50376
1994
323

Avec ma gratitude

A mes parents

A Howaida

A Mahmoud

Remerciements

Cette étude a été effectuée à l'Institut d'Electronique et Microélectronique du Nord (IEMN. UMR-CNRS No.9929), dirigé par Monsieur le Professeur E.Constant, dans le Department Hyperfréquences et Semiconducteurs (DHS) dirigé par Monsieur le Professeur G.Salmer. Je tiens à témoigner a Monsieur Salmer toute ma reconnaisance pour avoir assuré la direction de ce travail.

Je remercie Monsieur le Professeur Y.Crosnier d'avoir accepté de présider la commision d'examen de cette thèse et également pour les discussions et ses encouragements tout au long de ce travail.

Je remercie Monsieur le.Professeur M.J.Howes, directeur de departement de Genie Electrique et Electronique à l'Université de Leeds (Grande Bretagne), Monsieur E.Caquot Chef de Département au Centre National d'Etudes des Télécommunications (Bagneux), qui me font l'honneur de juger ce travail et d'en être les rapporteurs.

Je remercie Monsieur O.L.El-Sayed, Professeur à l'Universtié du Caire et Conseiller Culturel d'Egypte en France, Monsieur P.Hesto Professeur à l'Institut d'Electronique Fondamentale (I.E.F. Paris XI), Monsieur J.Favre de Thomson C.S.F. et Monsieur J.C.DeJaeger Professeur à l'Université des Sciences et Technologies de Lille qui me font l'honneur de participer à la commision d'examen.

Je remercie la Mission Francaise de Recherche et de Coopération au Caire pour m'avoir proposé une bourse d'étude tout au long de mon travail depuis 1988.

Les simulations ont été effectuées sur le IBM-3090 de l'ancien Centre Inter Régional de Calcul Electronique CIRCE (actuellement IDRIS), je tiens à remercier les personnels d'assitance pour leurs aides.

Les traitements graphiques ainsi que la rédaction de ce memoire ont été effectués au Centre Interuniversitaire de Traitement de l'Information (CITI). Sans l'aide du personnel de ce centre, ce travail n'aurait jamais abouti; je tiens à remercie G.Dubus Chef du service, P.Billa, E.Cassette, Y.Tinel, D.Marant, P.Declerque pour leur étroite coopération. Ils n'ont jamais hésité à m'apporter de l'aide chaque fois qu'il était nécessaire. Je ne saurais jamais oublier les personnels de permanence technique: J.M. Delacroix, A.Haddoux, B.Hadoux, qu'ils trouvent ici mes plus vifs remerciements.

J'adresse également mes remerciements aux collegues qui m'ont toujours soutenus et avec qui j'ai beaucoup discuté: O.Vanbesien, H.Happy, G.Dambrine, Y.Gobert, D.Chauvel, et L.Picheta.

Mes remerciements sont aussi adressés à Mesdames J.Charvet, N.Castelein et M.Farge pour leur sympathie et leur coopération.

Enfin, je tiens à remercier Monsieur J.P.Dehorter qui a assuré la reprographie de ce mémoire.

This page is intentionally left blank

This page is intentionally left blank

PREFACE

The High Electron Mobility Transistors have shown unprecedented performance in microwave low noise and high power applications.

Thanks to advanced layers growth techniques and efficient lithographic methods, it is possible to fabricate transistors with complex epilayer structure and extremely short gates and inter electrode spacing. The use of nouvelle material compositions and the techniques of band-gap tailoring gave way to the high performance margins demonstrated by such transistors. All these have rendered the HEMT's the favourite devices when microwaves are considered. However, with the continuous demand on the frequency rise, the use of such transistors has shown that there exists always an upper limit for a specific device in the operating frequencies. In most cases it is found that the reason is unoptimised structures. In order to optimise HEMT's, a complicated procedure of investigating the effect of different structural parameters has to be performed. Also, the optimisation approaches differ with respect to the foreseen application whether it is low noise, power or non-linear mixing applications.

Today the advances in material engineering yield up a wide range of material systems that could be used for HEMT's fabrication. Other than the conventional system AlGaAs/GaAs we have AlInAs/InGaAs lattice matched to InP substrates, InGaAs pseudomorphic channels, metamorphic systems and the list is unlimited.

The advances in lithographic technology permitted also the fabrication of extremely small gates and inter electrodes spacing scaling down the device dimensions to a largely sub-half-micron range. The advances in this domain permit the fine tailoring of recess shapes and give full liberty in the positioning of the gate.

Concerning the epilayer growth, the existing layer deposition techniques give unlim-

ited powers in fabricating epilayers with fine thickness, alternat layer systems and introducing unlimited range of doping profiles and densities. This is not limited to the donor layers only, but extends to the channel and buffer layers as well.

It is obvious then that a unique investigation based upon an experimental approach would be exhaustive both from the economical as well as the time point of view.

The solution would then be in the use of a simulation procedure to get the required investigation done. This simulation method has to satisfy two major criteria : *rapidity and precision.*

In the rush of technological advancements, device simulation has evolved from just a research tool used by physicists to an indispensable tool for device engineers integrated in CAD environments. Both physicists and device engineers rely nowadays strongly on simulation each for his reasons. Whereas the physicist adopts the physical insight perspective of device simulation to better understand its behaviour and material properties with a goal of developing new ones, the device engineer seeks information about the device terminal characteristics and looks for the optimisation of existing ones. Thus, the first needs a device simulation tool as physical as possible, and the second needs a tool physical enough to be accurate and at the same time not too slow.

Scanning the existing possibilities for simulation, the Two Dimensional Hydrodynamic Energy Modelling is the most adequate for HEMT simulation since its structure is fairly simple and it could present a good physical frame for FET study.

Throughout the years, there has been a lot of work done to design and to master a reliable hydrodynamic energy model. We tried to add to this work in an attempt to improve the capacity and the capabilities of hydrodynamic modelling for device simulation.

The purpose of this work

1- To develop a reliable precise physical simulation tool suitable for pseudomorphic HEMT's for power applications; it should account for as much as possible of the physical phenomena existing in the complex device.

2- Use the simulator in a systematic way to investigate the effect of various parameters upon the device performance, let them be geometric or material.

3- Derive broad outlines concerning the device optimisation for efficient power operation in millimetre wave range.

In the following chapter, a review of HEMT's principle of operation will be presented. The advantages of using HEMT's for power applications are highlighted. We also point out the limitations of conventional structures. A review of different solutions is presented then the concept of pseudomorphic devices is discussed. The state of art in power performance is then pointed out where the limitations of actual devices underlined. Following, a presentation of some particular characteristics of pseudomorphic HEMT's is made prior to the discussion of the simulation procedures done to include them.

In chapter two the model developed is presented and discussed, its main features and capabilities are highlighted from the three major aspects, physical, numerical and computational. Special emphasis will be given to the treatment of the new physical phenomena introduced as well as the features added to simplify the model handling.

In chapter three, we present a study of pseudomorphic HEMT's. We compare them to conventional HEMT's, and we compare uniform doping with pulse doping as well as a study of the buffer effect. We then present a study of the transient time instabilities in HEMT's in general with our analysis and conclusions.

Chapter four deals with the structural optimisation of pseudomorphic HEMT's where the effect of varying different parameters is presented together with the major conclusions of such a study.

In chapter five, we discuss the avalanche multiplication phenomena in HEMT's, where we present the modifications done to permit studying such phenomena. The effect of introducing a recess offset is studied and major results discussed.

Finally in chapter six we present the major conclusions of this work as well as the possible extensions.

This page is intentionally left blank

Contents

1 Les HEMTs de puissance **xxvii**

1.1 Introduction xxviii

1.2 Spécifications D'un HEMT de puissance xxviii

1.3 Composant de Choix xxviii

2 Moyens de Simulation **xxxi**

2.1 Choix de Modèle xxxii

2.2 Le modèle physique xxxii

2.3 Nouveaux effets physiques introduits xxxiii

2.3.1 Caractère bidimensionnel des électrons xxxiii

2.3.2 Effet d'écran xxxiv

2.3.3 L'ionisation par choc xxxiv

2.4 Simulation des phénomènes pris en compte xxxv

2.4.1 Introduction des effets quantiques xxxv

2.4.2 Simulation de l'effet d'écran xxxvi

2.5 Simulation de la multiplication par avalanche xxxvii

2.5.1 Taux de génération dû aux électrons xxxvii

2.5.2 Taux de génération dû aux trous xxxviii

2.6 Quelques éléments particuliers du modèle xxxviii

2.7 Rèsumé du modèle physique final xxxix

2.8 Traitement numérique et algorithmique xxxix

3 Etude des transistors HEMT pseudomorphiques **xli**

3.1 Comparaison avec le transistor conventionnel xlii

3.2 Structures à plan de dopage xliii

3.3 Nature de la couche tampon xliv

3.4 Les instabilités temporelles dans les HEMT xliv

4 Optimisation des HEMT pseudomorphiques **xlvii**

4.1 Le but de l'optimisation xlviii

4.2 L'épaisseur de la couche non dopée xlviii

4.3 Longueur de grille xlix

4.4 Offset de recess l

4.5 Optimisation de n_s l

5 Simulation des HEMTs en claquage **liii**

5.1 Analyses physiques liv

5.2 Effet des dimensions des recess liv

6 Conclusion **lvii**

6.1 Limites actulles et Ameliorationes potentielles lix

I Power FET's : A look through **1**

I.1 Introduction 2

I.2 Power requirements 2
I.3 From MESFET's to MODFET's 4
I.4 Modulation doping principle of operation 6
I.5 Alterations on HEMT structures 8
I.6 State of the art in power performance 12
I.7 Pseudomorphic Structures 13
I.7.1 Synthesis of pseudomorphic layers 13
I.8 Optimisation problem of Pseudomorphic HEMT's 16
I.9 Particularities in pseudomorphic HEMT's modelling 17
I.9.1 Carriers properties in strained layers 17
I.9.2 Carriers quantisation 18
I.9.3 Screening effect 20
I.9.4 Impact ionisation 20
I.9.4.1 What is impact ionisation 20
I.9.4.2 The principle of Ionisation energy 22
I.10 Conclusion 24

