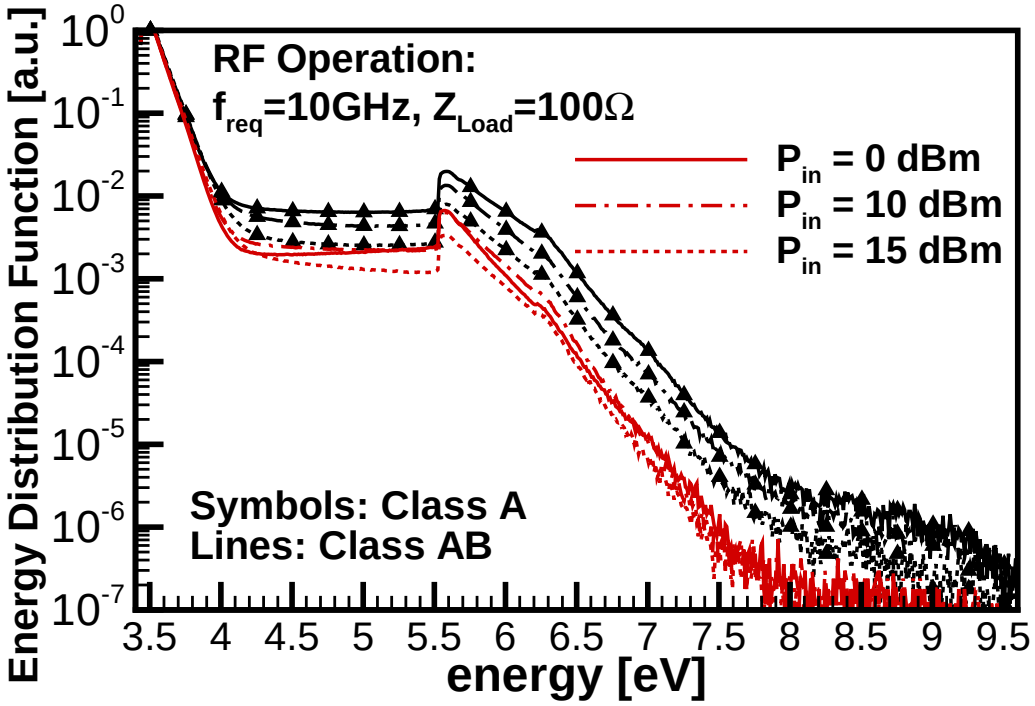

Figure 4.14: Electron energy distribution function under Class A (lines with symbols) and Class AB (lines) PA operation, at 3 input powers P_{in} = 0 dBm, 10 dBm and 15 dBm.

AB conditions where the EDF first increases with P_{in} reaching a peak at 5 dBm and then decreases with the input power, reaching its minimum at 15 dBm.

At low P_{in}, the device operates around the LSOP as shown in figure 4.13 (a), which in Class A presents both high current I_{DS} and high electric field E_x simultaneously, allowing for a high hot-carrier generation as compared to Class AB where the LSOP is close to an off-state of low I_{DS}. As the input power increases, under Class A the load-line excursion pushes the device towards regions where only either I_{DS} or E_x is high, reducing the hot-carrier generation. In contrast, under Class AB the EDF first increases with P_{in}, reaching its peak when I_{DS} and E_x are maximized, which is achieved in the semi-on region. From this point, as P_{in} keeps increasing, the EDF in Class AB starts decreasing due to the same reasons explained for Class A. It most be noticed that for all P_{in}, the hot electron concentration is nearly one order of

magnitude higher under Class A than in Class AB because for all possible values of E_x along the load-lines, the current I_{DS} is always higher in Class A than in AB.

Simulations at high frequency of 40 GHz were performed for Class A operation, while keeping the LSOP and Z_{Load} constant. A comparison of the EDFs obtained at 10 GHz and 40 GHz under Class A is shown in figure 4.15 for two input powers of P_{in} = 0 dBm and 18 dBm, showing that the hot-electron generation is higher at 40 GHz than at 10 GHz for a given P_{in}. Furthermore, the global peak hot-carrier generation is observed at 0 dBm for 40 GHz which is 27 % higher than that at 10 GHz throughout all the energy range. As the input power increases, the EDF decreases as it was observed at low frequency. However, when the device is driven into compression corresponding to the 18 dBm curves, the EDF is higher at 40 GHz than at 10 GHz, in particular an increment of 60 % is observed in the energy range of $4 < E < 6.5$ eV, whereas the EDF is 273 % (or 2.7 times) higher for energies higher than 7 eV, so that the 18 dBm curve reaches the peak generation obtained at 0 dBm. As it was described for the Class AB topology, the higher hot-carrier concetration at high frequency operation is due to the broadening of the load-lines caused by the higher capacitance of the device at high frequency, which presents an low impedance difficult to tune-out.

Finally, a comparison of the hot-electron generation under Class A and Class AB operation at high frequency is presented in figure 4.16, where curves of the electron EDFs are presented at 40 GHz for the input powers of P_{in} = 0 dBm and 18 dBm. In this case, the largest difference between the EDFs under Class A and Class AB conditions is observed for 0 dBm, corresponding to approximately one order of magnitude for all the energy range. However, since under Class AB operation at high frequency the EDF increases with the input power contrary to what is observed at low frequency, the peak hot-electron generation under Class AB reached at 18 dBm is practically the same as that calculated under Class A for all the energy range.

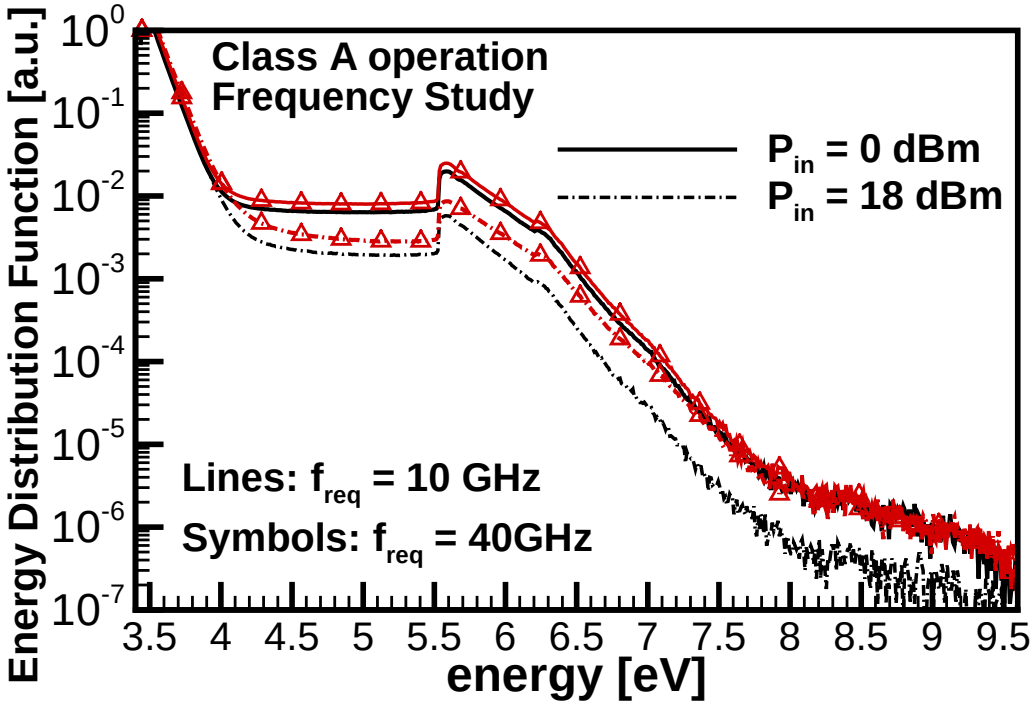

Figure 4.15: Frequency study of the electron energy distribution function in Class A operation for f_{req} = 10 GHz (lines) and 40 GHz (lines with symbols) at input power of P_{in} = 0 dBm and 18 dBm.

This analysis suggests that devices are more susceptible to hot-electron induced degradation when operating under Class A power amplifiers than with Class AB topology. This correlation is observed at all input powers when the frequency of operation is kept in the x-band. However, at mm-wave frequencies and high power operation, such as 40 GHz and P_{in} =18 dBm both topologies are expected to induce comparable hot-carrier degradation.

4.8 Conclusions

In this chapter, a hot electron generation study was presented in terms of the electron energy distribution function, calculated under large-signal RF, Class AB operation in GaN HEMTs. The results, consistent with experimental studies, suggest that under ideal conditions in the X-band with f_{req} = 10 GHz, hot-electron degrada-

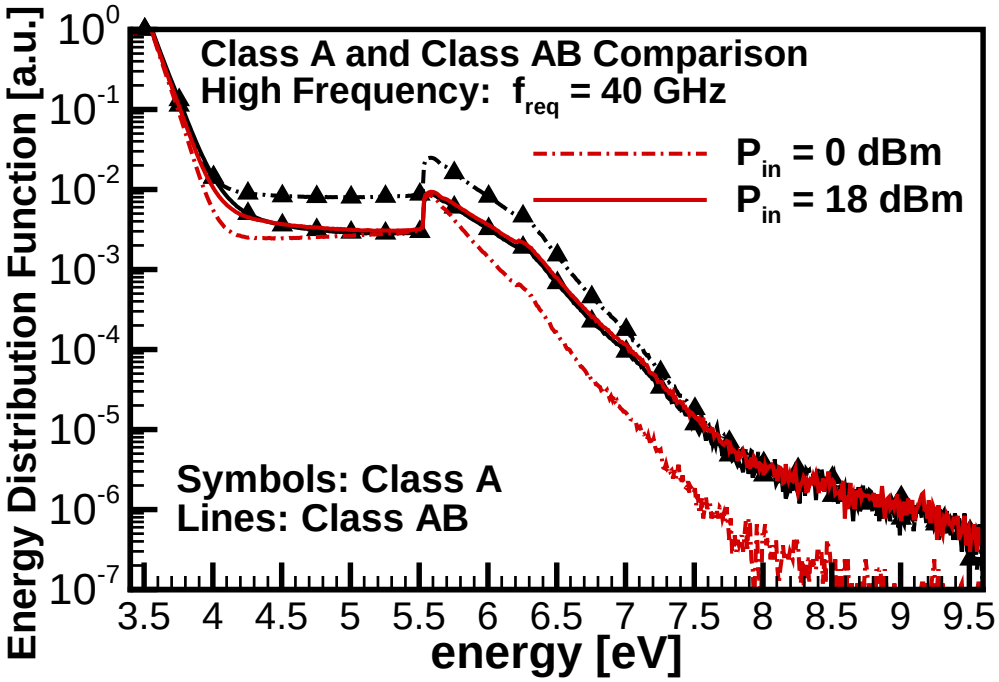

Figure 4.16: Comparison of electron energy distribution function at high frequency operation, under Class A (lines with symbols) and Class AB (lines) topologies at input power of P_{in} = 0 dBm and 18 dBm.

tion effects should be lower in RF drive than in DC, regardless of the input power or temperature of operation. However, in the mm-wave band up to 40 GHz, hot carrier generation matches the DC counterpart, even surpassing it at high energies when the power amplifier is driven into deep compression, suggesting stronger degradation under RF than in DC. Regarding the topology of the PA, it was shown that devices operating in Class A produce higher concentration of hot carriers regardless of the input power as compared to Class AB for frequencies in the X-band, whereas in the mm-wave regime Class A and AB topologies present comparable concentrations of hot-carriers particularly when the PA is driven into compression. Finally, in terms of reliability the RF analysis is more meaningful for realistic lifetime estimation given the practical application of these devices as mm-wave PAs, and the accurate calculation of the EDF under RF can be exploited in compact modeling-based lifetime and reliability analysis as it has been done for DC operation.

Chapter 5

A Π-SHAPED GATE DESIGN FOR REDUCING HOT ELECTRON GENERATION IN GAN HEMTS

The use of a Π-shaped gate structure is proposed for GaN HEMT devices, which effectively reduces the hot electron generation under all regimes of operation, while preserving device performance well into the lower millimeter-wave frequency range. Simulations in DC and large-signal RF conditions of the proposed Π-gate device, along with the corresponding electron energy distribution functions (EDFs), were obtained with a full-band Cellular Monte Carlo device simulator self-consistently coupled to a harmonic-balance circuit solver, and compared to simulations of a typical T-gate HEMT whose DC curves were calibrated to experimental data. The results show that the peak hot-carrier generation obtained with an asymmetric-Π-gate is about 41 %, 44 % and 75 % lower at DC and in Class AB mode at 10 GHz and 40 GHz, respectively, as compared to that observed in the T-gate devices. This new gate structure suggests that significantly higher reliability against hot electron induced device damage can be achieved with modest impacts on performance.

5.1 Introduction

The superior electrical and thermal properties of GaN based HEMTs have enabled their use in high-frequency Power Amplifiers (PAs). However, reliability issues are still a reason of concern and active research [101], and hot electron effects are systematically associated to degradation of device performance under both DC [91] and large-signal RF operation [94; 95; 71; 72], since highly energetic carriers with kinetic energy larger than a specific activation threshold can create electrically active

traps by dehydrogenation of passivated point defects [90]. Field-plates have been successfully adopted to reduce short-channel effects and enhance breakdown voltage by reducing the peak electric field [11; 12]. Also, slanted gate contact geometries have shown to reduce the effect of surface trap generation during off-state stress [102], and recent studies have relied on gate contact layout engineering to improve electrical device performance [103], decrease microwave noise [104] and achieve normally-off operation [105]. However, as of this writing, efforts that have been directed to reduce hot carrier generation and mitigate its effects have focused on asymmetric gate placement which affect performance [106], rather than the contact interface geometry.

With the aim of improving the reliability of GaN HEMTs by reducing the hot electron generation while preserving the DC and RF performance, a Π-shaped gate structure is proposed here as an alternative to traditional T-gate contacts. The concept is demonstrated with simulations of the DC, small signal AC, and large-signal RF characteristics performed using a full-band Cellular Monte Carlo particle-based device simulator (CMC) [37], self-consistently coupled to a harmonic-balance circuit solver [89]. In this approach, the carrier dynamics take into account scattering mechanisms due to deformation potential phonons, polar optical phonons, piezoelectric (polar acoustic) phonons, ionized impurities, impact ionization and thread dislocations. The results are compared to simulations of a GaN HEMT device with a typical T-gate contact, whose model was calibrated to experimental DC curves of a device fabricated in a MMIC process flow similar to the one reported by Fitch et al. [96].

The principles of operation of the proposed Π-gate contact are correlated to relevant microscopic quantities such as carrier velocity, average carrier concentration and total scattering rate as well as electric field distribution throughout the device, while design rules and projections of small-signal AC parameters are discussed to facilitate the design and fabrication of the proposed Π-gate contact geometry, which is com-

patible with the dielectrically defined optical T-gate process [107] and the two-step-recessed gate process [108]. The description of the highly non-linear, non-equilibrium carrier dynamics accurately captured by the CMC particle-based simulator allows to assess the impact of the gate layout on hot carrier generation by means of the accurate simulation of the electron EDFs in both DC bias and large-signal RF drive under load and drive conditions that are relevant for experiments [109]. The significant reduction of the hot carrier distribution observed with a Π-gate contact under all operating conditions suggests improved reliability of the proposed device.

5.2 Performance of Π-gated HEMTs

5.2.1 Principles of Operation

The layout of a conventional T-gate HEMT device is shown in figure 5.1 (a). In typical DC conditions, such as biasing for power amplifier operation, the device is subjected to high drain voltage V_{DS} that produces high electric fields on the drain-side of the gate depleting this region of electrons. Then, as a gate bias V_{GS} is applied to allow current flow, carriers that transit under the gate are accelerated by the high electric field gaining kinetic energy. As electrons continue traveling through the crystal, some of the gained energy is transferred to the lattice by electron-phonon scattering. However, a portion of electrons will achieve high energies expressed in fractions of an electron volt and referred to as hot electrons [29]. In order to reduce the generation of hot electrons while preserving the device performance, it is necessary to increase the number of electron-phonon scattering events which requires more carriers that undergo scattering as well as longer transit times, while keeping the peak electric field.

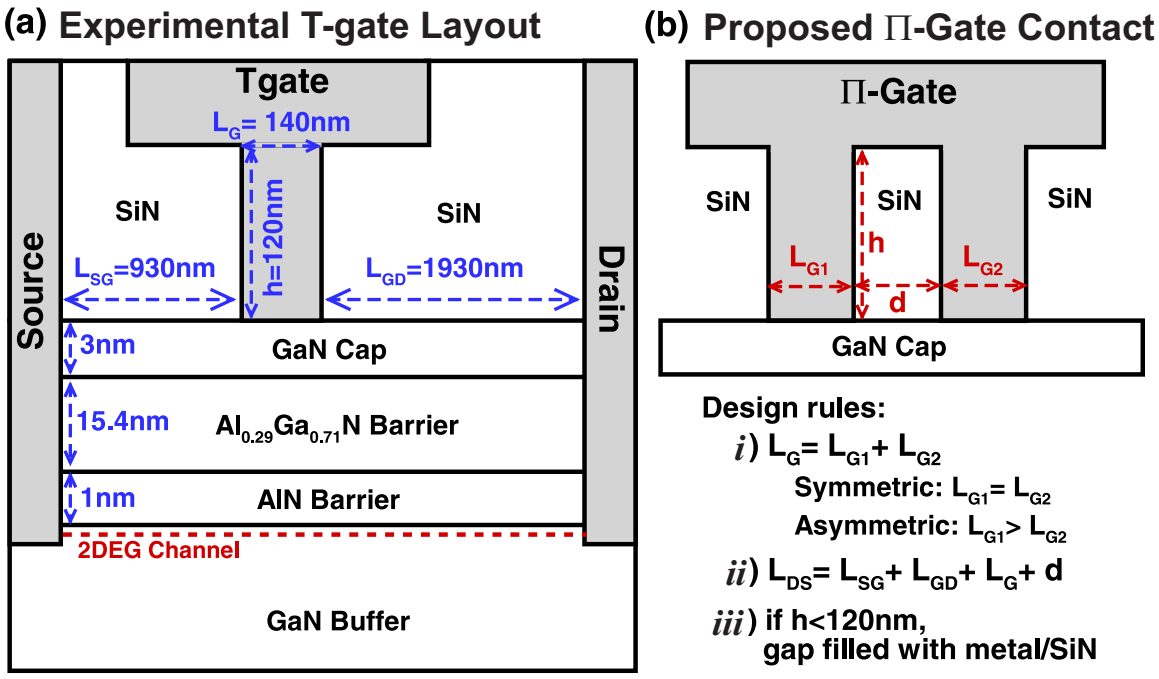

Figure 5.1: Device layout showing (a) the experimental T-gate geometry from fabricated devices (b) the proposed Π-gate, including design rules.

Such conditions can be achieved with a Π-shaped gate contact as proposed in figure 5.1 (b), which is formed by basically splitting the gate length L_G of T-gate in two stems L_{G1} and L_{G2} and adding a spacer region with length d, while keeping the rest of the access regions and the epitaxial stack structure unchanged. This gate layout effectively reduces the hot electron generation by stepping the conduction band E_C under the gate which increases the carrier concentration in this region, and by increasing the transit time of carriers under the gate with the spacer, which increments the total number of scattering events. In addition, since L_{G1} and L_{G2} are shorter than L_G in the T-gate, the Π-gate contact reduces the length of the region of high electric field on the drain-side of the gate depleted of carriers where there is nearly no scattering. Nevertheless, by keeping the length L_{GD} constant, the peak electric field is the same as that of T-gate preserving the device performance.

The discussed mechanisms require $L_{G1} >= L_{G2}$ since the opposite situation would not produce a significant change in the conduction band that modulates the carrier concentration. Furthermore, if L_{G2} were designed to be the longest stem, the region depleted of carriers would remain unchanged. However, it should be pointed out that making $L_{G2} << L_{G1}$ could make the fabrication impractical. Design rules for the Π-shaped gate contact, also shown in figure 5.1 (b), can be derived from the principles of operation as follows: *(i)* the gate length L_G is split in two stems of lengths L_{G1} and L_{G2}, which can be symmetric (Sym) with $L_{G1} = L_{G2}$ or asymmetric (Asym) with $L_{G1} > L_{G2}$, *(ii)* the two stems are separated by a spacer region with length d, designed to be long enough to allow electrons to scatter and thermalize, which increases L_{DS} while keeping L_{SG} and L_{GD} constant, and *(iii)* the stem in the spacer region is filled with a standard dielectric, SiN in this case, and its height h can be optimized to reduce the added gate capacitance and facilitate fabrication.

5.2.2 DC, small-signal AC and large-signal RF Operation

In order to demonstrate and evaluate the proposed device concept, an experimental HEMT is used as the baseline layout for design and comparison. It consists of an $AlGaN/AlN/GaN$ HEMT on SiC substrate as shown in figure 5.1 (a) with a typical T-shaped gate contact with $L_G = 140\,\text{nm}$, stem height $h = 120\,\text{nm}$, and access regions $L_{SG} = 930\,\text{nm}$ and $L_{GD} = 1930\,\text{nm}$ for a total device length of $L_{DS} = 3\,\mu\text{m}$. The CMC simulation model of the T-gate was validated with experimental DC data of the I_{DS}-V_{DS} and I_{DS}-V_{GS} curves as presented in figures 5.2 (a) and (b) respectively, showing excellent agreement at low V_{GS} where the drain current is small. However, since self-heating effects are not included, for high values of V_{GS} the current is overestimated for $I_{DS} > 0.8\,\text{A/mm}$ as it has been reported elsewhere [99; 110].

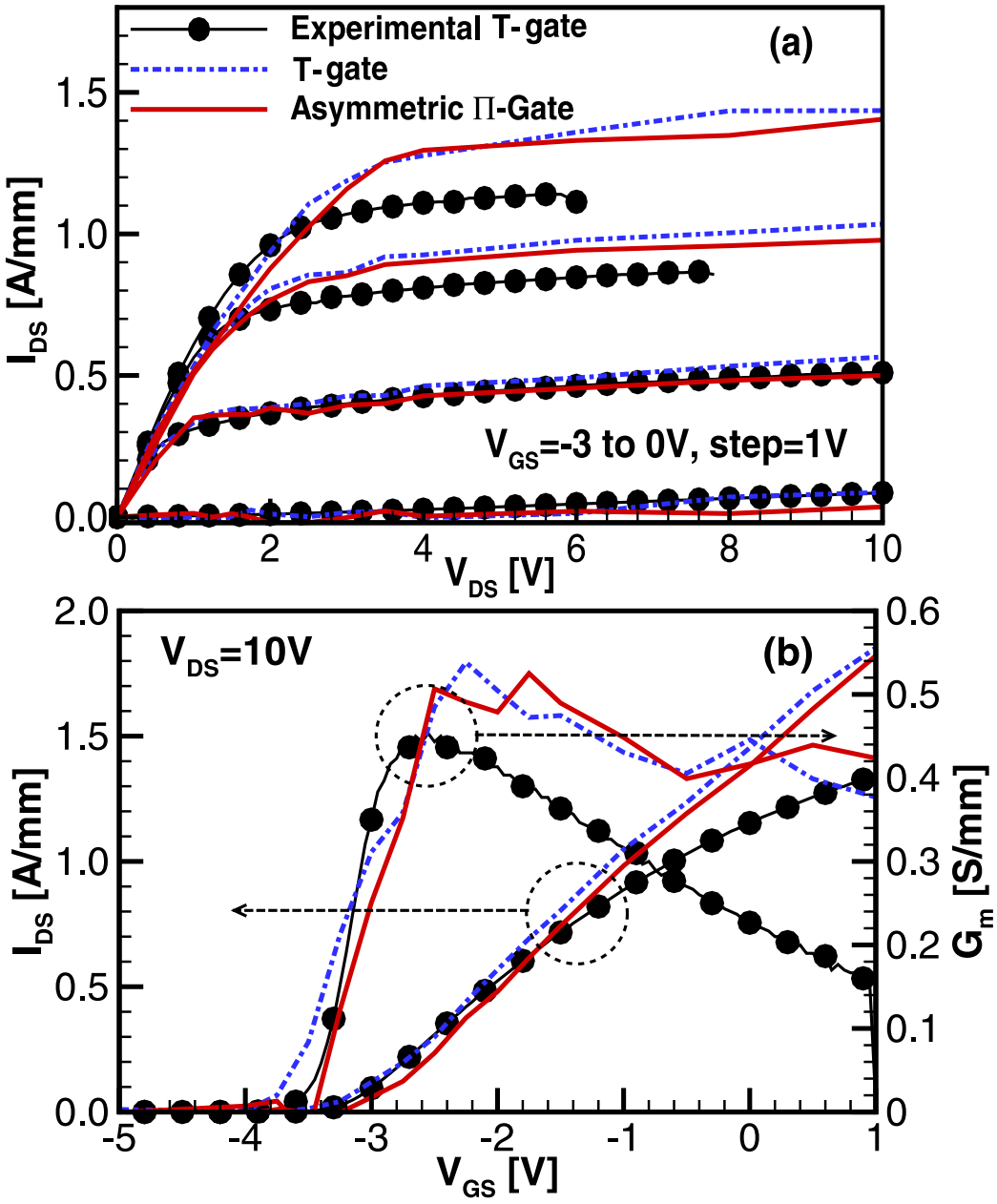

Figure 5.2: Simulated DC curves (dashed and solid lines) showing (a) output and (b) transfer (along with transconductance G_m) characteristics for T-gate and Asymmetric-Π-gate devices. Experimental data (symbols) corresponds to the T-gate device.

The Π-gate design chosen for comparison with the T-gate consists of an asymmetric contact with $L_{G1} = 100\,\text{nm}$, $L_{G2} = 40\,\text{nm}$, $d = 100\,\text{nm}$ and $h = 120\,\text{nm}$. This layout will be used for both DC and large signal RF characterization. Results in DC conditions for the asymmetric-Π-gate are also shown in figures 5.2 (a) and (b), indicating that the proposed device preserves the DC performance throughout the operating range. Nevertheless, I_{DS} is reduced $\sim$4 % at low V_{DS} due to the increased channel resistance added by the spacer length d, whereas at high V_{DS}, it decreases $\sim$10 % due to a barrier added under the gate that is modulated by the two stems of the Π-gate.

By means of small-signal AC simulations, the experimental cut-off frequency $f_t = 109\,\text{GHz}$ and maximum oscillation frequency $f_{MAX} = 234\,\text{GHz}$ were determined. In addition, figure 5.3 shows the effect on f_t and f_{MAX} of the symmetric and asymmetric Π-gate contacts, as a function of spacer region length d. As the length d increases, both parameters decrease due to the added distributed gate capacitance. The symmetric case shows better performance than the asymmetric layout for f_t with 19 % and 32 % degradation at $d = 160\,\text{nm}$ respectively, whereas f_{MAX} decreases up to 40 % for both layouts, showing nearly no reduction at short d for the asymmetric-Π-gate. Varying the stem height h in the spacer from 120 nm to 50 nm, produces the same trend with a slight improvement for the symmetric case of nearly 5 % higher f_t and f_{MAX} than the asymmetric case at $d = 100\,\text{nm}$.

The RF characterization was performed for a power amplifier in Class AB topology at frequencies in the X and K band of $f_{req} = 10\,\text{GHz}$ and 40 GHz respectively, with a large signal operating point (LSOP) of $V_{GS} = -3\,\text{V}$ and $V_{DS} = 18\,\text{V}$ with a tuned-load $Z_{Load} = 100\,\Omega$ for both the T-gate and the asymmetric-Π-gate devices. Figure 5.4 (a) shows the dynamic load lines at 12 dBm for $f_{req} = 10\,\text{GHz}$ and since it presents the same excursion for both devices, it is demonstrated that the asymmetric-Π-gate device offers the same power rating of the experimental T-gate. Moreover, the typical figures

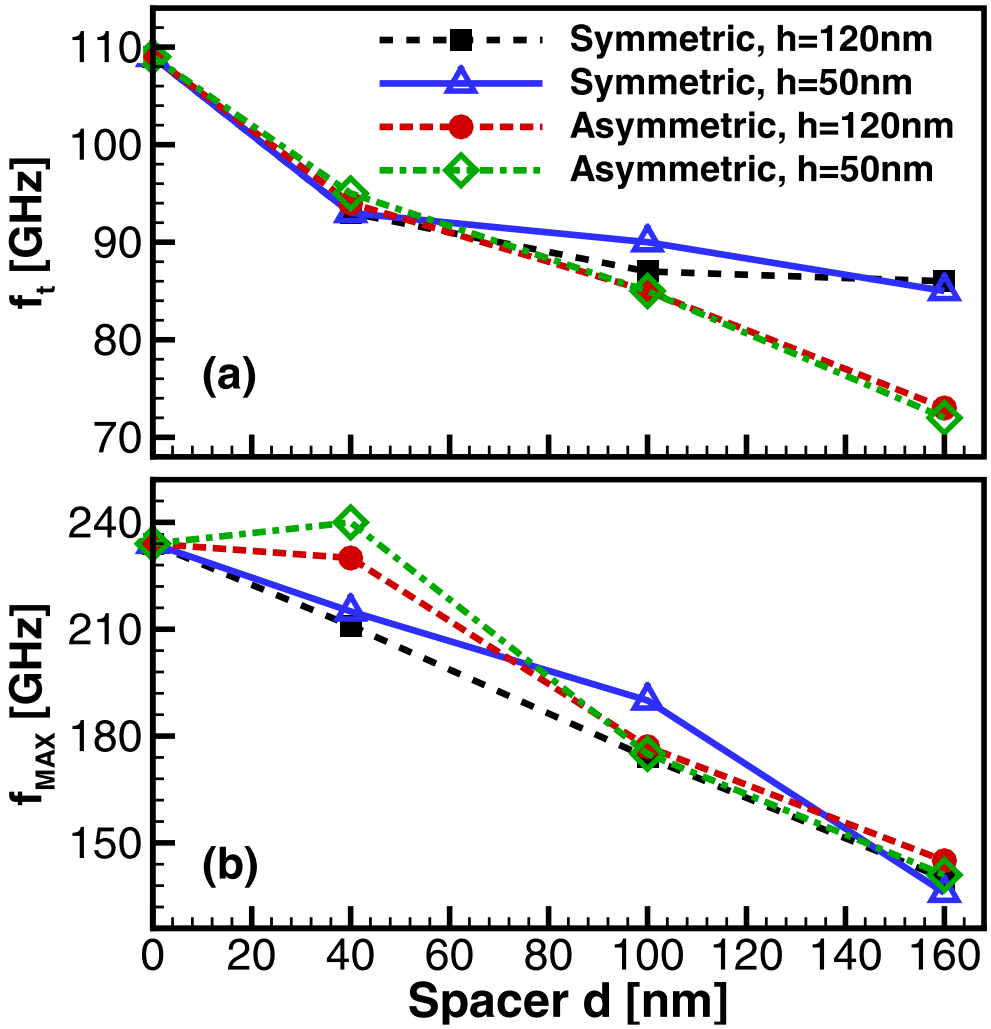

Figure 5.3: Small-Signal AC parameters (a) Cut-off frequency f_t and (b) Maximum Oscillation Frequency f_{MAX} as a function of spacer region length d for T-gate ($d = 0$), symmetric-Π-gate ($L_{G1} = L_{G2} = 70$ nm) and asymmetric-Π-gate ($L_{G1} = 100$ nm, $L_{G2} = 40$ nm), corresponding to two stem heights h.

of merit (FOM) shown in figure 5.4 (b) indicate a 14 % higher power added efficiency (PAE) for the asymmetric-Π-gate at low P_{in} as compared to the T-gate, due to the lower DC drain current at the LSOP measured at $V_{DS-LSOP}$ in figure 5.4 (a), which reduces the DC power. It must be noticed that the asymmetric-Π-gate device exhibits stronger electrostatic control than the T-gate, evidenced by lower drain current at low V_{GS} as the drain voltage increases. In particular, figure 5.4 (a) shows that for low $V_{GS} = -3\,\mathrm{V}$ the current I_{DS} at high $V_{DS} = 30\,\mathrm{V}$ of the asymmetric-Π-gate is 44 % lower than that of the T-gate, which can be explained by the modulation of the conduction band under the Π-gate added with the two stems.

This effect is shown in figure 5.5 (a) which depicts a close-up under the gate of the profile of E_C along the channel, where it can be seen that the conduction band is stepped and that in the case of high V_{DS}, an extra barrier is added under the Π-gate region, while the peak electric field remains constant in both devices as indicated by the slope of E_C for both curves. The additional barrier under the Π-gate reduces the x-component of the electron velocity in the channel, as shown in figure 5.5 (b), which translates into lower drain current. Since field-plates reduce short-channel effects by reshaping the peak electric field [11], a field-plated Π-gate design could be optimized for enhanced break down voltage by combining the effects of an additional barrier with the reshaping of the peak electric field.

5.3 Reduced Hot Electron Generation

The generation of high energy carriers is studied through the accurate simulation of the electron EDF under both DC bias and large-signal RF drive, which enables a qualitative description of hot-electron induced defect generation associated with degradation of device performance. In all the simulations, the EDF was calculated in a region on the drain side of the gate extending vertically from the AlN barrier

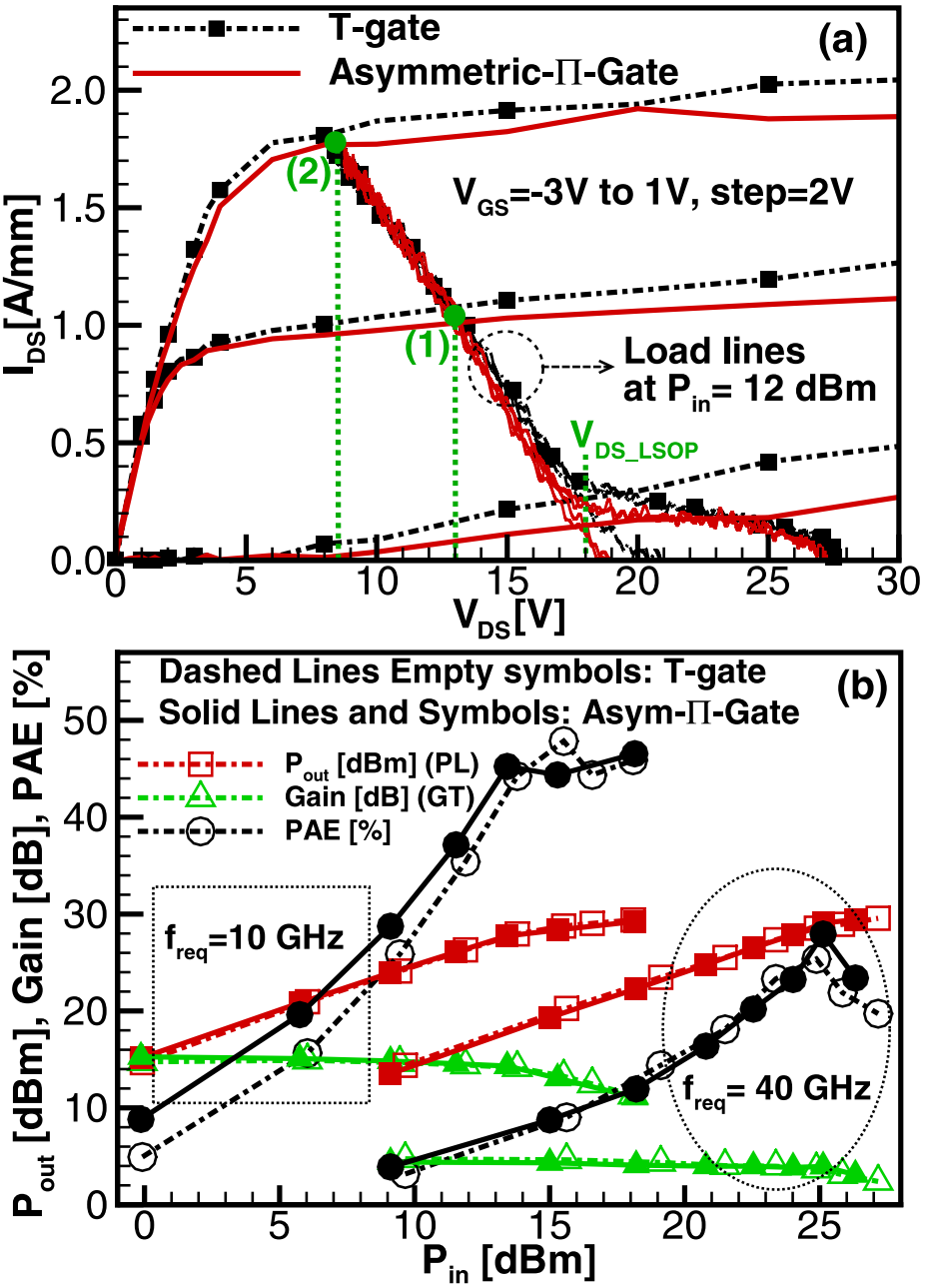

Figure 5.4: Power amplifier characterization under Class AB (RF) operation, showing (a) dynamic load-lines for P_{in}=12 dBm at f_{req}=10 GHz along with $I_{DS} - V_{DS}$ curves and (b) P_{out}, Gain and PAE for f_{req} = 10 GHz and 40 GHz with Z_{Load}=100Ω.

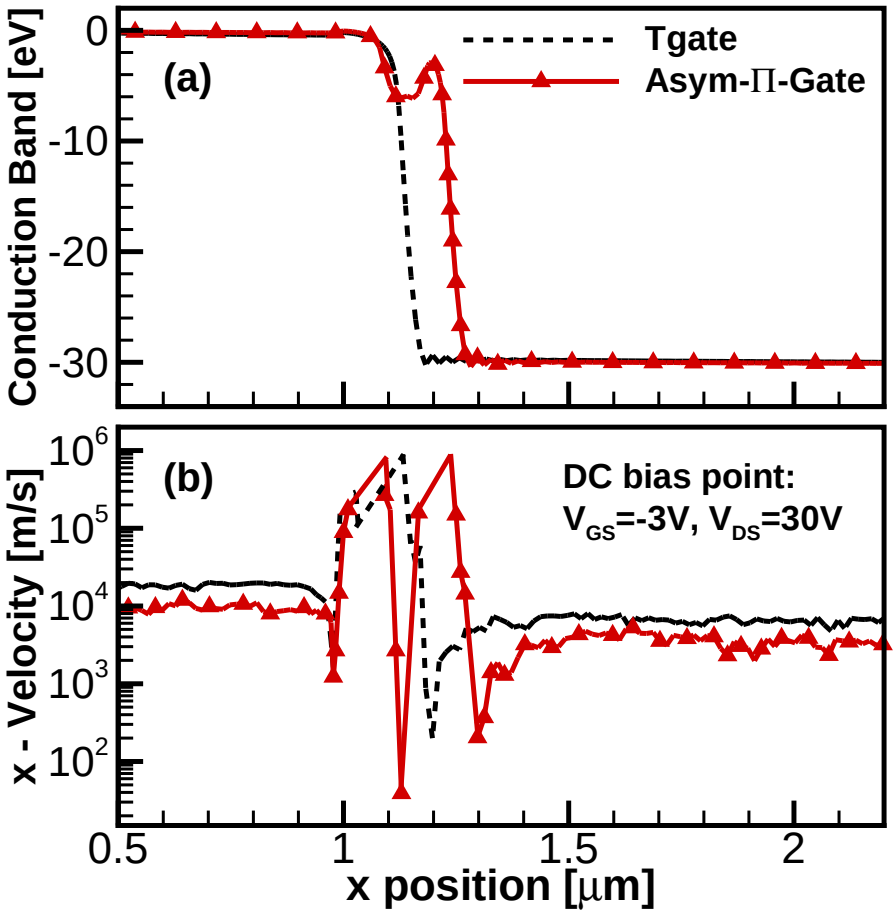

Figure 5.5: Close up of the profile along the channel (x-component) of (a) the conduction band E_C and (b) the electron velocity at $V_{GS} = -3\,\mathrm{V}$ and $V_{DS} = 30\,\mathrm{V}$ for T-gate and Asymmetric Π-gate devices.

to the bottom of the buffer, and the results are plotted on a semi-log scale where the 0 eV energy corresponds to the top of the valence band.