II SIMFETII - The device simulator **27**

II.1 Introduction 28
II.2 The Boltzmann Transport Equation 30
II.3 Simplifications to solve BTE 31
II.3.1 Relaxation Time Approximation 31
II.3.2 Displaced Maxwellian approximation 32

II.4 The Non Stationary model . . . 33

II.4.1 The general model . . . 33

II.4.2 Collision Terms . . . 35

II.4.3 Single equivalent gas concept . . . 36

II.4.3.1 Physical reasoning . . . 36

II.4.3.2 The average total energy concept . . . 37

II.5 The generalised single gas model . . . 37

II.6 The hydrodynamic energy model . . . 39

II.7 Discussion . . . 41

II.8 Enhancements of the model . . . 43

II.8.1 Simulation of carrier's quantization . . . 43

II.8.2 Simulation of the screening effect . . . 45

II.8.3 Simulation of avalanche multiplication . . . 48

II.8.4 Simulation of minority carriers transport . . . 53

II.8.5 The complete physical model . . . 54

II.8.6 Main Characteristics of the physical model . . . 55

II.9 Numerical approach . . . 56

II.9.1 Spatial discretisation . . . 59

II.9.2 Temporal Discretisation . . . 59

II.9.3 Boundary Conditions . . . 60

II.10 Computational Model . . . 63

II.10.1 The algorithm . . . 65

III A study of Pseudomorphic HEMT's **71**

III.1 Introduction 72

III.2 Device Structure 72

III.3 The device physics: Distribution of main quantities 76

III.3.1 Charge Distribution 76

III.3.2 Energy Distribution 77

III.3.3 Electrostatic Potential 84

III.4 Analysis of Terminal Characteristics 85

III.4.1 Transconductance g_m 85

III.4.2 The gate Capacitance C_{gs} 87

III.4.3 Intrinsic Current Gain cut Off Frequency f_c 89

III.4.4 Drain conductance g_d 91

III.5 Pulse doped HEMT's 93

III.5.1 Introduction 93

III.5.2 Distribution of physical quantities 93

III.5.3 Comparison with uniform doping 95

III.5.3.1 Drain current 95

III.5.3.2 The Gate capacitance 96

III.5.3.3 Effect on the intrinsic Transconductance 97

III.5.3.4 Effect on f_c 97

III.5.4 Recapitulation 101

III.6 Effect of buffer type on device performance 102

III.6.1 Distribution of physical quantities 102

III.6.2 Terminal characteristics. 102

III.6.2.1 Transconductance and current gain cut off frequency 102

III.6.3 Effect on feedback parameters 107

III.7 Recapitulation 109

III.8 On The Transient Instabilities in HEMT's 110

III.8.1 Existence of Gunn oscillations 110

III.8.2 Oscillations in HEMT's 111

III.8.3 Probability of oscillations in HEMT's 120

III.8.4 Effect of loading circuit 123

III.8.5 Discussion 126

IV Optimisation of pulse doped pseudomorphic HEMT's 131

IV.1 Optimisation of HEMT's 132

IV.2 The Device Structure 133

IV.3 Particularities about the Simulation 133

IV.4 Optimisation of Aspect Ratio 137

IV.4.1 Why should we optimise the aspect ratio? 137

IV.4.2 Effect of Undoped AlGaAs Layer Thickness 137

IV.4.2.1 Effect on g_m 137

IV.4.2.2 Effect on Gate capacitance 139

IV.4.2.3 Effect on f_{ci} 141

IV.4.2.4 Recapitulation 141

IV.4.3 Effect of Gate Length 141

IV.4.3.1 Effect on g_m 143

IV.4.3.2 Effect on Gate capacitance 143

IV.4.3.3 Effect on f_{ci} 148

IV.4.3.4 Effect on g_d 150

IV.4.3.5 Recapitulation 154

IV.4.4 Effect of Recess shape 155

IV.4.4.1 Interest in recess offset 155

IV.4.4.2 Effect on f_{ci} 155

IV.4.4.3 Effect on g_d 158

IV.4.4.4 Impact on device performance 158

IV.5 Optimisation of n_s 160

IV.5.1 Fore Plan Doping approach 160

IV.5.2 Rear Plan approach 162

IV.5.2.1 Recapitulation 166

IV.6 Summary 169

V Simulation of HEMT's in breakdown **173**

V.1 Introduction 174

V.2 Breakdown in quasi planar HEMT 176

V.2.1 Analysis of breakdown 180

V.3 Study of the effect of recess offset 186

V.3.1 Motivation of the study 186

V.3.2 Breakdown in recessed gate HEMT's 186

V.4 Conclusion 193

VI Summary and Conclusions **197**

VI.1 Motivation of this work 198

VI.2 The Simulator 198

VI.2.1 The physical model 198

VI.2.2 Mathematical and computational models 199

VI.3 Main results obtained 199

VI.4 Limitations and extensions 200

VI.5 Closing remark 201

A Vectorisation and Multitasking **207**

A.1 Vector processing 207

A.2 Multitasking (parallel processing) 212

List of Figures

I.1 GaAs MESFET; planar and recessed structures. The recessed gate structures were imposed by practical necessity. 5

I.2 Conduction band diagram of a heterojunction formed between wide band gap, heavily doped AlGaAs and non intentionally doped small band gap GaAs at room temperature. 6

I.3 A planar AlGaAs/GaAs MODFET 7

I.4 A typical Single quantum well AlGaAs/GaAs/AlGaAs planar MODFET 9

I.5 Maximum output power for different field effect transistors versus operation frequencies, ref.[2] 12

I.6 Output power density with operation frequency, a general drop in the output power density for operation frequency higher than 60 GHz 13

I.7 Lattice constant, energy bandgap and wavelength for different semiconducting materials 14

I.8 Strained layer hetero epitaxi. Lattice mismatch could be accommodated as an elastic strain so that the tension upon the buffer layer is minimal 15

I.9 A typical pseudomorphic InGaAs HEMT 16

I.10 Schematic presentation of the division of a conduction band into a sequence of subbands 19

I.11 Low field mobility versus 2DEG concentration for different HEMT's (courtesy of D.Pons, Thomson CSF) 21

II.1 Approximations used for the simulation of quantum wells: Triangular well approximation and the square well approximation 44

II.2 Effect of energy quantification on the drain current in a pulse doped pseudomorphic AlGaAs/InGaAs device with $3 * 10^{12} cm{-}2$ doping level and $0.3\mu m$ gate 46

II.3 The variation of n_s with the low field mobility, the fine dotted line gives the fitting used to simulate the phenomena in a first approach. 47

II.4 Effect of 2D electron gas screening on f_c in a pulse doped pseudomorphic AlGaAs/InGaAs device with $3 * 10^{12} cm{-}2$ doping level and $0.3\mu m$ gate 49

II.5 Variation of the ionisation coefficient with energy 52

II.6 The non-uniform mesh nomenclature adopted in SIMFET II. Scalar quantities are defined on the main node points (O), vectorial quantities are defined on the secondary points (X) 58

II.7 A schematic presentation of an actual device, the considered portion for the simulation is the one bounded by the dashed line. The rest of the electrodes is considered as a parasitic resistance, negligible current is assumed to be passing through the imaginary (dashed) boundaries 60

II.8 SIMFET II data flow chart (version 3.0) 66

III.1 A pseudomorphic AlGaAs/InGaAs/GaAs uniformly doped HEMT; there exist a plenty of adjustable paramteres 73

III.2 The critical InGaAs layer thickness function of Indium content in a pseudomorphic AlGaAs/InGaAs/GaAs pseudomorphic HEMT. . . . 75

III.3 Distribution of physical quantities: Electrons charge density, Average total energy, and electrostatic potential in a $0.3\mu m$ gate pseudomorphic AlGaAs/InGaAs HEMT with uniform doping $1*10^{18}cm^{-3}$ over $360\AA$. 78

III.4 Distribution of physical quantities: Electrons charge density, Average total energy, and electrostatic potential in a $0.3\mu m$ gate AlGaAs/GaAs conventional HEMT with uniform doping $1*10^{18}cm^{-3}$ over $360\AA$ 79

III.5 Distribution of physical quantities: Electrons charge density, Average total energy, and electrostatic potential in a $0.3\mu m$ gate pseudomorphic AlGaAs/InGaAs HEMT with uniform doping $1*10^{18}cm^{-3}$ over $360\AA$ near pinch off 80

III.6 Distribution of physical quantities: Electrons charge density, Average total energy, and electrostatic potential in a $0.3\mu m$ gate pseudomorphic AlGaAs/InGaAs HEMT with uniform doping $1*10^{18}cm^{-3}$ over $360\AA$ in open channel operation 81

III.7 Distribution of physical quantities: Electrons charge density, Average total energy, and electrostatic potential in a $0.3\mu m$ gate pseudomorphic AlGaAs/InGaAs HEMT with uniform doping $1*10^{18}cm^{-3}$ over $360\AA$. V_{gs}=0.3 V and V_{ds}=2V 82

III.8 Distribution of the charge density along the channel in the middle of the well ($470\AA$ from the gate) in a pseudomorphic AlGaAs/InGaAs $0.3\mu m$ gate HEMT uniformly doped at $1*10^{18}cm^{-2}$ at V_{gs}=0.3 V, and V_{ds}=2.5 V 83

III.9 Comparison of the intrinsic transconductance of conventional lattice matched and pseudomorphic uniformly doped HEMT, the devices had similar gate length of $0.3\mu m$ and a doping level of $10^{18}cm^{-2}$ . . 86

III.10 Comparison of the gate source capacitance of conventional lattice matched and pseudomorphic uniformly doped HEMT, the devices had similar gate length of $0.3\mu m$ and a doping level of $10^{18}cm^{-2}$ . . 88

III.11 Comparison of the intrinsic current gain cut off frequency of conventional lattice matched and pseudomorphic uniformly doped HEMT . 90