5.3.1 Energy Distribution in DC Regime

The EDFs in DC were built over 100 ps with 8x10^4 super-particles, for bias points (1) and (2) as indicated in figure 5.4 (a), corresponding to the peak DC power of the dynamic load lines for P_{in} = 6 dBm and 12 dBm, respectively. Figure 5.6 presents the resulting EDFs for both T-gate and proposed asymmetric-Π-gate layouts, showing that the peak generation of hot carriers, evidenced by the large tail of high energy

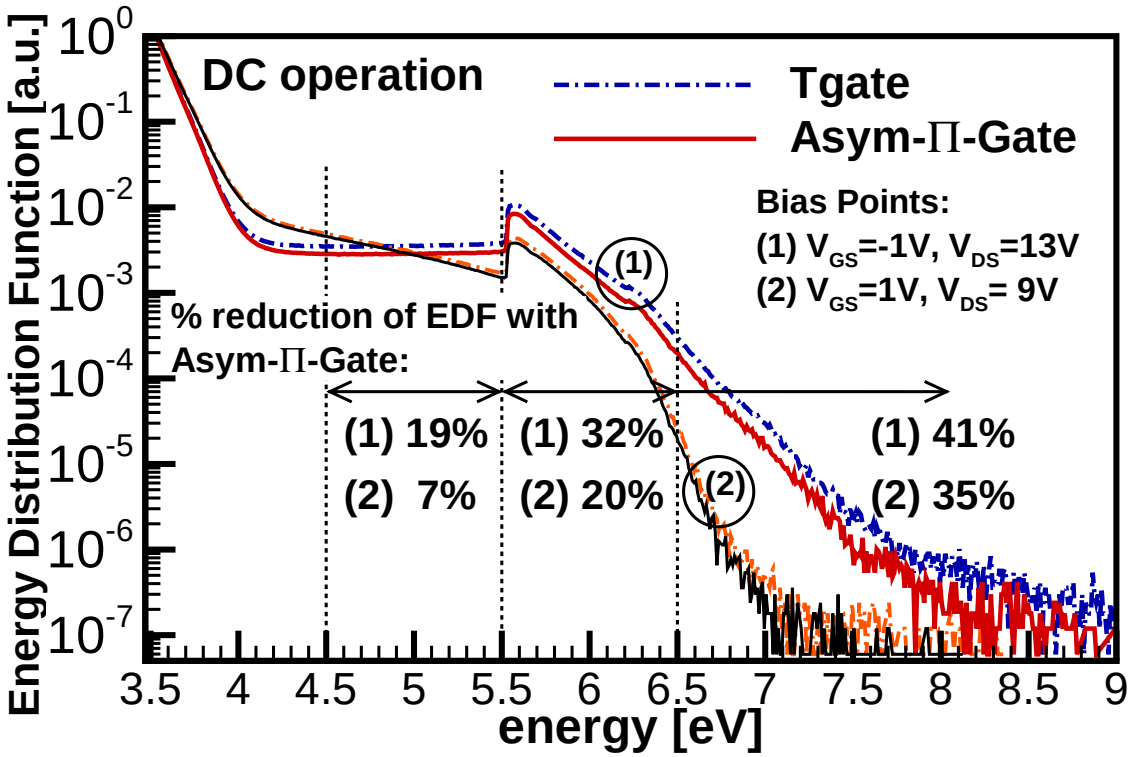

Figure 5.6: Electron energy distribution function for the DC bias points (1) and (2) shown in figure 5.4 (a) for T-gate and asymmetric Π-gate devices.

electrons including the transfer to satellite valleys at 5.5 eV, corresponds to bias point (1) where moderate current and high electric field are simultaneously present, also known as the semi-on state.

However, for the asymmetric-Π-gate, the peak EDF is 19 % lower than that of the T-gate for energies between 4 eV and 5.5 eV, approaching a 32 % reduction in the range of $5.5 < E < 6.5$ eV and reaching its maximum reduction of 41 % for $E > 6.5$ eV. For bias point (2), where V_{GS} is positive and the drain current is high, the EDF decreases for both layouts due to a lower lateral electric field. Under this condition as well, the Π-gate outperforms its T-gate counterpart in terms of the electron distribution, obtaining reductions of about 7 % for $4 < E < 5.5$ eV, 20 % for $5.5 < E < 6.5$ eV, and 35 % for $E > 6.5$ eV. These results suggest that improved reliability can be achieved with a Π-shaped gate contact since less hot-electron induced traps will be generated.

The reduction of the energy distribution observed with a Π-gate device is due to the change in the trajectories of the electrons along the channel, and it can be explained in terms of the microscopic physical quantities presented in figures 5.7 (a) to (d) showing a close-up profile under the gate along the channel. Splitting the gate contact into two stems (L_{G1}, L_{G2}) and adding a spacer region in between with length d, reshapes the electric field under the Π-gate with respect to the T-gate device as indicated by figure 5.7 (a), producing a stepped conduction band that increases the carrier concentration under the gate of the asymmetric-Π-gate as evidenced by figure 5.7 (b). In addition, when $(L_{G1} + d) > L_{G2}$ the transit time of electrons under the Π-gate increases, which enhances the total scattering under the gate as shown in figure 5.7 (c), while also reducing the length of the region depleted of carriers (L_{G2} on the drain-side of the gate) where scattering is substantially reduced due to the high peak electric field, this being the same in both devices.

A larger concentration of carriers under the Π-gate undergoing more scattering results in an overall lower velocity as shown in figure 5.7 (d), effectively reducing the number of carriers with kinetic energies high enough to transfer to satellite valleys and becoming hot as compared to the T-gate device. This effect is directly proportional to the spacer length d, which should be long enough to increase the transit time, with the trade-off of degrading small-signal AC performance, nevertheless preserving the DC and large-signal RF rating. Also, it is reported here that this effect is still observed, even though to a lesser extent, with symmetric-Π-gates where $L_{G1} = L_{G2}$ and also with lower stem heights h, which can be advantageous for fabrication.

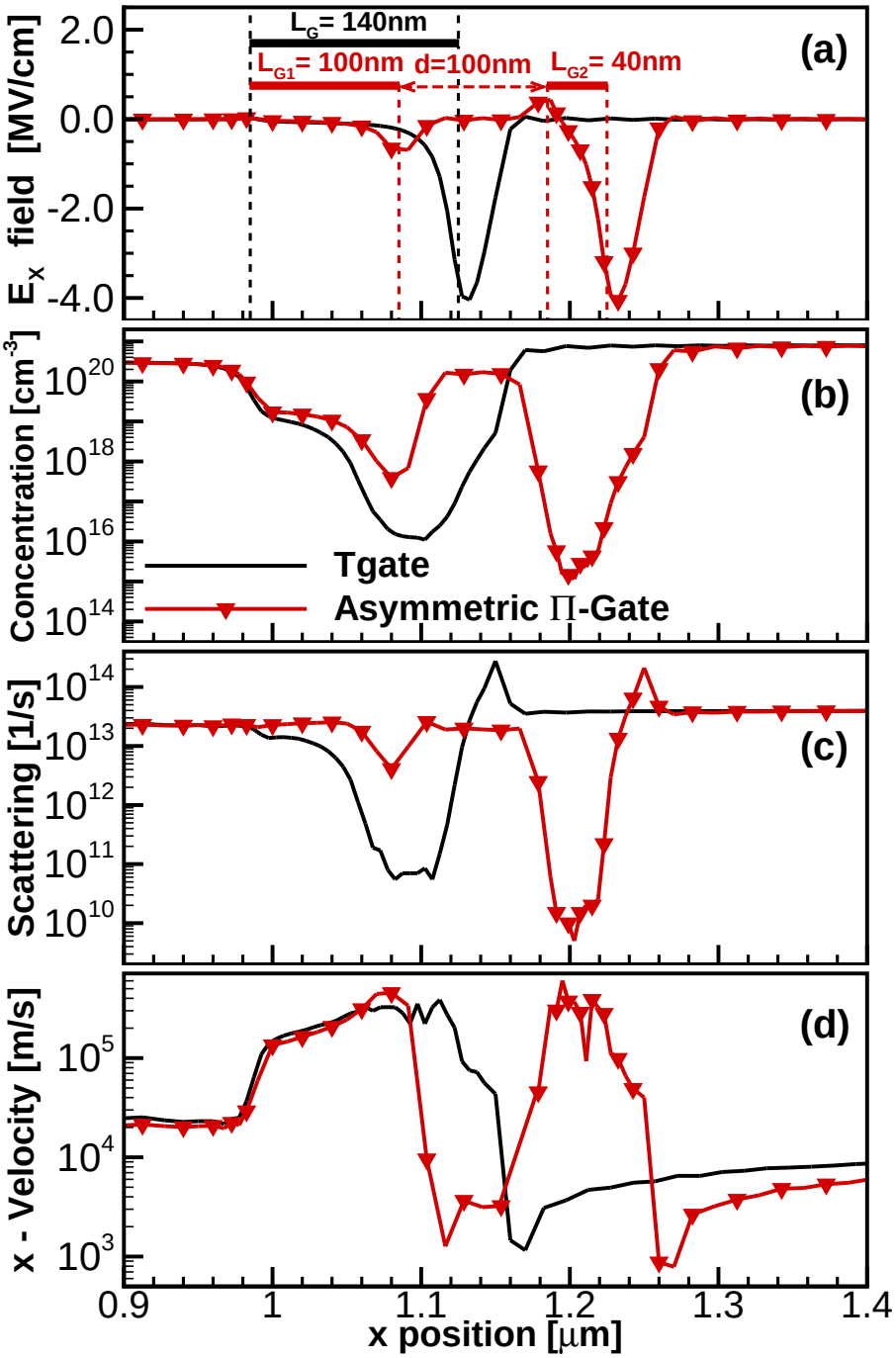

Figure 5.7: Profile along the channel of (a) x-component of the electric field, (b) electron concentration, (c) total scattering rate and (d) x-component of the electron velocity for T-gate and asymmetric-Π-gate at DC bias point (1) from figure 5.4. Solid bars in (a) show the position of the metallurgical gate contacts for each layout.

5.3.2 Energy Distribution under RF Operation

The electron EDF can also be obtained under large-signal RF conditions as previously reported [109]. Simulations were run for 200 ps of the RF signal, building the distribution function during the last 100 ps by updating the field distribution every 1.5 fs (Poisson solver time step) and normalizing to the number of iterations. Figures 5.8 (a) and (b) show the resulting EDFs at f_{req} = 10 GHz and 40 GHz, respectively, keeping the LSOP and Z_{Load} constant for both the T-gate and asymmetric-Π-gate layouts for P_{in} of 6 dBm and 18 dBm. At low frequency, figure 5.8 (a) shows that for both layouts the hot electron generation peaks at 6 dBm drive and it decreases as P_{in} increases. As compared to the T-gate at 6 dBm, the asymmetric-Π-gate shows a 32 % reduction in the EDF for $4 < E < 6.5$ eV, reaching up to 44 % reduction of hot carriers for $E > 6.5$ eV. At high P_{in} = 18 dBm, the Π-gate energy distribution is 13 % and 19 % lower than that of the T-gate for the energy ranges $4 < E < 5.5$ eV and $5.5 < E < 6.5$ eV, respectively, whereas it reaches a 46 % reduction for $E > 6.5$ eV.

At high frequency f_{req} = 40 GHz, as shown in figure 5.8 (b), the hot carrier generation increases with P_{in} contrary to what is observed at X-band [111]. For P_{in} = 6 dBm, the high frequency distribution of the Π-gate reaches that of the T-gate at 10 GHz, nonetheless, it is 25 % lower than that of the T-gate at 40 GHz for $4 < E < 6.5$ eV, reaching up to a 35 % reduction for higher energies. As the PA is driven into compression at 18 dBm drive, the EDF reaches its highest peak particularly for $E > 6$ eV, however it is in this case that the maximum reduction of the distribution is observed with the asymmetric-Π-gate being 50 % and 75 % lower than the T-gate EDF for $6.5 < E < 7.5$ eV and $E > 7.5$ eV respectively, suggesting improved robustness to failures from the generation of traps [90].

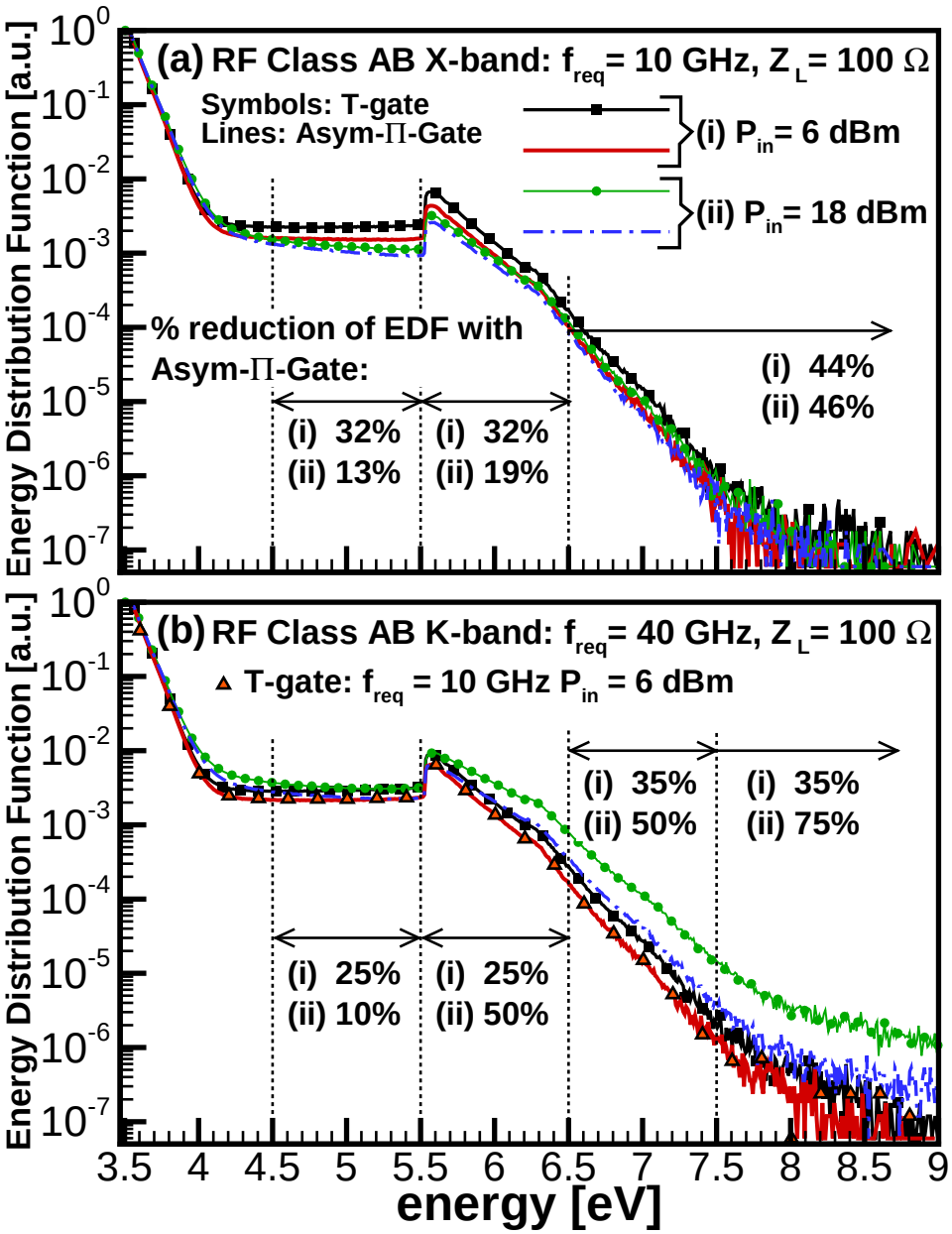

Figure 5.8: Electron energy distribution function under RF Class AB operation at (a) X-band 10 GHz and (b) K-band 40 GHz, for different input powers.

5.4 Conclusions

A new Π-shaped gate contact was proposed for improved reliability of GaN HEMTs in terms of reduced hot electron generation under all regimes of operation while preserving the device performance. The device concept was demonstrated by Monte Carlo simulations and compared to simulations of an experimental T-gate device where the simulated IV curves were first calibrated to measured data. The principles of operation were discussed and used to derive design rules, which along with projections of AC parameters f_t and f_{MAX} were presented as guidelines for design and fabrication. Furthermore, simulations of the electron EDF in both DC bias and large-signal RF drive show that under the same operating conditions, the peak of hot-carrier generation obtained with an asymmetric-Π-gate is lower up to 41 % in DC, 44 % and 75 % in Class AB at 10 GHz and at 40 GHz respectively compared to EDFs of T-gate HEMTs. These results suggest improved hot-electron induced failure conditions in both DC and large-signal RF operation of the device with an asymmetric-Π-gate, by reducing the number of highly energetic carriers that can create electrically active traps.

Chapter 6

ELECTRO-THERMAL CHARACTERIZATION OF *GAN* HEMT ON SI THROUGH SELF-CONSISTENT ENERGY BALANCE-CELLULAR MONTE CARLO DEVICE SIMULATIONS

In order to assess the mechanisms of self-heating observed in GaN HEMTs on Si substrates, the electro-thermal characterization of an experimental device is carried out in terms of the simulation of its DC performance by means of an expanded full band Cellular Monte Carlo particle-based simulator (CMC). In this new approach, the thermal effects are included through an energy-balance equation for phonons which allows to self-consistently couple the charge and heat transport. As a result, accurate temperature maps are obtained for the acoustic and optical phonon modes, showing that the location of the hot spot in the channel is not at the peak of the electric field, but instead it is shifted towards the drain up to 32 nm. Additionally, the modeled $I_{DS}(V_{GS}\text{-}V_{DS})$ space is significantly improved with respect to the traditional isothermal simulations, as a result of including the self-heating effects which modify the charge transport in the active layer of the device through the temperature dependence of the scattering mechanisms considered in the simulations. Finally, the electro-thermal model of the experimental HEMT is used to evaluate the impact that lateral scaling, i.e. reducing the source-to-gate L_{SG} and gate-to-drain L_{GD} dimensions, has on self-heating effects, showing that scaled devices offer improved DC and small-signal AC performance but are subject to higher temperatures in the channel as compared to the original non-scaled device when dissipating the same DC power.

6.1 Introduction

AlGaN/GaN HEMTs have become a prominent technology for high-power and high-frequency operation, due to their scalability and high efficiency. In particular, GaN HEMTs fabricated on Si substrates have been the object of intense study both at a research and commercial level because of the advantage of reduced cost and the capability of integration with Si MOS processes [112; 14; 113], which are both limiting factors with GaN on SiC technologies. However, the use of Si substrates raises reliability concerns in terms of crystal quality and thermal management, which degrade device performance. Although the thread dislocation density (N_{TDs}) in the GaN buffer grown on Si substrates has been reduced to non-prohibitive levels, thermal management and self-heating effects constitute reliability concerns due to the low thermal conductivity of Si, and therefor are the subjects of active research.

Although Monte Carlo techniques have been widely utilized to study electrical transport in HEMTs, the possibility to include thermal effects is less widespread and self-heating is often treated using a Relaxation Time Approximation (RTA) approach [114] or with the macroscopic Joule Heating term [115]. In this chapter, the DC electro-thermal characterization of an experimental GaN on Si HEMT is studied, by means of accurate full-band CMC simulations. The capabilities of the traditional CMC framework have been expanded to include thermal effects by means of self-consistently coupling the charge transport with a flux-based energy-balance equation for both optical and acoustic phonon modes. In this approach, instead of using the macroscopic Joule Heating term, a novel method to compute the forcing function of the heat equation is presented based on tracking the total energy exchange between electrons and phonons during simulation run-time, which does not require the conventional RTA, resulting in the full coupling between charge and heat transport.

Within this framework, the electrical performance in DC of the experimental device is accurately obtained. After the model of the experimental HEMT is validated with measured DC curves, it is used to asses the impact of lateral scaling on self-heating effects in the device.

6.2 Energy Balance Flux Approach

The underlying physical mechanism of heat generation in semiconductors, as illustrated in figure 6.1, consists of a transfer of energy from carriers, which are accelerated in the presence of an electric field, into the lattice by means of energetic optical phonons and to a lesser extent acoustic phonons, created through collisions or scattering events. Optical phonons, having very low group velocity resulting in low thermal conductivity, decay into acoustic phonons at a rate much slower than that at which these collisions occur, giving rise to hot spots. Eventually, as optical modes decay into acoustic phonons, these are able to move through the device by diffusion and be absorbed by the heat sink [114].