III.12 Comparison of the output conductance in two identical pseudomorphic AlGaAs/InGaAs and conventional AlGaAs/GaAs HEMT's. Gate length is $0.3\mu m$, doping level is $1 * 10^{18} cm^{-3}$, drain potential is 2.5V 91

III.13 Distribution of physical quantities in a pulse doped AlGaAs/InGaAs pseudomorphic HEMT 94

III.14 Output currents of $0.3\mu m$ gate devices; comparison between uniform and pulse doping at V_{ds}=2.5 V 95

III.15 Comparison of the gate to source capacitance of pulse doped and uniformly doped pseudomorphic AlGaAs/InGaAs HEMT with $0.3\mu m$ gate. The drain bias is 2.5V 96

III.16 Comparison of the intrinsic transconductance of a pulse doped and uniformly doped pseudomorphic AlGaAs/InGaAs HEMT with $0.3\mu m$ gate, drain bias is 2.5V 98

III.17 Comparison of the intrinsic current gain cut off frequency of conventional lattice matched and pseudomorphic uniformly doped pseudomorphic AlGaAs/InGaAs HEMT with $0.3\mu m$ gate, drain bias is 2.5V 99

III.18 Variation of the current gain cut off frequency with steady state drain current for a pulse doped and a uniformly doped pseudomorphic AlGaAs/InGaAs HEMT with a $0.3\mu m$ gate and at a drain bias 2.5V 100

III.19 Distribution of physical quantities in a pulse doped AlGaAs/InGaAs pseudomorphic HEMT with a n.i.d AlGaAs buffer layer 103

III.20 Distribution of physical quantities in a pulse doped AlGaAs/InGaAs pseudomorphic HEMT with a n.i.d. GaAs buffer layer 104

III.21 Intrinsic transconductance versus gate voltage: Comparison of buffer type in a pseudomorphic AlGaAs/InGaAs HEMT with $0.3\mu m$ gate length and planar doped at $3 * 10^{12} cm^{-2}$. 105

III.22 Intrinsic current gain cut off frequency versus gate voltage: Comparison of buffer type in a $0.3\mu m$ gate pseudomorphic AlGaAs/InGaAs HEMT pulse doped at $3 * 10^{12} cm^{-2}$ 106

III.23 Output drain conductance versus gate voltage: Comparison of buffer type in a pseudomorphic AlGaAs/InGaAs HEMT with $0.3\mu m$ gate length and planar doped at $3 * 10^{12} cm^{-2}$. 107

III.24 Feedback capacitance versus gate voltage: Comparison of buffer type in a pseudomorphic AlGaAs/InGaAs HEMT with $0.3\mu m$ gate length and planar doped at $3 * 10^{12} cm^{-2}$. 108

III.25 Terminal currents in a $0.3\mu m$ gate pulse doped pseudomorphic HEMT with $3 * 10^{12} cm^{-2}$ doping level and 140Å InGaAs channel with 20% In. The gate bias is 0.5 V, and the drain is 3 V. 112

III.26 Travelling Gunn Domain in the gate drain region of a pulse doped HEMT, V_{gs}=0.5V , V_{ds}=3 V, time instants at 20,21 and 22 psec . . 113

III.27 Travelling Gunn Domain in the gate drain region of a pulse doped HEMT, V_{gs}=0.5V , V_{ds}=3 V, time instants 23,24 and 25 psec: the end of the cycle 114

III.28 Associated equipotential contours of the high energy-gigh field domain in a pulse doped 0.3μm HEMT, V_{gs}=0.5V , V_{ds}=3 V at the first three time instants 20, 21 and 22 psec. 115

III.29 Associated equipotential contours of the high energy-high field domain in a pulse doped 0.3μm HEMT, V_{gs}=0.5V , V_{ds}=3 V at time instants 23,24 and 25 psec. 116

III.30 Currents variation in a $0.3\mu m$ gate pulse doped pseudomorphic HEMT with $3 * 10^{12} cm^{-2}$ doping level and 140Å InGaAs channel with 20% In. The drain is 3 V, Gate bias=0.3 V. 117

III.31 Currents variation in a $0.3\mu m$ gate pulse doped pseudomorphic HEMT with $3 * 10^{12} cm^{-2}$ doping level and $140\AA$ InGaAs channel with 20% In. The drain is 3 V, Gate bias=0.45 V. 118

III.32 The longitudinal electric field in the channel, variation in the form outside the domain when V_{gs} is increased in a 0.3 μm gate pulse doped pseudomorphic HEMT with $3 * 10^{12} cm^{-2}$ doping density. . . 119

III.33 Charge control of different HEMT structures, the simulated devices had $0.3\mu m$ gate. a) is a $360\AA$ barrier layer, uniformly doped ($1 * 10^{18} cm^{-3}$ conventional HEMT, b)the same as (a) but with a $600\AA$ barrier layer, c)a uniformly doped ($1 * 10^{18} cm^{-3}$) pseudomorphic AlGaAs/InGaAs HEMT, d)a pulse doped conventional HEMT ($3 * 10^{12} cm^{-2}$) and e)a pulse doped (same as d) pseudomorphic AlGaAs/InGaAs HEMT with $330\AA$ barrier layer 121

III.34 Charge control of the devices described in the previous figure, drain bias is zero. Gate bias giving n_s more than $1.2 * 10^{12} cm^{-2}$ is more likely to give oscillating drain current when the drain is biased . . . 122

III.35 Charge control of two pulse doped ($3 * 10^{12} cm^{-2}$) pseudomorphic AlGaAs/InGaAs HEMT with 330 and $100\AA$ undoped barrier layer thickness. 124

III.36 Output drain and source currents in the pulse doped 0.3μm device, the transistors is coupled to a 3Ω series resistance and 1pH series inductance, the drain bias is 3.5 V 125

IV.1 The basic structure of the pseudomorphic AlGaAs/InGaAs/GaAs pulse doped HEMT 134

IV.2 Simulation of recessed structures, a more realistic approach is applied136

IV.3 Variation of transconductance with gate bias for different thickness of barrier layer. Gate length is $0.3\mu m$, doping level is $3 * 10^{12} vm^{-2}$ ($V_{ds} = 2.5V$) 138

IV.4 Variation of maximum Transconductance with the inverse of a in a $0.3\mu m$ gate pulse doped pseudomorphic HEMT with doping level $3*10^{12}cm^{-2}$ ($V_{ds}=2.5V$) 139

IV.5 Variation of gate-source capacitance with gate bias for different thickness of barrier layer in a pseudomorphic AlGaAs/InGaAs HEMT with a $0.3\mu m$ gate and doping level $3*10^{12}cm^{-2}$. ($V_{ds}=2.5V$) . . . 140

IV.6 Variation of intrinsic current gain cut off frequency with gate bias for different thickness of barrier layer in a pseudomorphic AlGaAs/InGaAs HEMT with $0.3\mu m$ gate and doping level $3*10^{12}cm^{-2}$. ($V_{ds}=2.5V$) 142

IV.7 Effect of gate length reduction (scaling) upon intrinsic transconductance in a pseudomorphic pulse doped HEMT, the undoped barrier layer is kept at $100\AA$ and the doping level is $3*10^{12}cm^{-2}$ 144

IV.8 Effect of gate length reduction (scaling) upon the intrinsic gate source capacitance in a pseudomorphic pulse doped HEMT, the undoped barrier layer is kept at $100\AA$ and the doping level is $3*10^{12}cm^{-2}$ 145

IV.9 The feedback capacitance variation with V_{gs} corresponding to 0.1 and $0.3\mu m$ gate lengths, the devices are planar doped with $100\AA$ undoped barrier layer thickness. 146

IV.10 The feedback capacitance variation with V_{gs}-Voff corresponding to 0.1 and $0.3\mu m$ gate lengths, the devices are planar doped with $100\AA$ undoped barrier layer thickness. 147

IV.11 Effect of gate length reduction (scaling) upon the intrinsic current gain cut off frequency in a pseudomorphic pulse doped HEMT, the undoped barrier layer is kept at $100\AA$ and the doping level is $3*10^{12}cm^{-2}$ 148

IV.12 Variation of the maximum current gain cut off frequency with the inverse of the gate length in a pseudomorphic pulse doped HEMT, the undoped barrier layer is kept at $100\mathring{A}$ and the doping level is $3 * 10^{12} cm^{-2}$ 150

IV.13 The intrinsic drain conductance of $0.3\mu m$ and $0.1\mu m$ gate devices with planar doping of $3.10^{12} cm^{-2}$ and the same gate channel separation 151

IV.14 Plot of the distribution of electrons charge density, average total electron energy and electrostatic potential for a $0.3\mu m$ gate pulse doped pseudomorphic HEMT, the fore plan doping is $3.10^{12} cm^{-2}$. 152

IV.15 Plot of the distribution of electrons charge density, average total electron energy and electrostatic potential for a $0.1\mu m$ gate pulse doped pseudomorphic HEMT, the fore plan doping is $3.10^{12} cm^{-2}$. 153

IV.16 An HEMT with a recessed gate structure. 156

IV.17 Variation of the maximum current gain cut off frequency with recess offset in a voltage swing (2-10 V) in a pseudomorphic pulse doped HEMT with $0.3\mu m$ gate, the undoped barrier layer is kept at $150\mathring{A}$ and the doping level is $3 * 10^{12} cm^{-2}$ 157

IV.18 Variation of the output conductance with gate bias for different recess offsets in a planar doped HEMT of $0.3\mu m$ gate and doping level of $3 * 10^{12} cm^{-2}$, the barrier layer thickness is $150\mathring{A}$ and the drain voltage is 2.5V 159

IV.19 Variation of the drain current in a $0.3\mu m$ gate pulse doped HEMT with two different doping levels. The drain bias is 2.5 V 161

IV.20 Variation of f_c in a $0.3\mu m$ gate pulse doped HEMT with two different doping levels, the barrier layer is kept at $100\mathring{A}$. The drain bias is 2.5 V 162

IV.21 A planar doped HEMT with a rear plan doping; gate length $0.3\mu m$, barrier layer thickness $100\mathring{A}$ 163

IV.22 Variation of the current gain cut off frequency with the drain steady state current for pulse doped devices: effect of buffer doping while having a fore plan doping fixed at $3 * 10^{12} cm^{-2}$ and a $0.3 \mu m$ gate . 164

IV.23 Variation of the current gain cut off frequency with the drain steady state current: effect of buffer doping level (replot of the previous figure) 165

IV.24 Plot of the distribution of electrons charge density, average total electron energy and electrostatic potential for a buffer doped pseudomorphic HEMT, the fore plan doping is $3.10^{12} cm^{-2}$ and the rear plan doping is $1.5.10^{12} cm^{-2}$ 167

IV.25 Plot of the distribution of electrons charge density, average total electron energy and electrostatic potential for a buffer doped pseudomorphic HEMT, the fore plan doping is $3.10^{12} cm^{-2}$ and the rear plan doping is $2.5.10^{12} cm^{-2}$ 168

V.1 Schematic presentation of I-V characteristics in a FET. The breakdown at pinch-off is the limiting value on V_{ds}. 174

V.2 The conventional HEMT considered for the study of breakdown characteristics. The gate length is $0.3 \mu m$, the doping level is $1 * 10^{18} cm^{-3}$ uniformly over a layer $360 \mathring{A}$ thick. The supply layer is AlGaAs with 30% Al, and the channel is n.i.d. GaAs 176

V.3 Drain current versus Drain potential characteristics of a quasi planar lattice matched HEMT with uniform doping $10^{18} cm^{-2}$ at different gate potentials 178

V.4 Experimental drain current versus Drain potential characteristics of a quasi planar lattice matched HEMT with uniform doping $10^{18} cm^{-2}$ at different gate potentials, gate length $0.3 \mu m$, 28%Al, barrier layer $380 \mathring{A}$ 179

V.5 Distribution of physical quantities inside the HEMT upon breakdown where the gate potential is near pinchoff. 181

V.6 Distribution of physical quantities upon breakdown with an open channel in a conventional AlGaAs/GaAs HEMT, uniformly doped at $3 * 10^{12} cm^{-3}$. The gate length is $0.3\mu m$, V_{gs}=0.3 V, V_{ds}=5.0V . . 182

V.7 The high energy domain variation with drain bias at an open channel operation 183

V.8 The high energy domain variation at the same drain bias and for the three operation regimes of the gate bias. 184

V.9 Position of the high energy domain at breakdown for different gate bias 185

V.10 Schematic presentation of the three devices considered for the study of the effect of recess offset. The first device is with a zero recess offset, the second device is with $0.1\mu m$ offset and the third device is with $0.2\mu m$ recess offset distance. The gate length is $0.3\mu m$ and the doping level is $1 * 10^{18} cm^{-3}$ uniform over $360\AA$ 187

V.11 I-V characteristics of a $0.3\mu m$ conventional uniformly doped at $1 * 10^{18} cm^{-3}$ AlGaAs/GaAs HEMT with different recess offset distance. The top curves are for a device with a zero recess offset, the middle curve is for a device with $0.1\mu m$ offset and the bottom curve is for the device with $0.2\mu m$ recess offset distance 188

V.12 Influence of recess offset upon breakdown potential for different gate bias 189

V.13 Variation in the shape of the high energy domain in the quantum well for different recess offsets at breakdown in an open channel . . 190

V.14 High energy domains in the quantum well for different recess offsets at the same drain bias (5V) and the same gate bias (3V) 191

V.15 Variation in the shape of the high energy domain in the quantum well for different recess offsets at breakdown for a near pinch off operation 192

A.1 Nomenclature of a classical finite difference mesh 209

A.2 Nomenclature of a finite difference mesh where the mesh points are arranged in a vector of length m*n rather than an array 210

This page is intentionally left blank

List of Tables

II.1 Material parameters for calculating ionisation coefficient in GaAs and AlGaAs with 30% Al 51

II.2 Typical performance figures of the simulator for different devices. LN=Low noise conditions, Vds $<$ 3 V. UD=Uniform Doping. PD=Pulse Doping. BD=Breakdown conditions Vds $>$ 5 V. 67

IV.1 Comparison of maximum obtained f_{ci} with published experimental results of similar devices, all the devices are pulse doped pseudomorphic AlGaAs/InGaAs/GaAs HEMT's 149

IV.2 The performance of different devices with different recess offsets, the three devices are pseudomorphic AlGaAs/InGaAs pulse doped at $3*10^{12}cm^{-2}$,Rs=Rd=0.2$\Omega.mm$, Ri=1.3$\Omega.mm$, Ls=2pH.mm. (†parameters calculated at the maximum current gain cut off frequency 160

IV.3 Characteristics of the different devices with different planar doping levels. The gate length is $0.3\mu m$, the undoped barrier layer is $100\mathring{A}$, the current and the current gain cut off frequency were obtained for V_{ds}=2.5 V 169

This page is intentionally left blank

Chapitre 1

Les HEMTs de puissance

1.1. Introduction

Depuis leur invention, les transistors à effet de champ n'ont cessé d'inspirer les ingénieurs pour rechercher des modes de fonctionnement à des fréquences plus élevées et à des puissances plus importantes. Les recherches ont été basées sur la conception et la mise en oeuvre de nouvelles structures epitaxiales et de différentes technologies. C'est dans un tel contexte qu'après le transistor MESFET, plusieurs variétés de HEMT sur substrat au GaAs ou à l'InP ont été proposées. Pour comprendre la motivation des différentes approches il est important de bien préciser ce qui est demandé à un transistor à effet de champ pour un fonctionnement en puissance en gamme millimétrique.

1.2. Spécifications D'un HEMT de puissance

Plusieurs spécifications sont en effet requises à la fois :-

- Une puissance de sortie maximale et pour cela il faut augmenter les deux éléments intervenant dans la formule approchée de la puissance; la tension de claquage V_B et le courant maximal Im.

$$P_s \approx 1/8 V_B I_m \tag{1.1}$$

- Un gain maximal, c'est à dire une fréquence de coupure du gain qui soit la plus élevée possible. Cela exige une augmentation de la fréquence de coupure du gain en courant et donc une diminution des éléments capacitifs (C_{gs} et C_{gd}).
- Une linéarité la meilleure possible, ce qui impose une fréquence de coupure du gain la plus indépendante possible de la polarisation.

1.3. Composant de Choix

Dans le domaine des frequences millimetrique, parmi les composants possibles le HEMT GaAlAs/GaInAs à couche pseudomorphique sur GaAs semble actullement le plus prometteur et c'est cette filiére que nous nous sommes proposeès d'étudier et d'optimiser.

Dans ce contexte, une étude bibliographique attentive permet de trouver quelques solutions, mais indique aussi qu'il y a des compromis à trouver :

- l'augmentation des tensions de claquage passe par l'utilisation de couches peu dopées, de matériaux à large gap, et de technologies types "recess" où les paramètres géométriques sont extrêmement critiques.
- l'augmentation des courants passe au contraire par l'utilisation de couches epitaxiales très dopées, de matériaux à mobilité élevée donc souvent à faible bande interdite et à des heterostructures à barrière d'énergie très élevée.
- l'amélioration des fréquences de coupure passe par l'utilisation de grilles ultra courtes et d'épaisseurs de couches faibles

Il est évident qu'un nombre important de paramètres est à considérer et que l'optimisation est baseè sur la prise en considération d'influences contradictoires. Il est donc impossible qu'une telle approche soit menée de façon exclusivement expérimentale. Dans ces conditions, l'utilisation de moyens de simulation physiques précis et bien adaptés peut constituer une aide précieuse pour mettre en oeuvre une telle optimisation.

Chapitre 2

Moyens de Simulation

2.1. Choix de Modèle

Parmi les différentes méthodes de simulation, la classe de modèles de type Monte Carlo basés sur un traitement microscopique offre la meilleure précision possible: ils tiennent en effet compte de la plupart des effets physiques qui interviennent dans le parcours des porteurs de la source vers le drain. Mais, ils restent très complexes et nécessitent des temps et des moyens de calcul très importants.

A l'autre extrémité, on trouve des modèles analytiques. Ils sont très rapides, et peuvent être implantés sur des petits ordinateurs et intégrés dans les simulateurs de circuit. Mais ils sont basés le plus souvent sur des simplifications très exagérées et ils sont incapables de prédire l'évolution des paramètres et performances en fonction des paramètres technologiques.

Entre ces deux extrêmes, on trouve des modèles basés sur la resolution du système d'équations des semi-conducteurs. Les modèles de transports quantiques sont le plus souvent trop complexes pour être utilisés dans des structures réelles et ne sont applicables que à des éléments quasi unidimensionnels tels que les diodes à effet tunnel résonnant. Les modèles de transport quasi bidimensionnels constituent une solution, mais ils ne peuvent pas permettre d'évaluer, dans tous les cas, toutes les caractéristiques (conductance de drain, tension de claquage notamment) et ne sont pas utilisables pour les structures trop complexes.

Un bon compromis entre les exigences contradictoires semble réside dans le modèles dits hydrodynamiques bidimensionnels. Ils sont basés sur la resolution bidimensionnelle des équations de transport dérivés de l'equation de Boltzmann. Ils présentent une précision satisfaisante ainsi qu'un temps de calcul et une complexité acceptables. Dans le cas des HEMT, ce dernier type de simulation semble être le plus intéressant pour guider la conception de composants optimisés en vue d'une utilisation en régime de forte puissance en ondes millimétriques.