In order to accurately capture the physical description of heat generation within the CMC simulation framework, instead of solving the Heat Transport Equation (HTE) [116], the adopted approach is based on a flux analysis of heat generation and transport, through the solution of an energy balance equation for phonons (EBE). The time evolution of the EBE is directly obtained from the Boltzmann Transport Equation (BTE), and given for each phonon mode μ by 6.1:

$$\frac{\partial W_\mu}{\partial t} = -\nabla \cdot F_\mu\left(r,T\right) + \frac{\partial W_\mu}{\partial t}\mid_{e-p} + \frac{\partial W_\mu}{\partial t}\mid_{p-p}, \tag{6.1}$$

where $W_\mu\left(r.t\right) = 1/\Omega\,\Sigma_k\left(E_\mu(k) f_\mu(r.k.t)\right)$ is the ensemble energy density within the volume Ω of reciprocal space, $F_\mu\left(r.t\right) = 1/\Omega\,\Sigma_k\left(\nu(k) E_\mu(k) f_\mu(r.k.t)\right)$ is the energy flux

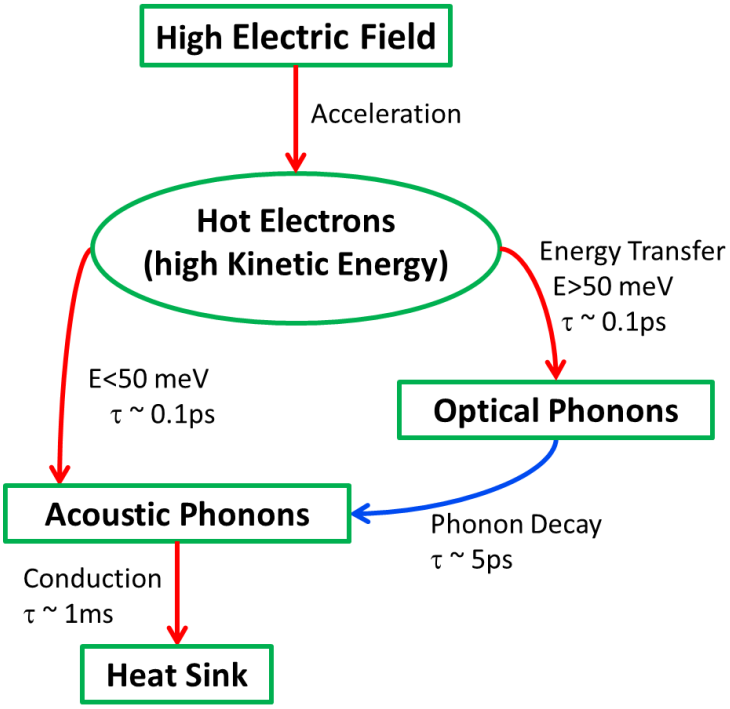

Figure 6.1: Flowchart of thermal energy transport in semiconductor devices.

or loss term, and the two partial derivatives on the right hand side (RHS) constitute the generation term due to electron-phonon (e-p) and phonon-phonon (p-p) interactions. The variables T and r correspond to the temperature and position respectively. At steady state, the LHS of 6.1 must be zero, and the EBE can be written for each phonon mode μ as:

$$\nabla \cdot (\kappa_\mu (T, r) \nabla T) = -\left(\frac{\partial W_\mu}{\partial t} \mid_{e-p} + \frac{\partial W_\mu}{\partial t} \mid_{p-p} \right) = -P_\mu (r) , \tag{6.2}$$

where the energy flux term has been approximated by Fourie's Law given by $F_\mu (r.t) = \kappa_\mu(T, r)\nabla T$, in which κ_μ is the thermal conductivity of phonon mode μ. It must be noticed that the generation term includes both sources of heat in semiconductors, namely the electron-phonon (e-p) interactions through scattering and the phonon-phonon (p-p) interactions, which account for the anharmonic decay of optical phonons into low energy acoustic phonons.

Since a robust and efficient elliptical solver for partial differential equations (PDE) is already available within the CMC framework, used to solve Poisson's equation for the electrostatic potential, equation 6.2 is further manipulated into the form of an elliptical PDE. First, it is necessary to impose the condition that the thermal conductivity is a piece-wise function of position, meaning it is constant with respect to the position within a given cell but it is allowed to change from one cell to another. Then, the term $\kappa_\mu(T, r)$ can be moved outside of the divergence operator in 6.2. Additionally, the temperature dependence of κ_μ is modeled by a simple power law of the form $\kappa_\mu(T) = \kappa_{ref}(T/T_{ref})^\alpha$, where κ_{ref} is the thermal conductivity value at a reference temperature T_{ref} and α is a fitting parameter [117], so that $\kappa_\mu(T)$ can be calibrated to experimental data of the material of interest.

The Kirchhof Transformation is then used to define a new variable Θ, which represents an "apparent" temperature and it is given by [118]:

$$\Theta_{\mu,C}(T) = T_{ref} + \frac{1}{\kappa_{\mu,C}(T_{ref})} \int_{T_{ref}}^{T} \kappa_{\mu,C}(\tau)\, \partial\tau, \tag{6.3}$$

so that μ denotes the respective phonon mode and C index of the computational cell 6.2, enabling the EBE equation 6.2 to be rewritten as a linear PDE in terms of Θ as:

$$\nabla\Theta_{\mu,c} = \frac{-P_\mu(r)}{\kappa_{\mu,c}(T_{ref})}, \tag{6.4}$$

where $P_\mu(r)$ is the heat generation term or the forcing function, and $\kappa_{\mu,c}$ is the reference value of the thermal conductivity since the temperature dependence is implicitly contained in the Θ variable. The linearity condition for 6.4 requires that both the temperature and its derivative in the normal direction, i.e. the heat flux, be continuous across a boundary [118]. This could be troublesome in the case of heterojunctions,

but it is overcome by imposing the same functional form $f(T)$ for the thermal conductivity across interfaces so that $\kappa_{\mu,c}\left(T_{ref}\right)=\lambda_c f\left(T\right)$, where λ_c is constant within a given cell, but is allowed to vary from one cell to another [118]. Even though this is a reasonable restriction for semiconductor materials, it is non-physical for metals.

It must be noticed that instead of using the macroscopic Joule Heating term, defined as $\vec{J}\cdot\vec{E}$, the forcing function $P_\mu\left(r\right)$ required in 6.4 is defined by the RHS of equation 6.2. The electron-phonon (e-p) term is computed in real-time by tracking the respective scattering mechanism, optical or acoustic, and recording the energy exchanged in the scattering event, overcoming the need of the usual assumptions such as elastic acoustic scattering events and the RTA for electron-phonon interactions as used in previous methods [114]. Nevertheless, the phonon-phonon (p-p) term is still computed using an RTA as:

$$\frac{\partial W_\mu}{\partial t}\mid_{p-p}= C_i\left(\frac{T_i-T_j}{\tau_{i-j}}\right), \tag{6.5}$$

where C_i is the heat-capacity of the phonon mode, $T_{i,j}$ is the temperature of the respective modes and τ_{i-j} is the decay time from i to j mode. In this sense, the anharmonic decay of optical phonons into acoustic phonons is described by 6.5. This novel approach, referred here as self-heating forcing function, requires the simultaneous solution of two equations as 6.4, one for the group of optical modes and one for the group of acoustic modes, each of them with a particular thermal conductivity $\kappa_{\mu,c}$. The result is two separate temperatures associated to the respective acoustic or optical phonon population.

The self-consistent electro-thermal simulation begins by first obtaining a steady-state solution of the device under study with a conventional isothermal CMC run cycle, so that the relevant electrostatic and transport quantities such as electric fields,

carrier concentration, distribution and velocity are resolved for a given potential at an initially constant temperature. The self-heating forcing function given by the RHS of 6.2 is then computed at each cell by taking averages on the net scattering energy exchanged (absorbed or emitted) between carriers and the lattice over N iterations, and the energy is normalized by the volume of the cell and the time step, resulting in a power density or $P_{\mu}(r)$. It must be highlighted that the length of the averaging window (or number of N iterations for averages) should be chosen sufficiently large to minimize the effect of noise associated to the CMC run-cycle. Once the forcing function is calculated for each cell, the heat equation 6.4 is then solved using a successive over-relaxation method to within a given tolerance, resulting in two separate temperature maps. Then the process is iterated, meaning another CMC run-cycle is carried out and the process of temperature mapping is repeated. The convergence is checked by testing if the change in temperatures between two consecutive solutions in every cell is sufficiently small. Since the local temperature obtained for each phonon mode in the grid is used to correct the corresponding acoustic and optical phonon scattering rates, then this method effectively captures the self-heating effect observed in devices.

6.3 Electro-Thermal Device Simulations

The experimental device characterized in this study was fabricated and reported by Altuntas et al. [112], and the experimental device layout is shown in figure 6.2 (a). It consists of a T-gate HEMT grown on Si 111 substrate by molecular beam epitaxy. The Si substrate is followed by a nucleation layer, typically AlN thick enough to allow the groth of a relaxed GaN buffer of 1730 nm. The epitaxial stack of the barrier consists of $Al_{0.29}Ga_{0.71}N/AlN/GaN$ with thicknesses of 10 nm and 1 nm for the $AlGaN$ and AlN layers respectively. The device is capped with a 0.5 nm thin

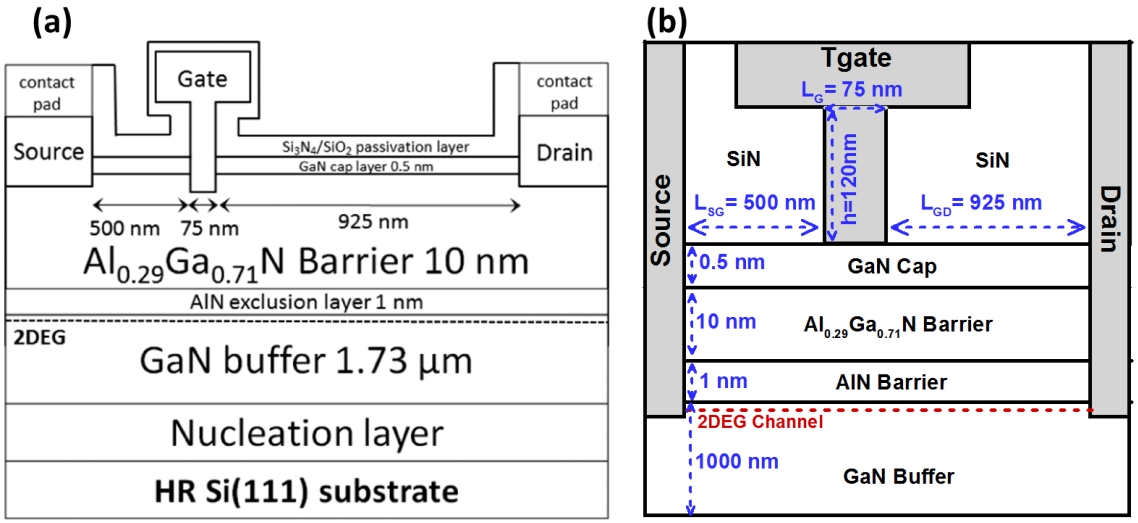

Figure 6.2: GaN-on-Si HEMT device layout showing (a) Experimental Layout by Altuntas [112] and (b) Simulated device layout.

GaN layer followed by 50 nm of SiN and 100 nm of SiO2 passivation. The gate length L_G is 75 nm, and the source-gate L_{SG} and gate-drain L_{GD} access region lengths are 500 nm and 925 nm respectively. Since the full thickness of the substrate including the nucleation layer and buffer has dimensions prohibitively large for Monte Carlo simulations, the simulated device layout shown in figure 6.2 (b) is restricted to a GaN buffer with thickness of 1000 nm.

Even though GaN-on-Si HEMTs typically have source-via grounding (SVG) to reduce the contact resistance, minimize the parasitic inductor caused by wire-bonding and improve thermal dissipation, its effects are already accounted for in the simulated device layout. The simulated device is intrinsic, meaning that the effects of packaging such as wire bonding are not included, except for the contact resistance which is calibrated with the doping of the contact layer. Concerning the improved thermal dissipation, by reducing the buffer thickness from several microns to 1000 nm, the improved efficiency of the SVG is emulated by setting the heat sink at the bottom of the buffer. In all the simulations, the carrier dynamics is captured by the

CMC framework [37], considering scattering processes due to polar optical phonons, piezoelectric (acoustic) phonons, deformation potential, ionized impurities and impact ionization. The device domain in real space is simulated by a particle-based dynamic kernel self-consistently coupled with a 2D multigrid Poisson solver.

GaN devices grown on Si substrates are subject to high thread dislocation densities (N_{TDs}) whose effect was included through threading dislocation scattering, calculated for the reported value of T_{DD}=5x10^{9}cm^{-2}, and by adding vertical sheets of charge in the buffer. The polarization charge is set as charge sheet layers at the heterointerfaces, calculated with Ambacher's formalism [45; 84]. The reference thermal conductivity at 300 K for GaN was set to $\kappa_{ref} = 130$ W/K·m with α = -0.43, while the bulk volumetric heat capacity was chosen to be 2.65x10^6 W/m^3 and the optical phonon frequency 98.6 THz. Regarding the optical phonon decay relaxation time, it was calibrated between 2 ps and 5 ps throughout the simulation range [119]. Whereas most electro-thermal studies focus on the GaN/Si or GaN/SiC substrate interface, here the focus is on the active region of the device. All simulations were performed with an initial lattice temperature of 300 K and the electro-thermal simulations included a thermal heat sink set to 300 K located at the bottom of the buffer.

Figures 6.3 and 6.4 show the maps for the acoustic mode temperature T_A and optical mode temperature T_{LO} respectively, plotted as a contour on the 3D conduction band simulated at a DC bias condition of V_{GS}=2.0V, V_{DS}=10V, where high current and high electric field are simultaneously present but at a moderate DC power. From the figure it can be seen that the peak temperature for both cases is located on the drain side of gate, particularly in the 2DEG channel where most of the phonons are generated through electron-phonon scattering. However, for the acoustic phonons the peak temperature in the channel reaches 370 K and it diffuses uniformly all the way through the device. On the other hand, the optical mode temperature reaches a high

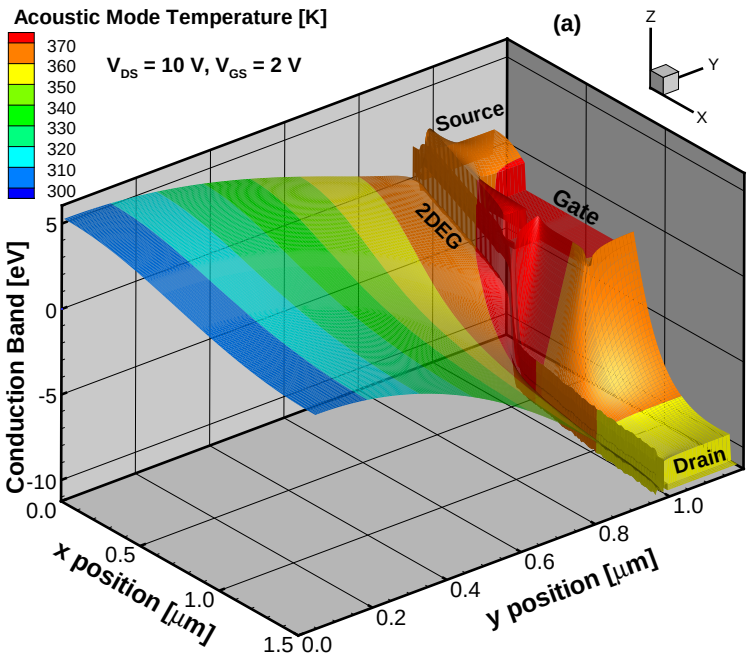

Figure 6.3: 3D conduction band of the HEMT with temperature profile for Acoustic phonon modes at DC bias of $V_{GS} = 2\,\mathrm{V}$, $V_{DS} = 10\,\mathrm{V}$.

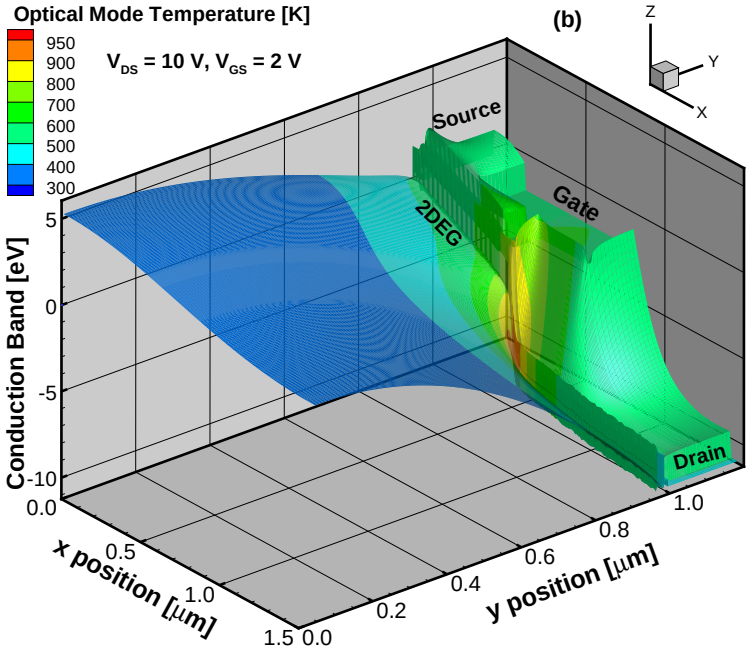

Figure 6.4: 3D conduction band of the HEMT with temperature profile for Optical phonon modes at DC bias of $V_{GS} = 2\,\mathrm{V}$, $V_{DS} = 10\,\mathrm{V}$.

peak of 950 K in the channel and the temperature rapidly drops several hundred K away from the peak. These results confirm that most of the energy is transferred from hot carries into the lattice through optical phonons, which are highly localized in the drain side of the gate creating a hot spot.

6.4 DC Electro-Thermal Characterization

The electro-thermal model of the device under study was validated to DC experimental data of the I_{DS}-V_{DS} and I_{DS}-V_{GS} curves reported by Altuntas [112] as the DC measurements (in contrast to the pulsed curves), and the results are presented in figures 6.5 (a) and (b) respectively. Isothermal simulations at 300K, i.e. regular CMC simulations with a constant lattice temperature, are presented in blue segmented lines showing a good agreement for both the output and transfer characteristics at low voltage, however it can be seen that for currents higher than 0.8 A/mm the value of I_{DS} is significantly overestimated because the self-heating of the device is not considered. This causes a higher drain current in the output and transfer characteristics, as well as a higher transconductance G_m which in turn leads to the overestimation of small and large signal parameters such as cut-off frequency f_t and output power, not presented here. When the self-consistently coupled energy-balance/CMC simulation is performed using the self-heating forcing function, I_{DS} is correctly predicted throughout the $I_{DS}(V_{GS},V_{DS})$ space, shown in red solid lines.