2.2. Le modèle physique

Le modèle que nous avons choisi pour mener à bien notre étude des composants HEMT de puissance pour ondes millimétriques (filière pseudomorphique GaAlAs/GaInAs)

avait été conçu pour la simulation physique de MESFETs et de HEMTs conventionnels GaAlAs/GaAs. Bien que constituant un très bon outil pour ces derniers types de composants, il incluait des hypothèses simplificatrices peu acceptables pour la simulation des composants à couches pseudomorphiques. Ces principales hypothèses sont :

- les effets liés à l'existence d'un gaz bidimensionnel d'électrons sont négligés: les électrons sont toujours considérés comme tridimensionnels. Cette hypothèse s'appuyait sur des simulations Monte Carlo qui montraient que dés que Vds était supérieur à 0.5 V les électrons pourraient être considérés comme tridimensionnels durant la plus grande partie du parcours.
- la mobilité des électrons est égale à celle obtenue en volume et indépendante de la concentration des porteurs.
- les phénomènes d'ionisation ne sont pas pris en compte.

Il nous a semblé important d'introduire ces effets dans le modèle, afin de mieux rendre compte de la réalité physique. Au préalable, nous allons en rappeler les éléments essentiels.

2.3. Nouveaux effets physiques introduits

2.3.1. Caractère bidimensionnel des électrons

Dans les structures à heterojonction telles que les HEMT, les porteurs confinés dans un puits quantique perdent un degré de liberté dans le sens perpendiculaire à l'heterojonction. Il en résulte une conséquence éssentielle: à la bande de conduction classique comprenant plusieurs vallées, est substituée une structure de sous bandes construites à partir de niveaux discrets, à l'intérieur des quelles la densité d'état est constante. Quand l'énergie moyenne des électrons est inférieure à la hauteur de la barrière, le caractère bidimensionnel des électrons doit être pris en compte. C'est beaucoup plus souvent le cas pour les composants pseudomorphiques où la hauteur est de 0,31 ev alors qu'elle n'est que 0,23 ev pour des composants conventionnels.

L'effet essentiel est que le niveau de référence de l'énergie des électrons n'est plus le bas de la bande du conduction, mais au minimum celui de la première sous bande ε_0; le transfèrt dans les vallées satellites s'en trouve facilite et la dynamique est donc différente.

2.3.2. Effet d'écran

Cet effet a été observé expérimentalement dans plusieurs types de transistors HEMT. Il se manifeste par une dépendance de la mobilité en fonction de la densité de charge dans le puits. La mobilité peut augmenter d'un facteur 3 quand n_s passe de $10^{11} cm^{-2}$ à $10^{13} cm^{-2}$. Des explications qualitatives attribuent cette amélioration à un effet d'écran exercé par le gaz 2D vers des impuretés de donneurs présentes dans l'AlGaAs. Il faut signaler que jusqu'aujourd'hui, il n'y a pas une interpretation quantitative qui permet de rendre compte de ce phénomène.

2.3.3. L'ionisation par choc

L'ionisation par choc est un phénomène qui se produit en présence d'un champ électrique très élevé dans une structure semiconductrice (ordre de quelques centaines de KV/cm). Jointe au phénomène de multiplication par avalanche, il est responsable de l'augmentation brutale du courant drain dans un transistor à effet de champs. Il peut entraîner une perte de fiabilité voire une destruction complète du composant. Puisque dans les applications de puissance, une tension importante constitue une condition normale de fonctionnement, l'apparition de l'ionisation par choc va introduire une limite sur les tensions utilisables et par suite sur la puissance de sortie.

Le phénomène de base est bien connu. Quand un électron est placé dans un champ électrique, sa vitesse et son énergie croissent jusqu'à ce qu'il subisse un choc (avec le réseau, une impureté,...). La fréquence des chocs augmente avec la vitesse des porteurs. Si le champ électrique est plus élevé, l'énergie cinétique des électrons peut être suffisante pour qu'un choc entraîne la création d'une paire électron trous, c'est l'ionisation par choc. Les deux porteurs créès peuvent à leur tour être accélérés et donner lieu à une nouvelle ionisation par choc. C'est ce qu'on appelle la multiplication par avalanche.

La conservation de l'énergie impose que le porteur initial ait une énergie cinétique supérieure à la largeur de la bande interdite. En fait, dans l'approximation d'une masse effective égale pour les deux types des porteurs, l'énergie de seuil pour l'ionisation est égale à 1,5 fois la largeur de la bande interdite.

Il est clair que l'effet l'ionisation par choc est un phénomène où le paramètre physique fondamental est l'énergie des électrons qui doit être supérieure à un certain seuil pour que le phénomène existe.

2.4. Simulation des phénomènes pris en compte

2.4.1. Introduction des effets quantiques

L'objectif est de calculer l'énergie de la première sous bande ε_0. Une simulation exacte nécessite la resolution auto consistante de l'equation de Schrodinger couplée à l'equation de Poisson: celle-ci devrait être faite à chaque point du puits quantique, ce qui serait extrêmement lourd. Il a semble préférable d'utiliser des approximations. Il est possible de calculer analytiquement le niveau ε_0 dans deux cas limites celui de puits triangulaire infini et de puits rectangulaire.

Pour les structures conventionnelles (AlGaAs/GaAs), nous avons considéré un puits triangulaire infini où le niveau ε_0 est calculé par:

$$\varepsilon_l = \left(\frac{\hbar^2}{2m}\right)^{1/3} \left(\frac{3\pi}{2} q E_s\right)^{2/3} (l + 3/4)^{2/3} \, l = 0, 1, 2, 3, ... \qquad (2.1)$$

ou Es est le champ électrique à l'interface.

Pour les composants HEMT de structure pseudomorphique, nous avons considère un puits rectangulaire où ε_0 est donne par

$$\varepsilon_l = \frac{\hbar^2 \pi^2}{8 m^* L_z^2} (l + 1)^2 \, l = 0, 1, 2, 3, ... \qquad (2.2)$$

où Lz est la largeur de puits.

ε_0 est alors considérée comme le niveau de référence de l'énergie potentielle des électrons dans le puits. Elle est ajoutée au dernier terme de l'équation de l'énergie.

Dans le cadre d'un gaz équivalent d'électrons, ceci va modifier la dynamique de transport.

Afin d'illustrer l'influence de l'introduction de cet effet, nous avons tracé l'évolution du courant de drain en fonction de Vgs dans un HEMT pseudomorphique en prennent en compte cet effet (figure II.2). Le courant diminue d'environ 10% quand on introduit ce phénomène.

2.4.2. Simulation de l'effet d'écran

Comme nous l'avons vu précédemment, il n'existe pas à notre connaissance un traitement quantitatif satisfaisant de cet effet. Pour l'introduire, nous nous sommes fondés sur les observations expérimentales effectuées à Thomson CSF: sur la figure(II.3) sont montrées les évolutions de la mobilité des porteurs en fonction de la densité de charge dans le puits pour différents types de transistors HEMTs. Pour en tenir compte, nous avons suppose que pour l'évolution générale de la dynamique électronique $\mu(\xi)$, la mobilité à faible champ est modifiée suivant une loi de type

$$\mu_{fc} = An_s + B \tag{2.3}$$

ou A et B sont: $1.6 * 10^{-9}$ et 2700

et n_s est donnè par

$$n_s(x) = \int_L ndl \tag{2.4}$$

Cet effet va contribuer à défavoriser la dynamique des électrons pour des tensions faibles pour lesquelles la densité superficielle de charge est très inférieure à la valeur maximale.

Afin de montrer l'influence de ces phénomènes, nous avons comparé pour un composant typique (HEMT conventionnel de longueur de grille de 0,3μ) l'évolution de la fréquence de coupure du gain en courant en introduisant ou non cet effet. La différence est très significative: en particulier en conditions faible bruit, la diminution de la fréquence de coupure atteint 25%. Elle est moins importante dans les zones de fréquence de coupure à courant maximum (figure II.4).

Nous avons pu comparer les résultats obtenus avec des donnees expérimentales disponibles dans la littérature ou obtenues sur des composants réalises par Thomson CSF. Les résultats sont présentés dans le tableau(IV.1). La bonne conformité montre la validité de notre approche.

2.5. Simulation de la multiplication par avalanche

Elle comporte principalement le calcul des taux de génération G qui est à introduire dans l'equation de continuité des électrons et des trous.

$$G = \alpha_n n v_n + \alpha_p P v_p \tag{2.5}$$

ou α_n et α_p sont respectivement les taux d'ionisation des électrons et des trous, et v_n, v_p leurs vitesses respectivement.

Signalons que la première partie de la formule $\alpha_n n v_n$ (taux de génération des électrons) constitue dans tous les cas le terme principal et que c'est celui que nous devons calculer avec le plus de précision.

2.5.1. Taux de génération dû aux électrons

Les densités des porteurs et les vitesses des électrons sont connues à partir des autres équations, il s'agit uniquement de calculer le taux d'ionisation. Afin de tenir compte des caractères non instantané de l'ionisation (existence de l'énergie de seuil, notion d'espace noir) nous admettons que *le taux d'ionisation des électrons est uniquement fonction de l'énergie moyenne totale (cinétique + potentielle)* des porteurs. Pour évaluer cette dépendance, nous faisons une hypothèse simplificatrice: nous admettons que la dépendance est la même qu'en régime stationnaire. On utilise d'une part des relations classiques entre les taux d'ionisation des électrons et le champ électrique;

$$\alpha = Aexp\left(\frac{B}{E}\right)^m \tag{2.6}$$

et d'autre part la dépendance en régime stationnaire entre l'énergie totale moyenne et le champ électrique.

Il faut signaler que pour des champs électriques très élevés (supérieurs à 500 KV/cm), l'évolution de l'énergie en fonction de champ électrique est très peu connue et peu de résultats ont été publies. Les résultats le plus convainquant ont été obtenus par Lippens et Vanbesien dans notre laboratoire; ceux que nous avons utilisés.

La forme finale de la dépendance des taux de l'ionisation en fonction de l'énergie est donnée sur la figure(II.5).

2.5.2. Taux de génération dû aux trous

Pour évaluer ce taux, il faut donc connaître la concentration et la vitesse des porteurs minoritaires. Il faut donc résoudre l'equation de continuité des trous et l'equation de courant de trous. Comme ces termes sont moins importants, ils sont évalués de même que α_p en prenant une approximation locale. On considère que la mobilité équivalente et le taux d'ionisation sont uniquement fonctions du champ électrique; l'évolution de la vitesse des trous en fonction de champ électrique est supposée similaire à celle des électrons dans le Silicium (mobilité de 300 cm/V.sec).