Once the optical and acoustic phonon mode temperatures are self-consistently computed, its values are used to determine the appropriate scattering rates, previously tabulated within a wide temperature range. As the temperature increases, both acoustic and optical scattering rates increase leading to higher total scattering, in turn reducing the electron velocity along the channel and therefore the total current. Figures 6.6 (a) and (b) present a profile along the channel of the x-velocity of electrons

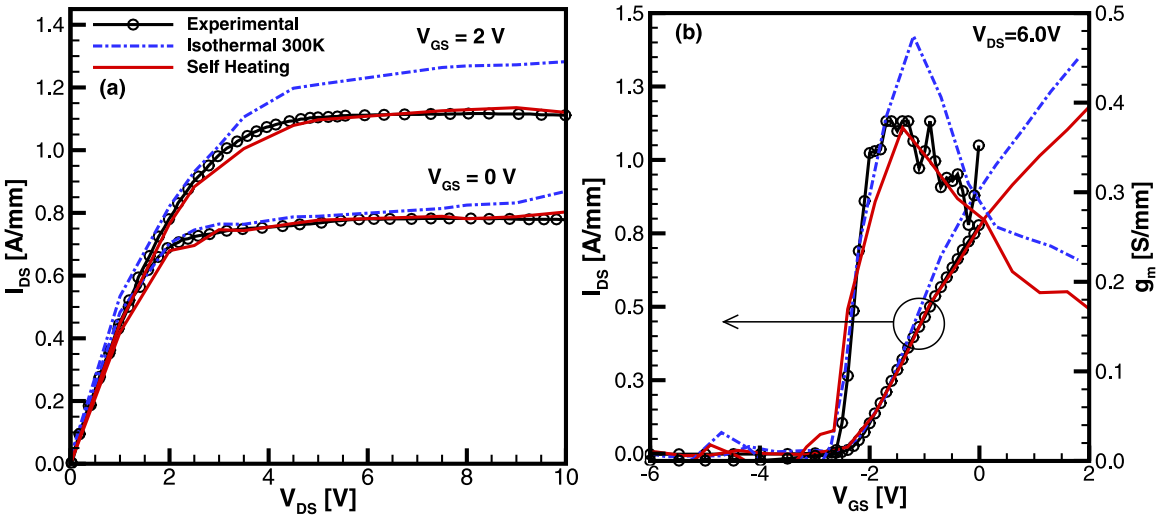

Figure 6.5: Validation of the electro-thermal model with experimental curves corresponding to (a) I_D-V_{DS} for V_{GS} = 0 V and 2 V, and (b) I_D-V_{GS} for V_{DS}=6 V. Isothermal simulations with T = 300 K are also shown.

and the total scattering in that order, for both the isothermal and self-heating cases at V_{GS}=2.0V, V_{DS}=5.0V, which corresponds to the the saturation bias point I_{Dsat} shown in figure 6.5 (a). It must be noticed that even though the overall scattering rate is higher with self-heating and the overall velocity is lower, the peak velocity remains unchanged with respect to the isothermal simulations. This is because self-heating does not significantly modify the electric field distribution, roughly given by the ratio V_{DG}/L_{GD} where $V_{DG} = V_{DS} - V_{GS}$.

A detailed profile of the x-electric field (E_x), acoustic mode temperature T_A and optical mode temperature T_{LO} along the channel is presented in figures 6.7 (a), (b) and (c) respectively, for V_{GS} = 2 V at three different values of V_{DS} which correspond to 3 points along the I_{DS}-V_{DS} curve shown in figure 6.5 (a). At low V_{DS} = 1 V, where the current is 0.4 A/mm and E_x along the channel is small and nearly constant, T_A increased up to 330 K above the initial 300 K while T_{LO} increased up 400K in the L_{GD} region. The temperature profile for this case is very uniform along the channel, and since the variation in temperature with respect to the isothermal case is relatively

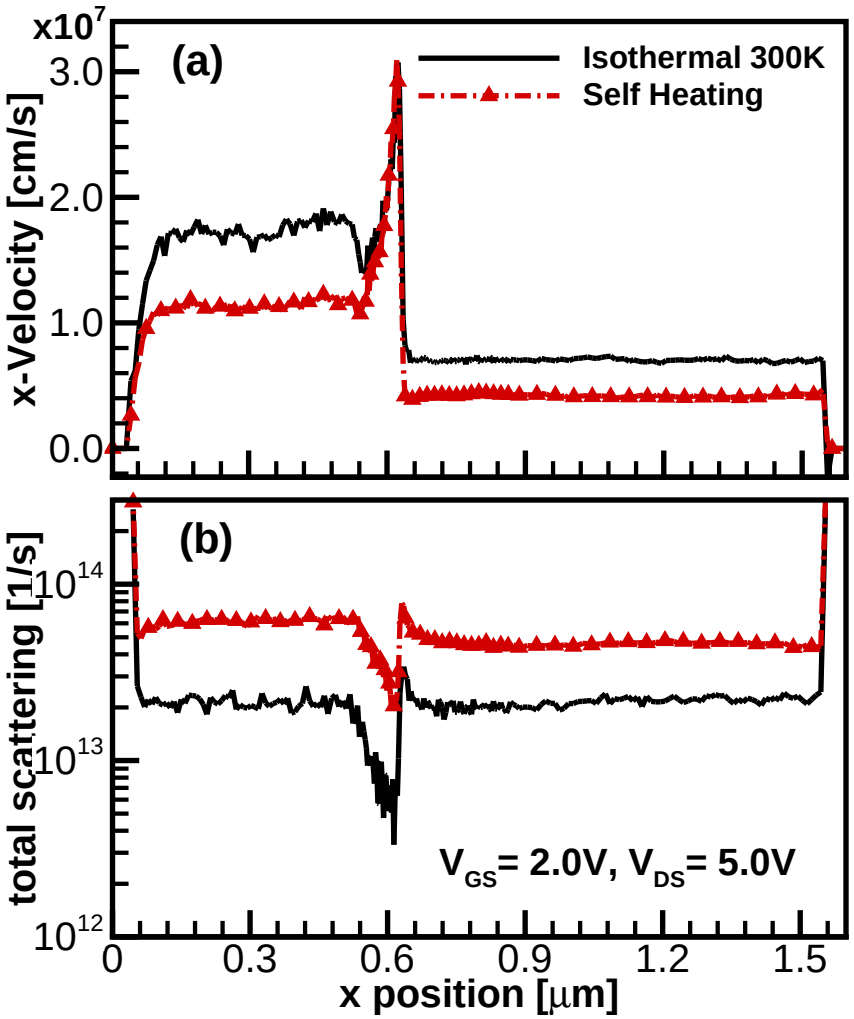

Figure 6.6: Profile along the channel of (a) electron x-velocity and (b) total scattering at DC bias point $V_{GS} = 2\,\mathrm{V}$, $V_{DS} = 5\,\mathrm{V}$, comparing isothermal simulations at $T = 300\,\mathrm{K}$ and self-heating curves.

small, the change in the scattering rate and therefore drain current is also small, reflected by the negligible difference among the plots in figure 6.5(a), which indicates that self-heating is not a dominant factor in this regime.

However, as the drain voltage increases so does the electric field and both acoustic and phonon mode temperatures. Moreover, at V_{DS}=10V a high peak E_x is present on the drain-side of the gate along with a peak T_A of 375 K and a high peak of 960 K for T_{LO}. A relevant result that is obtained due to the novel forcing function approach

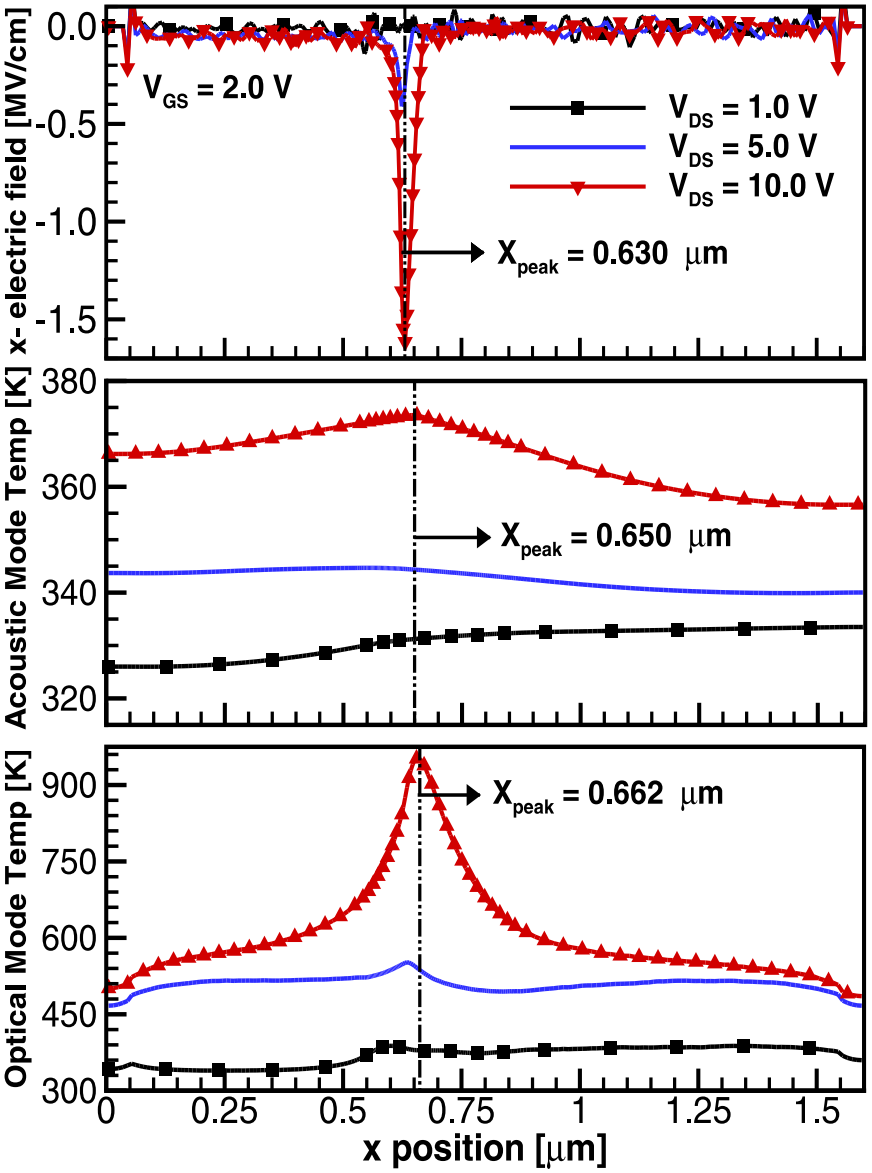

Figure 6.7: Profile along the channel of (a) x- component of the electric field, (b) acoustic mode temperature and (c) optical mode temperature for V_{GS}=2V at three different V_{DS}.

implemented in this work is that the position of the peak electric field in the channel does not correspond to the position of the peak temperature of either phonon mode, which is typically the case with other methods such as the Joule heating. This is due to the fact that the relaxation time of electrons is non-zero, i.e. once electrons are accelerated by the field gaining kinetic energy, they traverse a short region in the channel before they scatter and emit phonons. From the figure it can be seen that at $V_{DS} = 10\,\mathrm{V}$ there is a displacement between the peak E_x and peak temperature of 20 nm and 32 nm for the acoustic and optical modes respectively. This result provides relevant information to develop thermal management strategies to mitigate self-heating effects.

6.5 Impact of Lateral Scaling on Self-Heating

One of the main applications for GaN HEMTs is in the field of power amplifiers (PAs) operating at ultra high frequencies. In particular, the high electron velocities observed in the 2DEG of $AlGaN/GaN$ heterostructures make this a strong candidate for the millimeter-wave (mm-wave) range corresponding to $f_{req} > 30\,\mathrm{GHz}$. For this reason, several strategies have been developed to improve the frequency response of HEMTs. Among them, scaling the gate length L_G has proven to be very effective to increase the cut-off frequency f_t with the trade-off of increased short-channel effects, due to a lower aspect ratio of L_G/B_{Th} where B_{Th} is the thickness of the barrier. If the L_G scaling is accompanied by vertical scaling, i.e. B_{Th} is also reduced to preserve the aspect ratio, then effective suppression of short-channel effects is achieved but the gate leakage current is increased and the output current is degraded due to a lower conductivity of the 2DEG.

With the aim of overcoming the limitations of traditional vertical scaling, a significant amount of studies have focused on the effect of lateral scaling, i.e. reducing

the source-to-gate L_{SG} and gate-to-drain L_{GD} access region lengths [120; 121].The effectiveness of this strategy is well substantiated by the definition of the main metrics used to determine the frequency response of FETs, namely f_t which represents the maximum oscillation frequency of the device, and f_{MAX} which is the maximum frequency at which the transistor provides a power gain. These parameters are given by equations 6.6 and 6.7 for f_t and f_{MAX} respectively [122; 7]:

$$f_t = \frac{G_m/(2\pi)}{(C_{gs}+C_{gd}) + C_{gd}\cdot G_m\cdot(R_s+R_d) + g_d\cdot(R_s+R_d)\cdot(C_{gs}+C_{gd})}, \tag{6.6}$$

$$f_{MAX} = \frac{f_t}{2\sqrt{g_d\cdot(R_s+R_d+R_i) + R_g C_{gd}\,(2\pi f_t)}}, \tag{6.7}$$

where C_{gs}, C_{gd}, G_m, g_d, R_s and R_d represent gate-to-source and gate-to-drain capacitances, transconductance, output conductance and source and drain resistances, in that order. It must be noticed that R_s and R_d are parasitic resistances associated to the access regions of the device, and if they are negligible then equation 6.6 reduces to the typical expression $f_t = G_m/(2\pi\,(C_{gs}+C_{gd}))$ dominated by G_m and the capacitances, in which case f_t is clearly improved by increasing G_m and reducing the transit time under the gate which can be done by increasing the velocity of carriers.

The condition of low parasitic resistances can be achieved by either increasing the conductivity of the 2DEG or by simply reducing L_{SG} and L_{GD} of the transistor which is exactly the objective of lateral scaling. Even though the improvement in frequency response of HEMTs due to lateral scaling has been studied both with experiments and transport simulations [123; 124], the impact it has on reliability and particularly on self-heating effects is still a matter of interest. This issue is addressed in the following sections using the electro-thermal model of the experimental GaN-on-Si

HEMT of the previous sections, where the effect of L_{GD} and L_{SG} scaling has been treated separately.

6.5.1 Gate-to-Drain L_{GD} Scaling

The parasitic resistance R_d is dominated by the gate-to-drain region length. The effect of reducing the length L_{GD} from 925 nm of the experimental device down to 300 nm while keeping $L_{SG} = 500$ nm constant, can be seen in the output and transfer characteristics shown in figures 6.8 (a) and (b) respectively. A quick glance of panel (b) shows that L_{GD} has no effect on the $I_{DS} - V_{GS}$ curves and therefore on G_m, suggesting low impact on f_t. However, the output characteristic shows that as L_{GD} decreases so does the saturation or knee voltage V_{Kn}, going from approximately 5 V at 925 nm to 3 V for 300 nm. This in turn significantly reduces the on-resistance R_{on}, evidenced by the steeper slope of the linear region of the device as L_{GD} is reduced. On the other hand, the saturation or maximum current remains constant. Reducing V_{Kn} and R_{on} signify an improvement for PA design because it allows increasing the excursion of the load line and therefore the output power.

With the aim of assessing the impact on the reliability of the device, the bias point given by $V_{DS} = 5$ V and $V_{GS} = 2$ V on the $I_{DS} - V_{DS}$ curve was chosen because it corresponds to a point with constant DC power P_{DC} among devices with scaled L_{GD}, providing a good baseline to study the self-heating effects. First, the impact on the electric field and velocity distribution along the channel is presented in figures 6.9 (a) and (b) respectively. As expected, the peak electric field in the channel rises as L_{GD} decreases, because the same potential ($V_{DG} = V_{DS} - V_{GS}$) is dropped over a shorter length, while the potential lost in the parasitic component R_d is also reduced. This in turn increases the peak velocity of electrons in the channel allowing for a higher f_t. However, since C_{gd} is inversely proportional to L_{GD} then f_{MAX} performance is

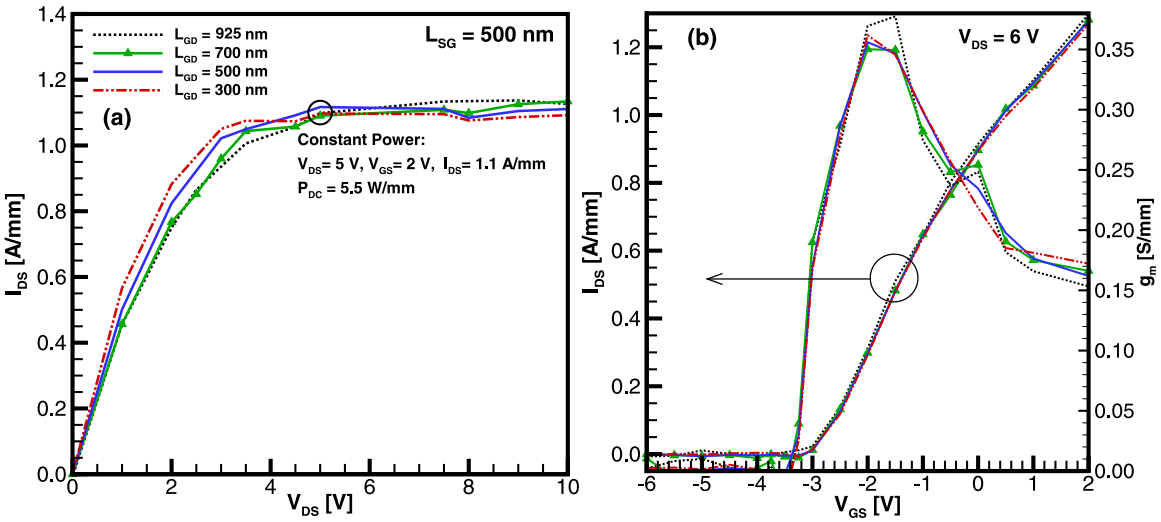

Figure 6.8: Effect of gate-to-drain L_{GD} lateral scaling on the (a) I_{DS}-V_{DS} for V_{GS} = 2 V and (b) I_{DS}-V_{GS} for V_{DS}=6 V, also shown the transconductance G_m in right-axis of (b). Self-heating effects included in all simulations.

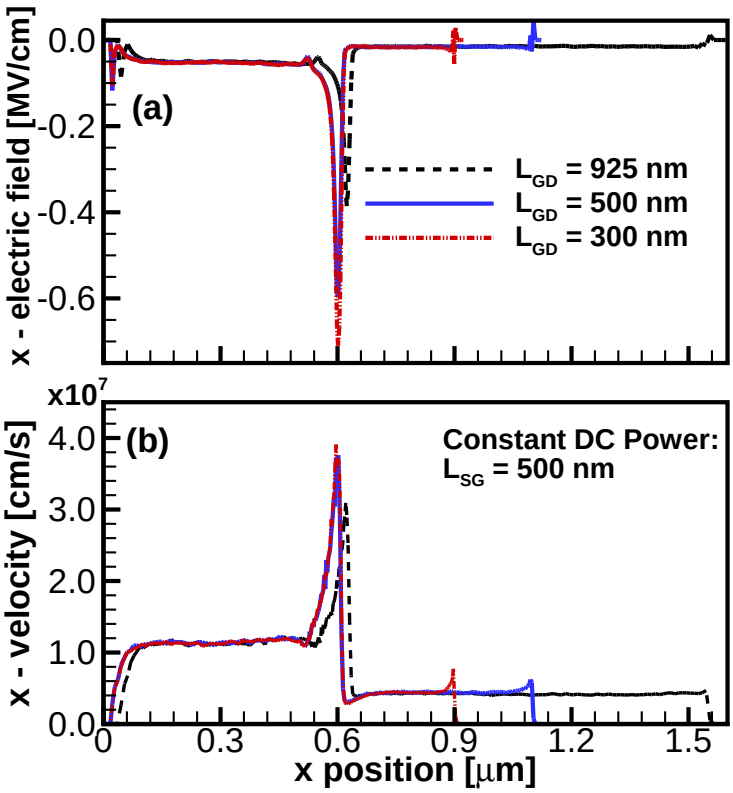

Figure 6.9: Profile along the channel of (a) electric field (b) electron velocity for three different values of L_{GD}. The DC bias point V_{GS} = 2 V, V_{DS} = 5 V corresponds to the point with constant DC power from figure 6.8.

degraded with this scaling. Even though these results indicate better f_t performance, increasing the peak field also implies a reduction of the breakdown voltage BV limiting the load-line excursion, which is a competing effect with the improvement achieved with the lower V_{Kn} also observed.

In terms of self-heating, the temperature profiles of the acoustic and optical phonon modes along the channel are shown in the left and right axis of figure 6.10 respectively, corresponding to the point with constant P_{DC}. Due to the higher electric field in the channel, as L_{GD} is reduced from 925 nm to 300 nm the peak T_{LO} goes up from 550 K to 630 K which represents a 15 % increment. Since optical phonon modes undergo anharmonic decay into acoustic modes, the peak T_A also increases going from 345 K to 356 K which represents a 3 % increment. Even though it seems a modest increment, it should be kept in mind that T_A is close to the lattice temperature effectively measured in a real device, and the 3 % increment in Kelvin represents a rise of temperature from 72°C to 83°C in the channel. Also, it must be noticed that asymmetries in the device, i.e. L_{GD} different to L_{SG}, induce a non-uniform distribution of T_A along the channel having as a general trend that the shorter region will hold the highest temperature, whereas a uniform T_A is observed for $L_{GD} = L_{SG}$ =500 nm.

From this analysis it can be concluded that scaling of L_{GD} yields lower V_{Kn} and R_{on} which can be used to improve the load-line excursion in PA applications, while practically no effect is seen in G_m. Nevertheless, the frequency response of the device given by f_t still increases because the peak electric field in the channel is higher as L_{GD} decreases, inducing a higher peak velocity which in turn reduces the transit time of carriers. However, this effect is detrimental in terms of BV and represents a competing effect with the lower V_{Kn} for maximum load-line excursion. In terms of self-heating, an increment of 15 % and 3 % is observed for T_{LO} and T_A respectively when L_{GD} is reduced by a factor of 3 under constant power dissipation. This represents a

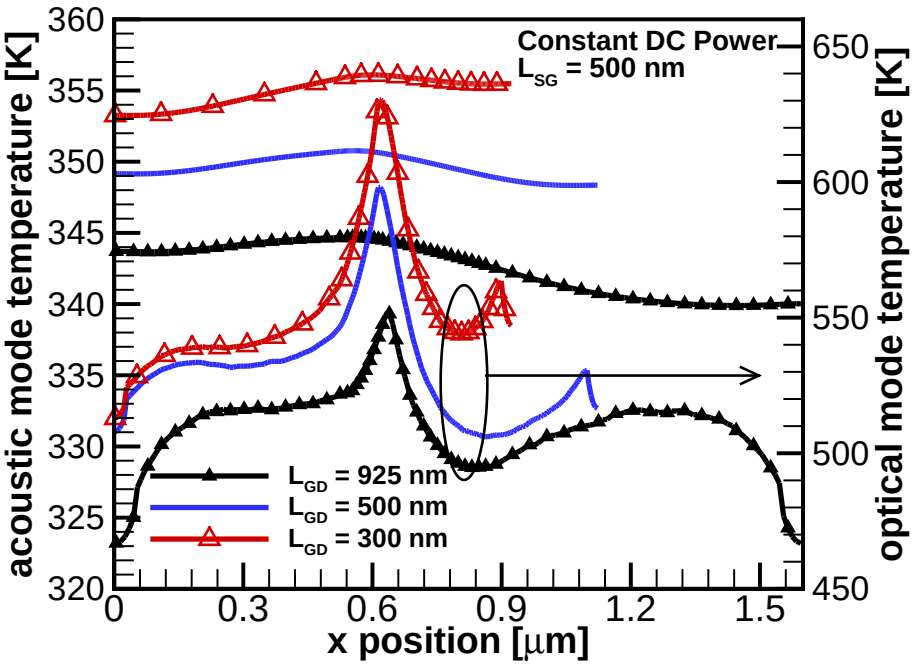

Figure 6.10: Effect of lateral scaling of L_{GD} on the profile along the channel of acoustic mode temperature (left-axis) and optical mode temperature (right-axis) for DC bias point from figure 6.8 corresponding to constant DC power.

rise in temperature from 72°C to 83°C in the channel, suggesting that lateral scaling of the gate-to-drain region requires implementing thermal management strategies to take advantage of the improved electrical performance while preserving reliability.

6.5.2 Source-to-Gate L_{SG} Scaling

The other lateral dimension that can be scaled is the source-to-gate access region length L_{SG} which is correlated to the parasitic R_s. Its effect can be seen in the output and transfer characteristics of the HEMT shown in figures 6.11 (a) and (b) respectively, where L_{SG} was reduced from the experimental value of 500 nm down to 200 nm while keeping L_{GD} = 925 nm constant. Unlike what happens with L_{GD} scaling, the $I_{DS} - V_{GS}$ curves show that as L_{SG} is reduced I_{DS} increases for the same applied voltage V_{GS}, which translates into a higher G_m while the threshold voltage remains constant. The improvement in G_m is due to the fact that a higher intrinsic

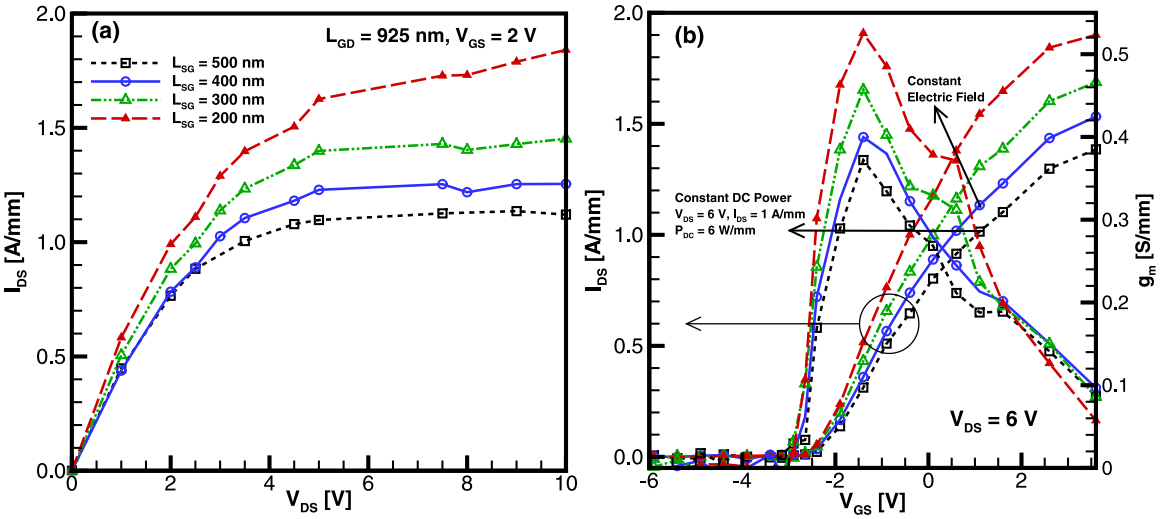

Figure 6.11: Effect of source-to-gate L_{SG} lateral scaling, while keeping L_{GD} constant, on the (a) I_{DS}-V_{DS} for V_{GS} =2 V and (b) I_{DS}-V_{GS} for V_{DS}=6 V, also shown in right-axis of (b) the transconductance G_m. Self-heating effects included in all simulations.

V_{GS} is applied to the device as R_s is reduced, since less potential is lost in the access region, allowing for a stronger modulation of charge in the channel. Furthermore, these results imply that higher f_t can be achieved due to the higher G_m and lower R_s obtained with this strategy. In addition, significant variations are also observed in the $I_{DS} - V_{DS}$ space, where the current I_{Dsat} increases and R_{on} decreases as L_{SG} is reduced, with the trade-off of higher V_{Kn}, which can be deleterious since a higher potential is required to drive the device into saturation.

Concerning the impact on reliability, two relevant conditions should be considered when studying the effect of reducing L_{SG} as indicated in figure 6.11 (b). The first corresponds to comparing self-heating along DC points with constant P_{DC}, which in this case corresponds to V_{DS} = 6 V while the value of V_{GS} was chosen to keep I_{DS} =1 A/mm as L_{SG} is decreased, providing a constant P_{DC} =6 W/mm. It must be noticed that as L_{SG} decreases from 500 nm to 200 nm, the V_{GS} required to keep constant current decreases from 1 V to -0.5 V. The reason is that a higher V_{DS} is dropped across a shorter length as L_{SG} decreases, increasing the peak electric field

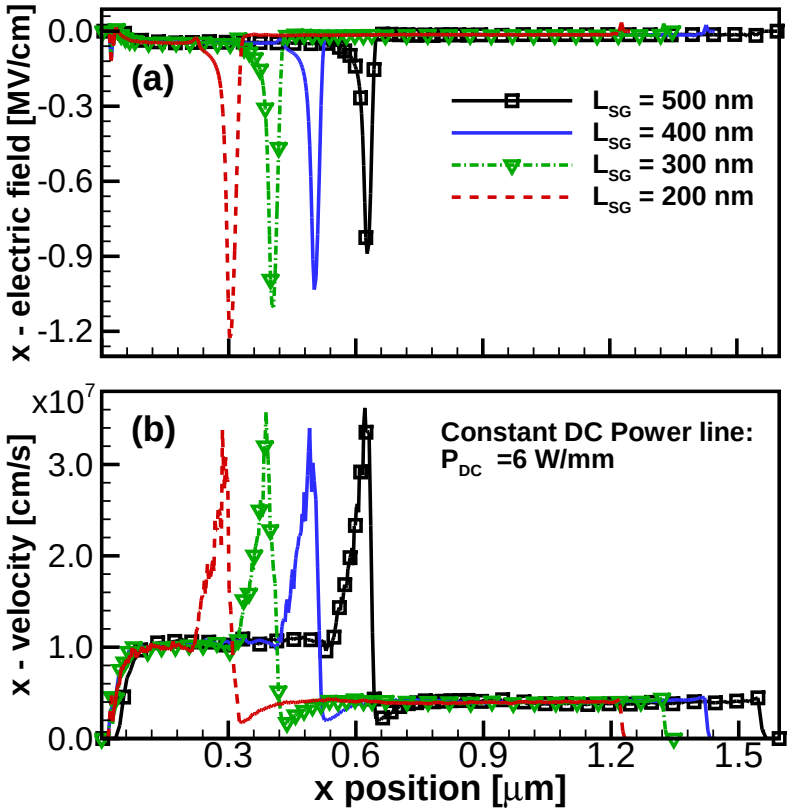

Figure 6.12: Profile along the channel of the electric field for two different L_{SG} values. DC bias points along the line of constant DC power from figure 6.11 (b), implying different values of V_{GS} were used, while keeping V_{DS} and I_{DS} constant.

in the channel as indicated in figure 6.12 (a). This in turn allows for a high electron velocity as shown in figure 6.12 (b), resulting in constant I_{DS}.

Since the peak electric field increases as L_{SG} is scaled down, then more electrons will gain high kinetic energy that will be transferred to the lattice through scattering, resulting in higher temperatures for both acoustic and optical temperature modes for scaled devices dissipating the same power. This is demonstrated in figure 6.13 where a profile along the channel of T_A and T_{LO} is shown for devices with L_{SG} = 200 nm and 500 nm under constant P_{DC}. In this case, the peak T_{LO} increases from 640 K to 740 K when L_{SG} goes from 500 nm to 200 nm, representing a 15 % increment, while the peak T_A rises from 348 K to 360 K under the same scaling, representing an increment of

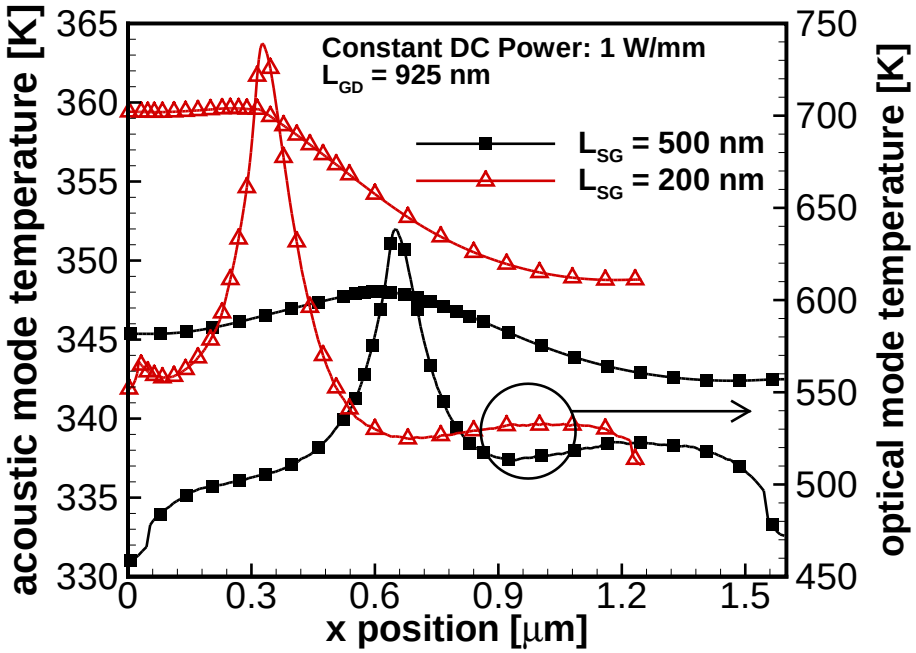

Figure 6.13: Effect of lateral scaling of L_{SG} on the profile along the channel of acoustic mode temperature (left-axis) and optical mode temperature (right-axis) for DC bias points from figure 6.11 (b) corresponding to the line with constant DC power.

temperature from 75°C to 87°C in the channel. It must be highlighted that in the scaled device the profile of T_A is non-uniform showcasing a temperature gradient from 360 K in the source-to-gate access region to 350 K on the gate-to-drain region.

The second condition of interest to assess the impact of L_{SG} scaling on self-heating corresponds to keeping the electric field constant instead of the dissipated power, as it is shown in figure 6.11 (b) with the constant electric field line. In this condition the DC bias point allows for higher I_{DS} as L_{SG} decreases. Figure 6.14 shows the profile along the channel of the electric field (left-axis) and the carrier velocity (right-axis) for L_{SG} = 200 nm and 400 nm. Even though the peak electric field is constant for both devices, the peak velocity is higher for the scaled transistor producing higher I_{DS}, which translates into a higher DC power dissipation. Therefore, since a higher DC power is dissipated in a smaller area, the temperature is expected to increase as L_{SG} is reduced.

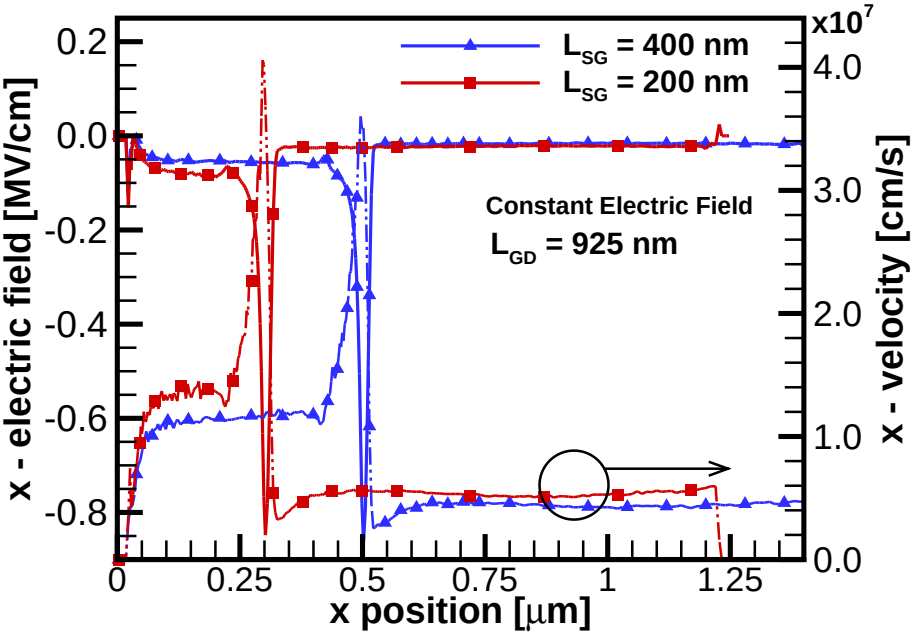

Figure 6.14: Profile along the channel of (a) electric field (b) electron velocity for two different L_{SG} = 200 nm and 400 nm where the DC bias points were chosen along the constant electric field line from figure 6.11 (b).

Figure 6.15 presents the temperature profile along the channel for the acoustic modes (left-axis) and optical modes (right-axis) under constant electric field conditions for 2 devices with L_{SG} = 200 nm and 400 nm. As expected, the peak T_A rises from 355 K to 372 K in the scaled device resulting in a peak channel temperature of 99°C. Once again T_A exhibits a non-uniform profile along the x direction due to the asymmetric lenghts of the access regions. Regarding the optical modes, the peak T_{LO} increases from 670 K to 790 K as L_{SG} goes from 400 nm to 200 nm.

From both conditions studied to asses the impact of L_{SG} scaling on self-heating it can be concluded that a scaled devices experience a significant increment in temperature as a trade-off to improved electrical performance, which can be detrimental to short-term and long-term reliability. In this sense, in order to fully take advantage of lateral scaling strategies, it is necessary to develop efficient thermal management techniques to go with it.

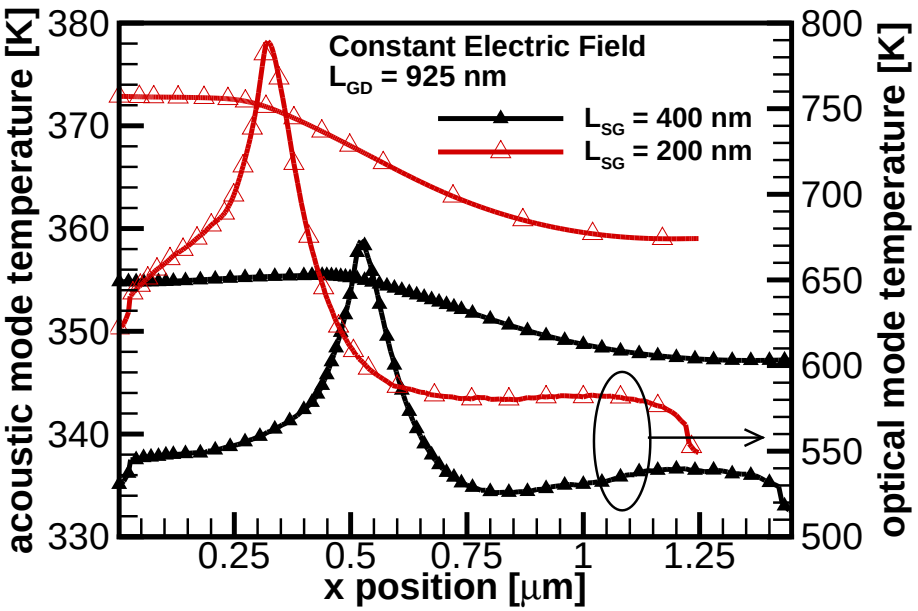

Figure 6.15: Profile along the channel of acoustic mode temperature (left-axis) and optical mode temperature (right-axis) under lateral scaling of L_{SG}, along the constant electric field line from figure 6.11 (b).

6.6 Conclusions

The electro-thermal characterization of a GaN HEMT on Si was presented. Simulations of the DC characteristics were obtained with an expanded Cellular Monte Carlo framework, where thermal effects are included through an energy-balance equation for phonons. The forcing function is computed by tracking the energy exchange during scattering events, which self-consistently couples the charge and heat transport. The electro-thermal model was validated to DC experimental data showing excellent agreement throughout the entire $I_{DS}(V_{GS} - V_{DS})$ space. Unlike what is typically reported, through the accurate thermal mapping of the device it was found that the position of the hot spot in the channel is not located at the peak electric field, but is actually shifted towards the drain by up to 32nm. Additionally, the electro-thermal model of the device was used to study the impact lateral scaling has on self-heating effects, showing that a significant improvement is observed in terms of electrical per-

formance but it is accompanied by the significant increase of the temperature in the channel when devices operate under constant DC power or constant electric field, highlighting the need of efficient thermal management strategies in GaN-HEMTs.

Chapter 7

SUMMARY

In this work, a simulation study of reliability in millimeter-wave (mm-wave) GaN Devices for power amplifier (PAs) applications was presented. The study was performed by means of a full band Cellular Monte Carlo particle-based device simulator (CMC), and it consisted in the systematic characterization of the performance of GaN devices operating under DC, small signal AC and large-signal radio-frequency (RF) conditions from which the microscopic properties of the device were correlated to effects such as generation of hot carriers and self-heating typically associated to degradation of performance.

The conceptual framework relevant for this work was introduced in the first two chapters. In particular, a short review of semiconductor device modeling was provided in the introduction, making emphasis on the CMC simulator. This was followed in chapter 2 by a brief review of concepts concerning GaN technology, focusing on the polarization effects in wurtzite GaN, principles of operation of devices such as HEMTs and HETs and reliability mechanisms relevant to this technology. Also, the principles of operation and design of PAs were discussed.

The first contribution of this work was presented in chapter 3, focusing on the study of non-idealities that limit the DC electrical performance of isotype $AlGaN/GaN$ heterojunction diodes. CMC simulations were performed to reproduce the measured IV curves of an experimental device, however significant discrepancies were found when the layout of the ideal device was used. This result was confirmed by means of hydrodynamic simulations performed with commercial software Sentaurus by Synopsys, suggesting that non-ideal effects should be included in the model to account for

the degraded electrical performance. Then by means of hydrodynamic simulations it was demonstrated that variations in the mole fraction of the $AlGaN/GaN$ barrier along with the presence of non-ideal Schottky contacts were responsible for the reduction of current drive capability of the experimental device. In the context of the development of HETs technology, these results provide valuable information about deleterious effects that degrade HETs' performance.