2.6. Quelques éléments particuliers du modèle

Afin d'améliorer la qualité du modèle nous avons introduit un certain nombre d'améliorations qui permettent de mieux rendre compte des composants réels:

- La simulation des structures creusées, "ou recess", est effectuée en considérant une structure équivalente planaire; dans ce cas la zone non creusée est représentée par une partie très dopée ($N_d = 2 * 10^{18} cm^{-3}$) où le potentiel de surface est sans effet.
- Dans les structures à plan de dopage la diminution de mobilité liée aux interactions avec les impuretés ionisées est prise en compte non seulement au niveau du plan mais à son voisinage pour introduire l'influence des interactions coulombiennes à large distance et la diffusion des dopants.

- Les mêmes effets sont introduits au niveau du spacer.

2.7. Rèsumé du modèle physique final

1. Simulation en régime dynamique transitoire.
2. Modèle bidimensionnel prenant en compte les effets d'injection et de transfèrt dans l'espace réel.
3. Effets physiques introduits:
   (a) Localisation des électrons dans le puits.
   (b) Effets d'écran.
   (c) Ionisation par choc.
   (d) Existence des deux types des porteurs.
   (e) Piégeages sur centre DX.
   (f) Effets du potentiel de surface.

2.8. Traitement numérique et algorithmique

Les équations sont discrètisées dans un maillage bidimensionnel aux différence finies à pas variable. Les équations forment un système non linéaire et peuvent être écrites sous une forme matricielle.

Les conditions aux limites choisies sont classiques. Les conditions de Dirichlet sont appliquées sur les contacts ohmiques. Sur les autres interfaces (à l'exception de la grille), les conditions de Neumann sont appliquées. Certaines de ces interfaces ne sont pas réelles, (interfaces imaginaires), elles sont choisies avec soin en s'arrangent pour que le courant qui les traverse soit nul ou extrêmement faible. En ce qui concerne le contact Schottky de grille, le problème est pour l'instant très complexe: la concentration de charge est une fonction du courant grille. En pratique, dans la plupart des cas, l'utilisation restant limitée aux régimes où le courant grille est nul, nous avons donc impose cette hypothèse pour les électrons. La densité de charge en est automatiquement déduite.

La discrètisation temporelle est effectuée en ayant un pas de temps variable et adaptable au fur et à mesure.

Les équations sont résolues comme suit:

- L'equation de Poisson est résolue en utilisant une méthode directe de double balayage matriciel qui constitue une variante de la méthode de Choleski. L'avantage est qu'elle nécessite un temps de calcul plus faible que les autres méthodes. L'inconvennient majeur réside dans le fait qu'elle ne permet pas des modifications de maillage en cours de calcul.
- Les équations de conservation sont d'abord linéarisées en utilisant une méthode de Newton, ensuite elles sont résolues au moyen d'une méthode itérative de sur relaxation. Les facteurs d'accélération sont 1,2 et 1 respectivement pour les équations de continuité et de l'énergie.
- Les équations sont mises sous forme vectorielle, en éliminant tout interdépendance de données, afin de profiter pleinement des avantages des processeurs vectoriels.

Nous avons utilise des techniques d'exécution en multitâches en profitant des valeurs différentes des temps de relaxation. Nous avons finalement développé des techniques d'accélération àfin de réduire le temps total de simulation.

Chapitre 3

Etude des transistors HEMT pseudomorphiques

3.1. Comparaison avec le transistor conventionnel

Pour bien réussir une bonne optimisation, une compréhension approfondie du fonctionnement des composants est nécessaire, ce qui est possible en utilisant notre modèle. Nous avons donc commencé à étudier le fonctionnement d'un transistor HEMT pseudomorphique à dopage uniforme en le comparant avec un transistor conventionnel de performances optimales. Dans ce cadre, nous avons repris les mêmes paramètres structurels. La longueur de grille est de 0.3μ, le dopage de la couche d'AlGaAs est $1 * 10^{18} cm^{-3}$ et son épaisseur est $360 \AA$. En ce qui concerne le choix des taux d'indium et d'aluminium, nous nous sommes basés sur les observations pratiques et la nécessité d'avoir un maximum de concentration de charges dans le puits. Ces deux éléments nous ont guidé vers un taux d'Al de 20% et d'In de 20% qui nous donnent une discontinuité de bande de conduction de 0,31 ev. Le taux d'indium nous impose une épaisseur maximale de la couche active de $150 \AA$ selon le critère de Matthew; pour obtenir le maximum de charges dans le puits nous avons choisi une épaisseur de $140 \AA$. Une couche tampon de $500 \AA$ a étè considéré, ce qui semble suffisant compte tenu de la faible injection des porteurs dans le substrat.

Nous avons remarqué lors de la simulation de ce transistor que le n_s (la concentration totale de la charge dans le puits) passe de $1 * 10^{12} cm^{-2}$ pour le transistor conventionnel à $1,34 * 10^{12} cm^{-2}$ pour le transistor pseudomorphique. La répartition des grandeurs physiques montre un meilleur confinement des porteurs dans le puits pour les structures pseudomorphiques. Ce confinement est accompagne par la formation d'un domaine d'énergie élevée comme on le voit sur la figure(III.3). Ce domaine de haute énergie est une caractéristique des transistors à effet de champ a base de matériaux III-V. La formation de ce domaine, dans le cadre d'un gaz équivalent d'électrons, est due à la caractéristique vitesse champ électrique des porteurs qui présente une zone de mobilité différentielle négative. Les courbes équipotentielles montrent que les lignes des potentiels sont perpendiculaires à l'axe source drain dans la plus grande partie du canal. Ceci veut dire que le transfèrt dans l'espace k est privilègié par rapport au transfert dans l'espace réel.

La comparaison des caractéristiques de sortie fait apparaître des performances pour les transistors pseudomorphiques nettement meilleures par rapport aux transistors

conventionnels. Cette amélioration peut s'expliquer comme étant le résultat d'un meilleur confinement dans le puits. Ceci est considéré comme le cause principale car la dynamique de transport électronique dans les couches InGaAs contraintées est suppose ́identique à celle des couches de GaAs en volume. La transconductance augmente de 600 mS/mm dans le HEMT conventionnel à 750ms/mm dans le HEMT pseudomorphique. La capacitance grille source est légèrement plus élevée: elle passe de 0.8 pF/mm (à Vgs=0) á 0.9pF/mm pour les composants de même longueur de grille $(0, 3\mu)$ et la même tension de pincement. Mais on remarque que la capacitance varie moins en fonction de Vgs dans le transistor pseudomorphique: ceci s'explique aussi par le meilleur confinement des porteurs. La fréquence de coupure du gain du courant (fc) fait apparaître une amélioration de 16% dans les transistors pseudomorphique. La comparaison des conductances de sortie montre clairement qu'il y a une amélioration très importante due au confinement des porteurs dans le puits pour les structures pseudomorphiques.

3.2. Structures à plan de dopage

L'utilisation des plans de dopage a un double intérêt. Elle offre en effet la possibilité de placer une quantité important de dopants assez proches du puits en évitant les risques dû à la présence d'un dopage élevé au voisinage de la grille. Nous avons donc poursuivi notre étude en simulant une structure pseudomorphique AlGaAs/InGaAs où la couche barrière de AlGaAs est à plan de dopage de $3*10^{12}cm^{-2}$. L'augmentation de n_s est remarquable, elle passe de $1,34*10^{12}cm^{-2}$ à $2,1*10^12cm^{-2}$ obtenue précédemment. Cette augmentation de n_s entraîne un décalage important de la tension de pincement; pour avoir deux transistors comparables, nous avons réduit l'épaisseur de la couche barrière à $180\AA$ dont $30\AA$ pour le plan de dopage. La comparaison des caractéristiques des deux composants montre que le transistor à plan de dopage donne un courant maximum supérieur, bien sur en relation avec l'augmentation de n_s. La réduction de la distance grille-puits entraîne une augmentation de la transconductance maximale qui passe de 750 à 1200 mS/mm. La variation de la capacitance de la grille montre une dépendance plus forte en fonction de Vgs pour les structures à plan de dopage. Ceci peut être expliqué par la variation

relative de la position du gaz 2D par rapport à la grille.

La variation de la fréquence de coupure du gain de courant montre un meilleure fonctionnement pour le régime de faible bruit. Cette supériorité n'est plus valable quand on va vers un fonctionnement en canal ouvert. La raison de cette chute est liée à l'importance de la conduction parallèle dans la couche barrière (MESFET parasite), dans un composant à plan de dopage; le dynamique des électrons dans cette zone est relativement plus mauvaise à cause du dopage élevé qui est 10 fois plus grand que dans le transistor à dopage uniforme. Malgré cela, quand on représente la variation de f_c en fonction du courant, on trouve que le HEMT à plan de dopage donne une meilleur fonctionnement sur toute la gamme d'utilisation. Cela est d'extrême importance pour les composants de puissance.

3.3. Nature de la couche tampon

L'utilisation de couche tampon en AlGaAs (au lieu de GaAs) semble avoir un léger effet sur le fonctionnement de la structure. Une augmentation de n_s est remarquée ($2,3$ à $2,4 * 10^{12} cm^{-2}$); par contre la valeur maximum de la fréquence f_c ne varie pas. On peut noter une amélioration de la conductance de sortie (10%) à cause du meilleur confinement de porteurs. Un substrat AlGaAs donne un offset de bande ΔE_c de $0,31ev$ au lieu de $0,15ev$ pour un substrat GaAs. Il peut en résulter une augmentation de 6% de la valeur maximale du F_{max}.