In chapter 4, a study of hot electron generation in GaN HEMTs was carried out in terms of the electron energy distribution function (EDF), calculated under large-signal RF, Class AB PA operation. The results suggest that in the X-band at 10 GHz, RF hot-electron degradation effects should be lower than under DC stress, regardless of the input power or temperature of operation which is consistent with experimental studies. However, in the mm-wave band up to 40 GHz, hot carrier generation matches the DC counterpart, even surpassing the DC concentration of high energy carriers when the PA is driven into deep compression, suggesting stronger degradation under RF than in DC. The study was extended to compare hot carrier generation under Class A and Class AB topologies, resulting in EDFs that suggest devices are more susceptible to hot-carrier degradation when operating in Class A. The relevance of this study lies in the fact that in terms of reliability, the RF analysis is more meaningful for realistic lifetime estimation given the practical application of GaN HEMTs as mm-wave PAs, and in that the accurate calculation of the EDF under RF can be exploited in compact modeling-based lifetime and reliability analysis as it has been done for DC operation.

Using the same methodology of chapter 4, in chapter 5 a new Π-shaped gate contact was proposed for improved reliability of GaN HEMTs which reduces hot electron generation under DC and large-signal RF operation while preserving device performance. The device concept was demonstrated by CMC simulations and it was

benchmarked against simulations of a T-gate device whose model was calibrated to experimental data. Simulations of the electron EDF in both DC and RF operation show that under the same operating conditions, the peak hot-carrier generation obtained with an Asymmetric-Π-gate is lower up to 41% in DC, 44% in Class AB at 10 GHz and 50 % in Class AB at 40 GHz with respect to EDFs of T-gate HEMTs, which suggests improved robustness to failures from the generation of traps induced by hot-carriers. Also, the design rules along with projections of small-signal AC parameters were presented as guidelines for design and fabrication.

Finally, chapter 6 was dedicated to the electro-thermal characterization of an experimental GaN HEMT fabricated on Si substrate, focusing on the active area of the device. Simulations of the DC characteristics were obtained with an expanded CMC framework, where thermal effects are included through an energy-balance equation for phonons. In this approach, the heat generation term or forcing function was computed by tracking the energy exchange during scattering events, which self-consistently couples the charge and heat transport. The model was validated to experimental data of the DC characteristics showing excellent agreement throughout the entire $I_{DS}(V_{GS} - V_{DS})$ space. In addition, by means of the accurate extraction of thermal maps of the device it was found that the position of the hot spot in the channel is not at the peak electric field, but is actually shifted towards the drain by up to 32nm, unlike what is typically reported. The electro-thermal model of the experimental device was then used to asses the impact of lateral scaling in HEMTs, considering the cases of scaling L_{SG} and L_{GD} separately. The results indicate that decreasing the lateral dimensions of the device significantly improve the electrical performance, but at the same time devices are subject to higher temperatures which constitute a reliability issue that should be addressed through efficient thermal management techniques in order to fully take advantage of lateral scaling.

The proposed future work consists on expanding the electro-thermal studies of GaN-HEMTs considering the effects of different substrate materials such as SiC and Diamond, and also increasing the accuracy of the electro-thermal model by adding the effect of thermal resistances at material interfaces. Including the full effect of the substrates, will enable the possibility of evaluating thermal management strategies to reduce the self-heating effects. Another interesting expansion of this work consists in evaluating the reliability of novel HEMT devices fabricated in N-polar GaN, as well as studying complex configurations of HEMTs that include multiple fingers and cascode structures recently proposed to achieve enhancement mode devices.

REFERENCES

[1] N. Alliance, "5G white paper," *Next generation mobile networks, white paper*, 2015.

[2] C. J. Hansen, "WiGiG: Multi-gigabit wireless communications in the 60 GHz band," *IEEE Wireless Communications*, vol. 18, no. 6, 2011.

[3] U. K. Mishra, L. Shen, T. E. Kazior, and Y.-F. Wu, "GaN-based RF power devices and amplifiers," *Proceedings of the IEEE*, vol. 96, no. 2, pp. 287–305, 2008.

[4] I. Keith Benson, Analog Devices. GaN Breaks Barriers RF Power Amplifiers Go Wide and High. (Date last accessed March 2018). [Online]. Available: https://www.analog.com/en/analog-dialogue/articles/rf-power-amplifiers-go-wide-and-high.html

[5] P. Hower, J. Lin, S. Haynie, S. Paiva, R. Shaw, and N. Hepfinger, "Safe operating area considerations in LDMOS transistors," in *Power Semiconductor Devices and ICs, 1999. ISPSD'99. Proceedings., The 11th International Symposium on*. IEEE, 1999, pp. 55–58.

[6] M. Mudholkar, S. Ahmed, M. N. Ericson, S. S. Frank, C. L. Britton, and H. A. Mantooth, "Datasheet driven silicon carbide power MOSFET model," *IEEE Transactions on Power Electronics*, vol. 29, no. 5, pp. 2220–2228, 2014.

[7] S. M. Sze and K. K. Ng, *Physics of semiconductor devices*. John wiley & sons, 2006.

[8] R. S. Pengelly, S. M. Wood, J. W. Milligan, S. T. Sheppard, and W. L. Pribble, "A review of GaN on SiC high electron-mobility power transistors and MMICs," *IEEE Transactions on Microwave Theory and Techniques*, vol. 60, no. 6, pp. 1764–1783, 2012.

[9] R. Dingle, H. Störmer, A. Gossard, and W. Wiegmann, "Electron mobilities in modulation-doped semiconductor heterojunction superlattices," *Applied Physics Letters*, vol. 33, no. 7, pp. 665–667, 1978.

[10] J. Ibbetson, P. Fini, K. Ness, S. DenBaars, J. Speck, and U. Mishra, "Polarization effects, surface states, and the source of electrons in AlGaN/GaN heterostructure field effect transistors," *Appl. Phys. Lett.*, vol. 77, no. 2, pp. 250–252, 2000.

[11] Y.-F. Wu, A. Saxler, M. Moore, R. Smith, S. Sheppard, P. Chavarkar, T. Wisleder, U. Mishra, and P. Parikh, "30-W/mm GaN HEMTs by field plate optimization," *IEEE Electron Device Lett.*, vol. 25, no. 3, pp. 117–119, 2004, doi: 10.1109/LED.2003.822667.

[12] T. Palacios, A. Chakraborty, S. Rajan, C. Poblenz, S. Keller, S. DenBaars, J. Speck, and U. Mishra, "High-power AlGaN/GaN HEMTs for Ka-band applications," *IEEE Electron Device Lett.*, vol. 26, no. 11, pp. 781–783, 2005, doi: 10.1109/LED.2005.857701.

[13] U. K. Mishra, P. Parikh, and Y.-F. Wu, "AlGaN/GaN HEMTs-an overview of device operation and applications," *Proceedings of the IEEE*, vol. 90, no. 6, pp. 1022–1031, 2002.

[14] S. Chowdhury, Y. Wu, L. Shen, K. Smith, P. Smith, T. Kikkawa, J. Gritters, L. McCarthy, R. Lal, R. Barr *et al.*, "650V Highly Reliable GaN HEMTs on Si Substrates over Multiple Generations: Matching Silicon CMOS Manufacturing Metrics and Process Control," in *Compound Semiconductor Integrated Circuit Symposium (CSICS), 2016 IEEE*. IEEE, 2016, pp. 1–4.

[15] D. Gustafsson, A. Leidenhed, and K. Andersson, "Packaged 7 GHz GaN MMIC doherty power amplifier," in *Compound Semiconductor Integrated Circuit Symposium (CSICS), 2017 IEEE*. IEEE, 2017, pp. 1–4.

[16] G. V. Research. Gallium Nitride (GaN) Semiconductor Devices Market Analysis By Product (GaN Radio Frequency Devices, Opto-semiconductors, Power Semiconductors), By Wafer Size, By Application, By Region, And Segment Forecasts, 2018 - 2025. (Date last accessed March 2018). [Online]. Available: https://www.grandviewresearch.com/industry-analysis/gan-gallium-nitride-semiconductor-devices-market

[17] I. Cree. Cree, Inc. Celebrates 15 Years of GaN MMIC Technology Leadership. (Date last accessed March 2018). [Online]. Available: https://www.wolfspeed.com/downloads/dl/file/id/928/product/113

[18] J. Zúñiga-Pérez, V. Consonni, L. Lymperakis, X. Kong, A. Trampert, S. Fernández-Garrido, O. Brandt, H. Renevier, S. Keller, K. Hestroffer *et al.*, "Polarity in GaN and ZnO: Theory, measurement, growth, and devices," *Applied Physics Reviews*, vol. 3, no. 4, p. 041303, 2016.

[19] A. P. Edwards, J. A. Mittereder, S. C. Binari, D. S. Katzer, D. F. Storm, and J. A. Roussos, "Improved reliability of AlGaN-GaN HEMTs using an NH/sub 3/plasma treatment prior to SiN passivation," *IEEE electron device letters*, vol. 26, no. 4, pp. 225–227, 2005.

[20] S. Dasgupta, A. Raman, J. S. Speck, U. K. Mishra *et al.*, "Experimental demonstration of III-nitride hot-electron transistor with GaN base," *IEEE Electron Device Lett.*, vol. 32, no. 9, pp. 1212–1214, 2011.

[21] R. Soligo, S. Chowdhury, G. Gupta, U. Mishra, and M. Saraniti, "The Role of the Base Stack on the AC Performance of GaN Hot Electron Transistor," *IEEE Electron Device Lett.*, vol. 36, no. 7, pp. 669–671, 2015.

[22] D. Cheney, E. Douglas, L. Liu, C. Lo, Y. Xi, B. Gila, F. Ren, D. Horton, M. Law, D. J. Smith *et al.*, "Reliability studies of AlGaN/GaN high electron mobility transistors," *Semicond. Sci. Technol.*, vol. 28, no. 7, p. 074019, Jun. 2013, doi: 10.1088/0268-1242/28/7/074019.

[23] E. Zanoni, "GaN HEMT reliability research–a white paper–," 2017.

[24] I. Rossetto, M. Meneghini, A. Tajalli, S. Dalcanale, C. De Santi, P. Moens, A. Banerjee, E. Zanoni, and G. Meneghesso, "Evidence of Hot-Electron Effects During Hard Switching of AlGaN/GaN HEMTs," *IEEE transactions on electron devices*, vol. 64, no. 9, pp. 3734–3739, 2017.

[25] M. Meneghini, A. Tajalli, P. Moens, A. Banerjee, E. Zanoni, and G. Meneghesso, "Trapping phenomena and degradation mechanisms in GaN-based power HEMTs," *Materials Science in Semiconductor Processing*, 2017.

[26] M. Meneghini, A. Stocco, R. Silvestri, G. Meneghesso, and E. Zanoni, "Degradation of AlGaN/GaN high electron mobility transistors related to hot electrons," *Appl. Phys. Lett.*, vol. 100, no. 23, p. 233508, Jun. 2012, doi: 10.1063/1.4723848.

[27] K. G. Wilson, "Science, industry, and the new Japanese challenge," *Proceedings of the IEEE*, vol. 72, no. 1, pp. 6–18, 1984.

[28] M. Saraniti, A. Rein, G. Zandler, P. Vogl, and P. Lugli, "An efficient multigrid Poisson solver for device simulations," *IEEE transactions on computer-aided design of integrated circuits and systems*, vol. 15, no. 2, pp. 141–150, 1996.

[29] D. K. Ferry, *Semiconductors: Bonds and bands*. IOP Publishing Ltd, 2013.

[30] M. Lundstrom, *Fundamentals of carrier transport*. Cambridge University Press, 2009.

[31] C. Hamaguchi and C. Hamaguchi, *Basic semiconductor physics*. Springer, 2001, vol. 212.

[32] C. Jacoboni and L. Reggiani, "The Monte Carlo method for the solution of charge transport in semiconductors with applications to covalent materials," *Reviews of modern Physics*, vol. 55, no. 3, p. 645, 1983.

[33] S. Selberherr, *Analysis and simulation of semiconductor devices*. Springer Wien; New York, 1984.

[34] J. J. Liou, A. Ortiz-Conde, and F. Garcia-Sanchez, *Analysis and design of MOSFETs: modeling, simulation, and parameter extraction*. Springer Science & Business Media, 1998.

[35] Y. Zhao, J. Watling, S. Kaya, A. Asenov, and J. Barker, "Drift diffusion and hydrodynamic simulations of Si/SiGe p-MOSFETs," *Materials Science and Engineering: B*, vol. 72, no. 2, pp. 180–183, 2000.

[36] M. V. Fischetti and S. E. Laux, "Monte Carlo analysis of electron transport in small semiconductor devices including band-structure and space-charge effects," *Physical Review B*, vol. 38, no. 14, p. 9721, 1988.

[37] M. Saraniti and S. M. Goodnick, "Hybrid fullband cellular automaton/Monte Carlo approach for fast simulation of charge transport in semiconductors," *IEEE Trans. Electron Devices*, vol. 47, no. 10, pp. 1909–1916, Oct. 2000, doi: 10.1109/16.870571.

[38] J. R. Chelikowsky and M. L. Cohen, "Nonlocal pseudopotential calculations for the electronic structure of eleven diamond and zinc-blende semiconductors," *Physical Review B*, vol. 14, no. 2, p. 556, 1976.

[39] K. Kunc and O. H. Nielsen, "Lattice dynamics of zincblende structure compounds II. Shell model," *Computer Physics Communications*, vol. 17, no. 4, pp. 413–422, 1979.

[40] D. K. Ferry, "The onset of quantization in ultra-submicron semiconductor devices," *Superlattices and Microstructures*, vol. 27, no. 2-3, pp. 61–66, 2000.

[41] I. Knezevic, D. Z. Vasileska, and D. K. Ferry, "Impact of strong quantum confinement on the performance of a highly asymmetric device structure: Monte Carlo particle-based simulation of a focused-ion-beam MOSFET," *IEEE Transactions on Electron Devices*, vol. 49, no. 6, pp. 1019–1026, 2002.

[42] T. P. Chow and R. Tyagi, "Wide bandgap compound semiconductors for superior high-voltage unipolar power devices," *IEEE Trans. Electron Devices*, vol. 41, no. 8, pp. 1481–1483, 1994.

[43] I. Smorchkova, L. Chen, T. Mates, L. Shen, S. Heikman, B. Moran, S. Keller, S. DenBaars, J. Speck, and U. Mishra, "AlN/GaN and (Al,Ga)N/AlN/GaN two-dimensional electron gas structures grown by plasma-assisted molecular-beam epitaxy," *Journal of Applied Physics*, vol. 90, no. 10, pp. 5196–5201, 2001.

[44] T. Zhu and R. A. Oliver, "Unintentional doping in GaN," *Physical Chemistry Chemical Physics*, vol. 14, no. 27, pp. 9558–9573, 2012.

[45] O. Ambacher, J. Smart, J. Shealy, N. Weimann, K. Chu, M. Murphy, W. Schaff, L. Eastman, R. Dimitrov, L. Wittmer *et al.*, "Two-dimensional electron gases induced by spontaneous and piezoelectric polarization charges in N-and Ga-face AlGaN/GaN heterostructures," *Journal of applied physics*, vol. 85, no. 6, pp. 3222–3233, 1999.

[46] L. Esaki and R. Tsu, "Superlattice and Negative Differential Conductivity in Semiconductors," *IBM Journal of Research and Development*, vol. 14, no. 1, pp. 61–65, 1970.

[47] T. Mimura, S. Hiyamizu, T. Fujii, and K. Nanbu, "A new field-effect transistor with selectively doped GaAs/n-AlxGa1-xAs heterojunctions," *Japanese journal of applied physics*, vol. 19, no. 5, p. L225, 1980.

[48] M. A. Khan, J. Van Hove, J. Kuznia, and D. Olson, "High electron mobility GaN/Al x Ga1- x N heterostructures grown by low-pressure metalorganic chemical vapor deposition," *Applied Physics Letters*, vol. 58, no. 21, pp. 2408–2410, 1991.

[49] M. Asif Khan, J. Kuznia, D. Olson, W. Schaff, J. Burm, and M. Shur, "Microwave performance of a 0.25 μm gate AlGaN/GaN heterostructure field effect transistor," *Applied Physics Letters*, vol. 65, no. 9, pp. 1121–1123, 1994.

[50] L. Shen, S. Heikman, B. Moran, R. Coffie, N.-Q. Zhang, D. Buttari, I. Smorchkova, S. Keller, S. DenBaars, and U. Mishra, "AlGaN/AlN/GaN high-power microwave HEMT," *IEEE Electron Device Lett.*, vol. 22, no. 10, pp. 457–459, 2001.

[51] S. Karmalkar, M. S. Shur, and R. Gaska, "GaN-based power high electron mobility transistors," in *Wide Energy Bandgap Electronic Devices*. World Scientific, 2003, pp. 173–216.

[52] D. Bisi, A. Chini, F. Soci, A. Stocco, M. Meneghini, A. Pantellini, A. Nanni, C. Lanzieri, P. Gamarra, C. Lacam *et al.*, "Hot-electron degradation of AlGaN/GaN high-electron mobility transistors during RF operation: correlation with GaN buffer design," *IEEE Electron Device Letters*, vol. 36, no. 10, pp. 1011–1014, 2015.

[53] M. Hikita, M. Yanagihara, K. Nakazawa, H. Ueno, Y. Hirose, T. Ueda, Y. Uemoto, T. Tanaka, D. Ueda, and T. Egawa, "AlGaN/GaN power HFET on silicon substrate with source-via grounding (SVG) structure," *IEEE Transactions on Electron Devices*, vol. 52, no. 9, pp. 1963–1968, 2005.

[54] A. Winzer, R. Goldhahn, G. Gobsch, A. Link, M. Eickhoff, U. Rossow, and A. Hangleiter, "Determination of the polarization discontinuity at the Al Ga N/ Ga N interface by electroreflectance spectroscopy," *Applied Physics Letters*, vol. 86, no. 18, p. 181912, 2005.

[55] F. A. Marino, D. A. Cullen, D. J. Smith, M. R. McCartney, and M. Saraniti, "Simulation of polarization charge on AlGaN/GaN high electron mobility transistors: Comparison to electron holography," *J. Appl. Phys.*, vol. 107, no. 5, p. 054516, 2010.

[56] S. E. Laux, "Techniques for small-signal analysis of semiconductor devices," *IEEE Transactions on Electron Devices*, vol. 32, no. 10, pp. 2028–2037, 1985.

[57] H. Rao and G. Bosman, "Hot-electron induced defect generation in AlGaN/GaN high electron mobility transistors," *Solid-State Electronics*, vol. 79, pp. 11–13, 2013.

[58] A. Sozza, C. Dua, E. Morvan, S. Delage, F. Rampazzo, A. Tazzoli, F. Danesin, G. Meneghesso, E. Zanoni, A. Curutchet *et al.*, "Evidence of traps creation in gan/algan/gan hemts after a 3000 hour on-state and off-state hot-electron stress," in *Electron Devices Meeting, 2005. IEDM Technical Digest. IEEE International*. IEEE, 2005, pp. 4–pp.

[59] Y. Puzyrev, S. Mukherjee, J. Chen, T. Roy, M. Silvestri, R. D. Schrimpf, D. M. Fleetwood, J. Singh, J. M. Hinckley, A. Paccagnella *et al.*, "Gate bias dependence of defect-mediated hot-carrier degradation in GaN HEMTs," *IEEE Trans. Electron Devices*, vol. 61, no. 5, pp. 1316–1320, Mar. 2014, doi: 10.1109/TED.2014.2309278.

[60] S. T. Pantelides, Y. Puzyrev, X. Shen, T. Roy, S. DasGupta, B. R. Tuttle, D. M. Fleetwood, and R. D. Schrimpf, "Reliability of III–V devices–The defects that cause the trouble," *Microelectronic Engineering*, vol. 90, pp. 3–8, 2012.