3.4. Les instabilités temporelles dans les HEMT

Le phénomène d'oscillations parasites a été remarqué dans le passé dans plusieurs structures à effet de champ à la fois expérimentalement et à partir des simulations basées sur des modèles différents. L'apparition de ces oscillations est liées à l'existence d'un domaine de haute énergie età haut champ qui traverse périodiquement la zone entre la grille et le drain. Ce domaine est generé à cause du phénomène de mobilité différentielle négative. Les courants sur les deux électrodes (source est drain) montrent que dans notre cas, la fréquence est de l'ordre de 200GHz; cette fréquence correspond bien au temps de transit des porteurs entre grille et drain

(0,5μ), à une vitesse de déplacement de $1 * 10^7 cm/sec$. Nous avons aussi vérifié que ces oscillations sont d'origine physique est non pas numérique: la fréquence change en fonction de la distance grille drain, et les oscillations disparaissent pour des tensions grille drain faibles.

Afinde comprendre ce phénomène nous avons étudié une structure pseudomorphique à plan de dopage de $3 * 10^{12} cm^{-2}$ et une grille de $0,3\mu$. Le courant de sortie du HEMT restait stable jusqu'à une tension Vgs de 0,45 V. A Vgs=0,5V nous observons l'apparition des oscillations dans les courants de sortie. Quand on représente la composante longitudinale du champ électrique pour plusieurs valeurs de Vgs, nous remarquons que ce champ tombe toujours vers zéro à l'extérieur du domaine dans le cas où la fonctionnement est stable. Mais quand on se rapproche du régime instable, la distance pour laquelle le champ électrique retourne à zéro s'élargit pour se répartir sur toute la zone grille drain juste avant le démarrage des oscillations. Ceci veut dire qu'il faut que la forme du domaine de champ électrique dans la zone grille drain soit modifiée afin qu'un champ suffisant pour provoquer le mouvement du domaine existe à l'extérieur de celui-ci.

Par ailleurs, nous avons étudié la possibilité d'avoir des oscillations dans différentes structures HEMT. Nous avons simule des HEMT:

1. Conventionnels avec un rapport d'aspect optimal.
2. Conventionnels avec un mauvais rapport d'aspect (faible).
3. Pseudomorphiques à dopage uniforme.
4. Conventionnels à plan de dopage.
5. Pseudomorphiques à plan de dopage et un rapport d'aspect faible(couche supérieure épaisse, 330A).

Nous avons ensuite étudié le contrôle de charge dans ces composants et tenté de le coreller avec l'existence des oscillations. *Nous avons pu remarquer que pour les points de polarisation de grille qui donnent un n_s supérieur à $1 * 10^{12} cm^{-2}$, il existe une forte chance que les oscillations se déclenchent lors de la polarisation de drain.*

En ce qui concerne le HEMT à plan de dopage nous avons pu constater que la réduction de l'épaisseur de la couche supérieure non dopée peut avoir un effet positif sur le fonctionnement du transistor et sur la non-apparition des oscillations.

L'influence du circuit extérieur peut être déterminant quant à l'existence de ces oscillations parasites. Il semblait peu acceptable d'avoir des tensions Vgs et Vds fixes en présence d'oscillations de courant, ce qui tendait à dire que les impédances de charge et de source sont nulles. Nous avons tenté de rendre compte de l'influence d'une self parasite de source et d'une résistance source Rs. Pour un transistor de 100μ de grille, on obtenait une stabilité dès que la résistance source est supérieure à 3Ω et la self plus grande que 1pH. Dans ces conditions, la présence de la self à tendance à entraîner l'apparition d'instabilités numériques. De même, nous avons commence d'étudier l'influence d'un circuit de charge du coté drain. Des premiers résultats encourageants ont été obtenus dans le cas d'un circuit se présentant sous forme série, mais la résolution s'est avérée difficile et dépassait le cadre de ce travail.

En résumé, cette étude concernant les oscillations parasites montre que celles ci peuvent exister dans tout les composant planaires si certaines conditions sont satisfaites. Nous avons pu montrer aussi que les éléments parasites peuvent jouer un rôle dans la non-apparition de ces oscillations dans la pratique.

Chapitre 4

Optimisation des HEMT pseudomorphiques

4.1. Le but de l'optimisation

Pour avoir de meilleures performances avec les transistors HEMTs à plan de dopage et à couche d'InGaAs pseudomorphique, nous nous sommes intéressés principalement aux deux éléments: l'augmentation de la fréquence de coupure tout en gardant une bonne linéarité du gain, et l'augmentation du courant de sortie afin d'avoir une puissance élevée. Pour la première, une fréquence de coupure plus élevée passe généralement par la réduction des dimensions, notamment la longueur de grille. La fréquence de coupure est sensiblement proportionnelle à l'inverse de L_g. D'autre part pour augmenter le courant de sortie il faut augmenter la concentration de porteurs dans le gaz 2D (n_s).

4.2. L'épaisseur de la couche non dopée

Nous avons tout d'abord étudié l'influence de l'épaisseur de la couche barrière non dopée dans les structures HEMT pseudomorphiques à plan de dopage. La raison principale de cette étude était motivée par l'apparition des oscillations parasites qui nous obligeait, pour avoir des transistors 'normally-on', de réduire l'épaisseur de cette couche. Quatre valeurs différentes de l'épaisseur (L1) de cette couche ont étaient envisagées en gardant les autres paramètres identiques: la longueur de grille fixée à $0,3\mu$, le dopage à $3 * 10^12cm^-2$ et l'épaisseur de la couche active de $140\AA$ avec 20% d'indium. Nous avons pu remarquer qu'il existe une augmentation de la transconductance quand l'épaisseur de la couche diminue, cette augmentation correspond semblement à une évolution en 1/a où a est la distance totale de la grille au gaz 2D.

La capacitance de grille C_{gs} augmente et la variation relative en fonction de Vgs change aussi quand la distance L1 diminue; ceci peut s'explique comme précédemment par l'évolution de la position relative du gaz 2D.

La fréquence de coupure du gain de courant fait apparaître très peu de variations avec L1; il est donc intéressant d'utiliser des couches minces pour éviter les instabilités mais dans les limites de la tension de claquage imposées. Des études faites dans notre laboratoire ont indique une épaisseur minimale de $100\AA$. Puisque nous

envisageons par la suite une étude en fonction de L_g, l'utilisation d'une couche mince est nécessaire pour diminuer les effets de canal court. Une épaisseur de $100\AA$ nous est apparue comme optimale pour la suite de notre étude.

4.3. Longueur de grille

La diminution de la longueur de grille a un double intérêt. D'abord elle donne lieu à l'augmentation de la vitesse effective grâce aux effets de survitesse, et elle permet de réduire le temps de parcours moyen des porteurs sous la grille ce qui entraîne une réduction du temps de transit. L'augmentation de la vitesse effective des porteurs ainsi que la réduction de la longueur de grille peuvent donner une augmentation de la fréquence de coupure du gain de courant selon la formule simplifiée

$$fc \approx < v > /2\pi L_g \tag{4.1}$$

Nous avons donc simulé quatre composants de longueurs de grille différentes: 0,1, 0,15, 0,2, et 0,3 μ. Nous avons pu remarquer qu'il existe une augmentation de la fréquence de coupure quand la longueur de grille diminue; cependant cette augmentation ne semble pas linéaire comme prédite par la formule simplifiée. En particulier, nous remarquons que la montée de la fréquence est moins rapide quand la longueur de grille est de $0,1\mu$. Cela peut être expliqué comme une conséquence des effets de canal court. Ceci est encore plus visible en remarquant le décalage important de la tension de pincement (ce décalage a était réduit en utilisant une couche tampon AlGaAs). L'augmentation énorme de la conductance de sortie qu'on trouve quand la longueur de grille est réduite constitue une autre preuve des effets importants de canal court. Les répartitions de la concentration de charge, de l'énergie et du potentiel dans les deux composants à 0,1 et 0,3 μ montrent bien l'injection importante des porteurs dans la couche tampon pour des composants à grille courte. Il semble donc important de trouver une solution pour améliorer la conductance de sortie, une technique recommandée est l'utilisation d'une zone recess suffisamment large.

4.4. Offset de recess

Dans ce contexte, un troisième paramètre qui est d'extrême importance est la distance de séparation coin grille coin recess (DCGCR). L'utilisation des valeurs élevées peut entraîner une dégradation des performances en régime grand signal et il nous a semblé important d'étudier l'évolution des fréquences de coupure jusqu'à des tensions drain source élevées (10V). Nous avons trouvé que la valeur maximale de la fréquence de coupure du gain de courant varie très peu quand la DCGCR passe de zéro à $0,2\mu$ pour une tension de drain égale à 2,5 volts. Quand on a une tension plus élevée, 10V par exemple, une distance de $0,1\mu$ semble avoir une influence mineure sur f_{cmax}. Par contre, pour 0,2 μ de DCGCR, f_{cmax} présente une diminution relativement importante (15%) indiquant que cette distance constitue une limite pour ce type des composants.

D'autre part, nous avons étudié la variation de la conductance de sortie en fonction de ce paramètre; une DCGCR de 0,2 μ entraîne une diminution de la conductance de sortie par un facteur égal à 50%. Ceci va avoir un impact favourable direct sur les fréquence de coupure du gain maximal et la fréquence maximale d'oscillation.

4.5. Optimisation de n_s

Dans ce but, nous cherchons à augmenter la concentration des porteurs de gaz 2D tout en améliorant ou au moins en évitant de dégrader d'autres paramètres. Pour augmenter le n_s il y a deux approches: le premier est d'augmenter le dopage du plan avant, le deuxième est d'introduire un plan arrière de dopage. Dans le cadre de la première approche, nous avons simulé deux composants: le premier avec un plan avant dopé à $3 * 10^{12}$ et le deuxième dopé a $5 * 10^{12} cm^{-2}$. Nous avons trouve une augmentation de n_s qui passe de 2,3 a $2,9 * 10^{12} cm^{-2}$ et du courant maximum de 300 à 580 mA/mm. Par ailleurs, nous avons trouve une linéarité meilleure avec le dopage plus élevé; mais bien entendu le seul désavantage était d'avoir une stabilité plus réduite.