[61] R. Gaska, A. Osinsky, J. Yang, and M. S. Shur, "Self-heating in high-power AlGaN-GaN HFETs," *IEEE Electron Device Letters*, vol. 19, no. 3, pp. 89–91, 1998.

[62] S. Martin-Horcajo, J. W. Pomeroy, B. Lambert, H. Jung, H. Blanck, and M. Kuball, "Transient thermoreflectance for gate temperature assessment in pulse operated GaN-based HEMTs," *IEEE Electron Device Letters*, vol. 37, no. 9, pp. 1197–1200, 2016.

[63] M. O. Manasreh, *III-nitride semiconductors: electrical, structural and defects properties*. Elsevier, 2000.

[64] C. I. Ashby, C. C. Mitchell, J. Han, N. A. Missert, P. P. Provencio, D. M. Follstaedt, G. M. Peake, and L. Griego, "Low-dislocation-density GaN from a single growth on a textured substrate," *Applied Physics Letters*, vol. 77, no. 20, pp. 3233–3235, 2000.

[65] R. Joshi, V. Sridhara, B. Jogai, P. Shah, and R. Del Rosario, "Analysis of dislocation scattering on electron mobility in GaN high electron mobility transistors," *Journal of applied physics*, vol. 93, no. 12, pp. 10 046–10 052, 2003.

[66] M. Gurusinghe and T. Andersson, "Mobility in epitaxial GaN: Limitations of free-electron concentration due to dislocations and compensation," *Physical Review B*, vol. 67, no. 23, p. 235208, 2003.

[67] N. G. Weimann, L. F. Eastman, D. Doppalapudi, H. M. Ng, and T. D. Moustakas, "Scattering of electrons at threading dislocations in GaN," *Journal of Applied Physics*, vol. 83, no. 7, pp. 3656–3659, 1998.

[68] J. Joh and J. A. del Alamo, "Mechanisms for electrical degradation of GaN high-electron mobility transistors," in *Electron Devices Meeting, 2006. IEDM'06. International*. IEEE, 2006, pp. 1–4.

[69] J. A. del Alamo and J. Joh, "GaN HEMT reliability," *Microelectronics reliability*, vol. 49, no. 9-11, pp. 1200–1206, 2009.

[70] K. Smith, J. Haller, J. Guerrero, R. Smith, R. Lal, and Y.-F. Wu, "Lifetime tests of 600-V GaN-on-Si power switches and HEMTs," *Microelectronics Reliability*, vol. 58, pp. 197–203, 2016.

[71] T. Brazzini, M. A. Casbon, H. Sun, M. J. Uren, J. Lees, P. J. Tasker, H. Jung, H. Blanck, and M. Kuball, "Electroluminescence of hot electrons in AlGaN/GaN high-electron-mobility transistors under radio frequency operation," *Appl. Phys. Lett.*, vol. 106, no. 21, p. 213502, May 2015, doi: 10.1063/1.4921848.

[72] T. Brazzini, M. A. Casbon, M. J. Uren, P. J. Tasker, H. Jung, H. Blanck, and M. Kuball, "Hot-Electron Electroluminescence Under RF Operation in GaN-HEMTs: A Comparison Among Operational Classes," *IEEE Transactions on Electron Devices*, vol. 64, no. 5, pp. 2155–2160, 2017.

[73] F. H. Raab, P. Asbeck, S. Cripps, P. B. Kenington, Z. B. Popovic, N. Pothecary, J. F. Sevic, and N. O. Sokal, "Power amplifiers and transmitters for RF and microwave," *IEEE transactions on Microwave Theory and Techniques*, vol. 50, no. 3, pp. 814–826, 2002.

[74] S. C. Cripps, *RF Power Amplifiers for Wireless Communications*. Artech house Inc, 2006.

[75] D. M. Pozar, *Microwave engineering*. John Wiley & Sons, 2009.

[76] H. L. Krauss, C. W. Bostian, and F. H. Raab, *Solid state radio engineering*. Wiley New York, 1980, vol. 47103018.

[77] H. Statz, P. Newman, I. W. Smith, R. A. Pucel, and H. A. Haus, "GaAs FET device and circuit simulation in SPICE," *IEEE Transactions on Electron Devices*, vol. 34, no. 2, pp. 160–169, 1987.

[78] W. Lu, L. Wang, S. Gu, D. P. Aplin, D. M. Estrada, K. Paul, and P. M. Asbeck, "Analysis of reverse leakage current and breakdown voltage in GaN and InGaN/GaN Schottky barriers," *IEEE Trans. Electron Devices*, vol. 58, no. 7, pp. 1986–1994, 2011.

[79] M. Shur, A. Bykhovski, R. Gaska, M. Khan, and J. Yang, "AlGaN–GaN–AlInGaN induced base transistor," *Appl. Phys. Lett.*, vol. 76, no. 22, pp. 3298–3300, 2000.

[80] G. Gupta, M. Laurent, J. Lu, S. Keller, and U. K. Mishra, "Design of polarization-dipole-induced isotype heterojunction diodes for use in III–N hot electron transistors," *Applied Physics Express*, vol. 7, no. 1, p. 014102, 2014.

[81] M. Miyoshi, A. Imanishi, T. Egawa, H. Ishikawa, K.-i. Asai, T. Shibata, M. Tanaka, and O. Oda, "DC characteristics in high-quality AlGaN/AlN/GaN high-electron-mobility transistors grown on AlN/sapphire templates," *Japanese journal of applied physics*, vol. 44, no. 9R, p. 6490, 2005.

[82] S. W. Kaun, P. G. Burke, M. Hoi Wong, E. C. Kyle, U. K. Mishra, and J. S. Speck, "Effect of dislocations on electron mobility in AlGaN/GaN and AlGaN/AlN/GaN heterostructures," *Appl. Phys. Lett.*, vol. 101, no. 26, p. 262102, 2012.

[83] B. Mazumder, S. W. Kaun, J. Lu, S. Keller, U. K. Mishra, and J. S. Speck, "Atom probe analysis of AlN interlayers in AlGaN/AlN/GaN heterostructures," *Appl. Phys. Lett.*, vol. 102, no. 11, p. 111603, 2013.

[84] O. Ambacher, J. Majewski, C. Miskys, A. Link, M. Hermann, M. Eickhoff, M. Stutzmann, F. Bernardini, V. Fiorentini, V. Tilak *et al.*, "Pyroelectric properties of Al (In) GaN/GaN hetero-and quantum well structures," *Journal of physics: condensed matter*, vol. 14, no. 13, p. 3399, 2002.

[85] Y. Zhang and J. Singh, "Charge control and mobility studies for an AlGaN/GaN high electron mobility transistor," *Journal of Applied Physics*, vol. 85, no. 1, pp. 587–594, 1999.

[86] S. Keller, S. Heikman, I. Ben-Yaacov, L. Shen, S. DenBaars, and U. Mishra, "Indium-surfactant-assisted growth of high-mobility AlN/GaN multilayer structures by metalorganic chemical vapor deposition," *Appl. Phys. Lett.*, vol. 79, no. 21, pp. 3449–3451, 2001.

[87] C.-H. Chen, S. M. Baier, D. K. Arch, and M. S. Shur, "A new and simple model for GaAs heterojunction FET gate characteristics," *IEEE Trans. Electron Devices*, vol. 35, no. 5, pp. 570–577, 1988.

[88] Y. Lv, Z. Lin, T. D. Corrigan, J. Zhao, Z. Cao, L. Meng, C. Luan, Z. Wang, and H. Chen, "Extraction of AlGaN/GaN heterostructure Schottky diode barrier heights from forward current-voltage characteristics," *Journal of Applied Physics*, vol. 109, no. 7, p. 074512, 2011.

[89] D. Guerra, D. K. Ferry, S. M. Goodnick, M. Saraniti, and F. A. Marino, "Large-signal full-band Monte Carlo device simulation of millimeter-wave power GaN HEMTs with the inclusion of parasitic and reliability issues," in *Simulation of Semiconductor Processes and Devices (SISPAD), 2011 International Conference on.* IEEE, Oct. 2011, pp. 87–90, doi: 10.1109/SISPAD.2011.6 035 056.

[90] Y. Puzyrev, T. Roy, M. Beck, B. Tuttle, R. Schrimpf, D. Fleetwood, and S. Pantelides, "Dehydrogenation of defects and hot-electron degradation in GaN high-electron-mobility transistors," *J. Appl. Phys.*, vol. 109, no. 3, p. 034501, Feb. 2011, doi: 10.1063/1.3524185.

[91] S. Mukherjee, Y. Puzyrev, J. Chen, D. M. Fleetwood, R. D. Schrimpf, and S. T. Pantelides, "Hot-Carrier Degradation in GaN HEMTs Due to Substitutional Iron and Its Complexes," *IEEE Trans. Electron Devices*, vol. 63, no. 4, pp. 1486–1494, Mar. 2016, doi: 10.1109/TED.2016.2532806.

[92] J. Joh, J. A. Del Alamo, U. Chowdhury, and J. L. Jimenez, "Correlation between RF and DC reliability in GaN high electron mobility transistors," in *ROCS Workshop, 2008 [Reliability of Compound Semiconductors Workshop].* IEEE, Jun. 2008, pp. 185–194, doi: 10.1109/ROCS.2008.5 483 626.

[93] A. Chini, F. Fantini, V. Di Lecce, M. Esposto, A. Stocco, N. Ronchi, F. Zanon, G. Meneghesso, and E. Zanoni, "Correlation between DC and rf degradation due to deep levels in AlGaN/GaN HEMTs," in *Proc. IEEE Int. Electron Devices Meeting (IEDM)*. IEEE, Mar. 2009, pp. 1–4, doi: 10.1109/IEDM.2009.5 424 394.

[94] J. Joh and J. A. Del Alamo, "RF power degradation of GaN high electron mobility transistors," in *Proc. IEEE Int. Electron Devices Meeting (IEDM)*. IEEE, Jan. 2010, pp. 20–2, doi: 10.1109/IEDM.2010.5 703 397.

[95] M. Caesar, M. Dammann, V. Polyakov, P. Waltereit, W. Bronner, M. Baeumler, R. Quay, M. Mikulla, and O. Ambacher, "Generation of traps in AlGaN/GaN HEMTs during RF-and DC-stress test," in *Reliability Physics Symposium (IRPS), 2012 IEEE International*. IEEE, Oct. 2012, pp. CD–6, doi: 10.1109/IRPS.2012.6 241 883.

[96] R. C. Fitch, D. E. Walker, A. J. Green, S. E. Tetlak, J. K. Gillespie, R. D. Gilbert, K. A. Sutherlin, W. D. Gouty, J. P. Theimer, G. D. Via *et al.*, "Implementation of High-Power-Density X-Band AlGaN/GaN High Electron Mobility Transistors in a Millimeter-Wave Monolithic Microwave Integrated Circuit Process," *IEEE Electron Device Lett.*, vol. 36, no. 10, pp. 1004–1007, 2015.

[97] S. Yamakawa, R. Akis, N. Faralli, M. Saraniti, and S. M. Goodnick, "Rigid ion model of high field transport in GaN," *Journal of Physics: Condensed Matter*, vol. 21, no. 17, p. 174206, 2009.

[98] W. Hackbusch, "The frequency decomposition multi-grid algorithm," in *Robust Multi-Grid Methods*. Springer, 1989, pp. 96–104.

[99] F. A. Marino, N. Faralli, T. Palacios, D. K. Ferry, S. M. Goodnick, and M. Saraniti, "Effects of threading dislocations on AlGaN/GaN high-electron mobility transistors," *IEEE Trans. Electron Devices*, vol. 57, no. 1, pp. 353–360, 2010.

[100] H. Tokuda, J. Yamazaki, and M. Kuzuhara, "High temperature electron transport properties in AlGaN/GaN heterostructures," *J. Appl. Phys.*, vol. 108, no. 10, p. 104509, 2010.

[101] G. Meneghesso, M. Meneghini, I. Rossetto, D. Bisi, S. Stoffels, M. Van Hove, S. Decoutere, and E. Zanoni, "Reliability and parasitic issues in GaN-based power HEMTs: a review," *Semicond. Sci. Technol.*, vol. 31, no. 9, p. 093004, Aug. 2016, doi: 10.1088/0268-1242/31/9/093004.

[102] J. Möreke, M. Ťapajna, M. J. Uren, Y. Pei, U. K. Mishra, and M. Kuball, "Effects of gate shaping and consequent process changes on AlGaN/GaN HEMT reliability," *physica status solidi (a)*, vol. 209, no. 12, pp. 2646–2652, 2012.

[103] J. Ma and E. Matioli, "High Performance Tri-Gate GaN Power MOSHEMTs on Silicon Substrate," *IEEE Electron Device Lett.*, vol. 38, no. 3, pp. 367–370, 2017, doi: 10.1109/LED.2017.2661755.

[104] H. S. Yoon, B.-G. Min, J. M. Lee, D. M. Kang, H.-K. Ahn, H. Kim, and J. Lim, "Microwave Low-Noise Performance of 0.17 μm Gate-Length AlGaN/GaN HEMTs on SiC With Wide Head Double-Deck T-Shaped Gate," *IEEE Electron Device Lett.*, vol. 37, no. 11, pp. 1407–1410, Sep. 2016, doi: 10.1109/LED.2016.2612624.

[105] H. Y. Wong, N. Braga, and R. Mickevicius, "Normally-Off GaN HFET Based on Layout and Stress Engineering," *IEEE Electron Device Lett.*, vol. 37, no. 12, pp. 1621–1624, Oct. 2016, doi: 10.1109/LED.2016.2619978.

[106] J. Joh and J. A. Del Alamo, "Impact of gate placement on RF power degradation in GaN high electron mobility transistors," *Microelectron. Rel.*, vol. 52, no. 1, pp. 33–38, 2012, doi: 10.1016/j.microrel.2011.09.008.

[107] D. Fanning, L. Witkowski, J. Stidham, H. Tserng, M. Muir, and P. Saunier, "Dielectrically defined optical T-gate for high power GaAs pHEMTs," *2002 GaAs MANTECH Digest of Papers*, pp. 83–86, 2002.

[108] Y. Yamashita, A. Endoh, K. Shinohara, M. Higashiwaki, K. Hikosaka, T. Mimura, S. Hiyamizu, and T. Matsui, "Ultra-short 25-nm-gate lattice-matched InAlAs/InGaAs HEMTs within the range of 400 GHz cutoff frequency," *IEEE Electron Device Lett.*, vol. 22, no. 8, pp. 367–369, 2001, doi: 10.1109/55.936345.

[109] A. D. Latorre-Rey, F. Sabatti, J. D. Albrecht, and M. Saraniti, "Hot electron generation under large-signal radio frequency operation of GaN high-electron-mobility transistors," *Appl. Phys. Lett.*, vol. 111, no. 1, Jul. 2017.

[110] A. D. Latorre-Rey, K. Merrill, J. D. Albrecht, and M. Saraniti, "Electro-Thermal Characterization of GaN HEMT on Si through self-consistent Energy balance-Cellular Monte Carlo Device Simulations," in *Proc. Compound Semiconductor Integrated Circuit Symposium (CSICS), 2017,.* IEEE, Oct. 2017, pp. 1–4.

[111] A. D. Latorre-Rey, J. D. Albrecht, and M. Saraniti, "Generation of Hot Electrons in GaN HEMTs under RF Class A and AB PAs," in *Proc. Device Research Conference (DRC), 2017 75th Annual.* IEEE, Jun. 2017, pp. 1–2.

[112] P. Altuntas, F. Lecourt, A. Cutivet, N. Defrance, E. Okada, M. Lesecq, S. Rennesson, A. Agboton, Y. Cordier, V. Hoel *et al.*, "Power performance at 40GHz of AlGaN/GaN high-electron mobility transistors grown by molecular beam epitaxy on Si(111) substrate," *IEEE Electron Device Lett.*, vol. 36, no. 4, pp. 303–305, 2015.

[113] K. J. Chen, O. Haberlen, A. Lidow, C. lin Tsai, T. Ueda, Y. Uemoto, and Y. Wu, "GaN-on-Si Power Technology: Devices and Applications," *IEEE Trans. Electron Devices*, vol. 64, no. 3, pp. 779–795, 2017.

[114] K. Raleva, D. Vasileska, S. M. Goodnick, and M. Nedjalkov, "Modeling thermal effects in nanodevices," *IEEE Trans. Electron Devices*, vol. 55, no. 6, pp. 1306–1316, 2008.

[115] Q. Hao, H. Zhao, and Y. Xiao, "A hybrid simulation technique for electrothermal studies of two-dimensional GaN-on-SiC high electron mobility transistors," *Journal of Applied Physics*, vol. 121, no. 20, pp. 204 501–204 508, 2017.

[116] W. Dai and R. Nassar, "A finite difference scheme for solving the heat transport equation at the microscale," *Numerical Methods for Partial Differential Equations*, vol. 15, no. 6, pp. 697–708, 1999.

[117] V. Palankovski and S. Selberherr, "Thermal models for semiconductor device simulation," in *High Temperature Electronics, 1999. HITEN 99. The Third European Conference on.* IEEE, 1999, pp. 25–28.

[118] F. Bonani and G. Ghione, "On the application of the Kirchhoff transformation to the steady-state thermal analysis of semiconductor devices with temperature-dependent and piecewise inhomogeneous thermal conductivity," *Solid-state electronics*, vol. 38, no. 7, pp. 1409–1412, 1995.

[119] K. Tsen, D. Ferry, A. Botchkarev, B. Sverdlov, A. Salvador, and H. Morkoc, "Time-resolved Raman studies of the decay of the longitudinal optical phonons in wurtzite GaN," *Appl. Phys. Lett.*, vol. 72, no. 17, pp. 2132–2134, 1998.

[120] J. W. Chung, W. E. Hoke, E. M. Chumbes, and T. Palacios, "AlGaN/GaN HEMT With 300-GHz f_{max}," *IEEE Electron Device Letters*, vol. 31, no. 3, pp. 195–197, 2010.

[121] K. Shinohara, D. C. Regan, Y. Tang, A. L. Corrion, D. F. Brown, J. C. Wong, J. F. Robinson, H. H. Fung, A. Schmitz, T. C. Oh *et al.*, "Scaling of GaN HEMTs and Schottky diodes for submillimeter-wave MMIC applications," *IEEE Transactions on Electron Devices*, vol. 60, no. 10, pp. 2982–2996, 2013.

[122] P. J. Tasker and B. Hughes, "Importance of source and drain resistance to the maximum f/sub T/of millimeter-wave MODFETs," *IEEE Electron Device Letters*, vol. 10, no. 7, pp. 291–293, 1989.

[123] K. Shinohara, D. Regan, I. Milosavljevic, A. Corrion, D. Brown, P. Willadsen, C. Butler, A. Schmitz, S. Kim, V. Lee *et al.*, "Electron velocity enhancement in laterally scaled GaN DH-HEMTs with $f\_\{T\}$ of 260 GHz," *IEEE Electron Device Letters*, vol. 32, no. 8, pp. 1074–1076, 2011.

[124] R. Soligo, D. Guerra, D. K. Ferry, S. M. Goodnick, and M. Saraniti, "Cellular Monte Carlo study lateral scaling impact of on the DC-RF performance of high-power GaN HEMTs," in *Computational Electronics (IWCE), 2012 15th International Workshop on.* IEEE, 2012, pp. 1–4.

Biography Alvaro D. Latorre-Rey

Alvaro D. Latorre-Rey was born in Caracas, Venezuela. He received the Professional Electronics Engineering degree (Cum Laude) in 2009 and the Magister of Electronics Engineering degree (with Honors) in 2011, with a major in solid-state electronics, both from Universidad Simon Bolivar, Caracas-Venezuela. He obtained the Ph.D. degree in electrical engineering in April 2018 from Arizona State University in Tempe, AZ, USA. His research was focused on the modeling of reliability for mm-wave GaN HEMTs by means of Monte Carlo device simulations, working under the guidance of Dr. Marco Saraniti within the Center for Computational Nanoscience. From May to December 2017, he held an internship at ON Semiconductor Corporation within the Power Solutions Group where he focused on the compact modeling of TMBS and JBS power rectifiers. His research interests are in the field of design, simulation (Monte Carlo and drift-diffusion methods), physics-based compact modeling and characterization of high power and RF semiconductor devices such as HEMTs, MOSFETs and diodes, in silicon and wide band gap materials like GaN and SiC. He is an IEEE Eta Kappa Nu member and the author or coauthor of several articles published in technical journals and specialized conferences.