En introduisant un plan arrière de dopage égal à $1,5 * 10^1 2cm^-2$, le n_s max passe à $4,2 * 10^1 2cm^{-2}$ et le courant maximum obtenu est de 740mA/mm. Ces deux

grandeurs deviennent égales à $4,6 * 10^{12} cm^{-2}$ et 1160mA/mm quand le dopage du plan arrière passe a $2,5 * 10^{12} cm^{-2}$. Mais nous avons remarque que dans la zone de courant inférieur (350mA/mm), la fréquence de coupure de ce dernier composant présente une chute très importante. Cela veut dire que tant au niveau des performances faible bruit qu'en ce qui concerne le rendement, il semble qu'un dopage trop important du plan arrière ait des conséquences négatives. Une valeur optimale du dopage de ce plan semble donc de l'ordre de $2 * 10^{12} cm^{-2}$.

Chapitre 5

Simulation des HEMTs en claquage

5.1. Analyses physiques

Les modèles existants ne permettent pas une étude valable des HEMTs en régime de claquage, soit parce qu'ils sont trop simplifiés ou parce qu'ils ignorent la nature énergétique du phénomène de multiplication par ionisation et d'avalanche. Nous avons utilise notre modèle pour la simulation de transistors HEMTs conventionnels en régime de fonctionnement allant jusqu'au claquage. Les taux d'ionisations des électrons étaient pris comme fonction de l'énergie moyenne totale et celle des trous comme une fonction du champ électrique local en admettant que leur énergie suivait instantanément le champ électrique.

La première série des simulations a était effectuée pour un transistor quasi planaire où nous avons négligée les effets de surface. Les caractéristiques I-V de ce composant ont montre un accord satisfaisant avec les résultats expérimentaux et avec les conclusions d'un raisonnement physique approché. Seul les valeurs de la tension de claquage au pincement peuvent être différentes car on ne tient pas compte du courant tunnel de grille qui peut être le phénomène qui domine le claquage dans ces conditions de fonctionnement. Nous avons donc limité nos simulations à de valeur des Vgs assez négatives mais pas très proches de pincement par exemple à Vgs=-0,3V quand la tension de pincement est de Vp=-0,5V.

En faisant une analyse assez rigoureuse nous avons remarqué que les tensions de claquage les plus faibles sont corrélées à une forme assez large du domaine de haute énergie. Ceci est vérifie en regardant l'évolution de la forme du domaine en fonction de Vds à Vgs fixée et pour différentes valeurs de Vgs au claquage.

5.2. Effet des dimensions des recess

L'utilisation d'un recess large entraîne une chute du fc quand la tension de drain est élevée, mais il est connu qu'il peut en résulter aussi une amélioration de la tension de claquage. Il faut donc trouve un compromis entre ces exigences contradictoires. Nous avons simulé trois transistors HEMT conventionnels à longueur de grille de $0,3\mu$; nous avons étudié trois distances d'offset différentes: zéro, 0,1 et 0,2 μ. Un potentiel de surface de 0.5 V était introduit dans la zone de recess.

Nous avons remarqué qu'il existe une amélioration de la tension de claquage lors de l'augmentation de la distance d'offset de zéro à $0,1\mu$. Cette amélioration est due principalement à la modification de la forme de domaine de haute énergie qui se trouve inclus dans la zone de recess et non pas dans toute la zone grille drain. L'augmentation de l'offset de grille au delà de $0,1\mu$ entraine une amélioration de la tension de claquage à canal ouvert mais on ne trouve pratiquement plus d'amélioration proche du pincement. On trouve alors que la forme du domaine d'énergie est pratiquement inchangée quand la distance d'offset passe de 0,1 à 0,2 μ.

Nous avons donc conclu que l'introduction d'un recess assez large avec un offset de $0,1\mu$ peut entraîner une amélioration significative des tensions de claquage, mais qu'au-delà, compte tenu de la dégradation de la fréquence de coupure, l'utilisation des valeurs plus importantes ne semblaient pas intéressante.

Chapitre 6

Conclusion

Dans cette étude, nous avons modifiè un modèle hydrodynamique pour la simulation de transistors HEMTs pseudomorphiques. Nous avons pu introduire des effets physiques importants dans le fonctionnement de ces transistors. Ces effets sont le caractére bidimensionnel du gaz d'électrons dans le puits, l'effet d'"ecran, l'ionisation par choc et le courant de porteurs minoritaires. Les résultats obtenues sont en bon accord avec les résultats expérimentaux.

Des études systématiques ont montré que l'utilisation de plans de dopage est prèferable par rapport au dopage uniforme. La nature de la couche tampon a un faible effet sur le fonctionnement.

Dans toutes les structures HEMT planaires, il existe un phénomène d'oscillation parasite. Un circuit extérieur peut avoir un effet déterminant quant à l'apparition de ces oscillations.

L'épaisseur de la couche non dopée n'a qu'un faible effet sur la frèquence de coupure du gain de courant mais des couches minces servent a éviter les oscillations parasites et à limiter les effet de canal court.

Le limite minimale de la longueur de grille est égale à $0,15\mu$ si on veut bénéficier pleinement des potentialités des transistors. Des longueurs plus faibles entraînent des effets de canal court non désirables.

L'utilisation d'un offset de recess large améliore la conductance de sortie, mais on risque des dégradations des fréquences de coupure fc à des valeurs élevées de Vds.

Un plan arrière de dopage semble avoir un effet très important sur le fonctionnement des transistors de puissance. On trouve une augmentation de courant ainsi qu'une meilleure linéarité de gain, mais l'augmentation du niveau de dopage du plan arrière peut entraîner des effets secondaires indésirables concernant le fonctionnement faible bruit.

L'utilisation d'un offset modèrément large $(0,1\mu)$ entraîne une amélioration de la tension de claquage. Des distances plus larges n'entraînent pas forcement une amélioration significative sur tout en régime de pincement.

6.1. Limites actulles et Ameliorationes potentielles

Ce modèle peut être utiliser dans l'etat actuel pour la simulation de transistors HEMTs de structures assez complexes. La reduction des dimensions impose, d'une part, l'utilisation d'une methode plus exacte pour tenir compte du caractére bidimensionnel du gaz d'électrons. D'autre part il est important d'introduire les termes d'inertie dans l'equation de conservation du moment, au moins sous forme corrective.

L'existence de plusieres puits quantiques impose l'introduction du terme de gradient de chaleur dans l'equation de l'ènergie pour rendre compte de la variation de la temperature electronique à travers les heterojonctions.

Un effort important doit être dirigé vers l'amelioration des performances du modèle qui necessite l'usage de gros ordinateurs.

This page is intentionally left blank

Two Dimensional Simulation of High Electron Mobility Transistors: Contribution to the Optimisation of Millimetric Wave Devices

This page is intentionally left blank

Chapter I

Power FET's : A look through

Abstract

In this chapter, we will be reviewing the power requirements for FET's. The evolution of HEMT's is presented and its principle of operation discussed. The power limitations of HEMT's are outlined and the different variations used for better power performance presented. In addition, an investigation of the state of the art in power FET's will be made upon which the choice of pseudomorphic HEMT's is made. Throughout this review, we will present the concept of strained layers and the advantages in using them. Also a discussion concerning particularities in pseudomorphic HEMT's operation is presented.

I.1. Introduction

Since the proposal of using electric field for the current control was introduced by Shockley in 1952 and the devices based upon it, named field effect transistors, they have never ceased to inspire the engineers for higher power, faster operation and lower noise. And since the chase began, field effect transistors have undergone several stages of modifications. These were related to the advancement in technological process and discovery of new materials.

Whereas silicon and germanium, specially silicon, are the favourites for fast logic applications, due to the high compactness of their transistors suitable for VLSI, the compound semiconductors such as GaAs, InP, InGaAs, etc. established an unapproachable performance in microwaves applications. This performance extends to several fields of microwave applications, we find impressive low noise performance for MESFET's, AlGaAs/GaAs conventional HEMT's. Noise figures as low as 0.5 dB at 30 GHz and 0.9 dB at 60 GHz [1] were obtained using HEMT's[1].

In non-linear applications, MESFET's and HEMT's were extensively used in fabricating microwave mixers with impressive figures for gain conversion[2]. The FET's also demonstrated promising performance in microwave power applications. However difficulties were encountered due to the specific nature of power applications.

I.2. Power requirements

In order to understand why there is to be certain limitations talking of microwave power applications and how far could FET's go, we have to know what is required from a power FET.

A major criterion to evaluate FET power performance is the maximum output power. This is generally calculated as

[1]At 60GHz, this is for a $0.15\mu m$ AlInAs/InGaAs HEMT, a noise figure of about 0.6 dB was obtained for the same structure with $0.1\mu m$ gate

$$P_{max.} \approx 1/8 * I_{dss} * V_{br} \tag{I.1}$$

where I_{dss} and V_{br} are the saturation drain current and break-down potential respectively. Even though a static approach, this equation gives a good indication of the device output power in spite of external parasitic elements and current displacement under microwave operation conditions[1].

The second criterion is the power gain G. This parameter is in direct relationship with the current gain cut off frequency f_c and accordingly the maximum frequency of oscillation f_m,

$$G = K(fc/f)^2 \tag{I.2}$$

where K is nearly a constant An increase in these parameters (f_c and f_m) will give an increase in the gain.

The third criterion in measuring power performance is the power added efficiency given as :

$$\eta = \frac{P_{out} - P_{in}}{P_{dc}} = \eta_{dc}(1 - 1/G) \tag{I.3}$$

This gives a link between the power gain G, and the output power where η_{dc} is the drain efficiency P_{out}/P_{dc} which has a maximum theoretical maximum value of 50% for class A amplifiers and 78% for class B amplifiers. Since $1/G$ is a retarding factor, a high G is again desirable.

The fourth criterion is the linearity, this means that the gain should remain constant in a microwave signal swing. In class A amplifiers this is achieved when the small signal maximum available gain given by:

$$MAG = \frac{\left[\frac{f_c}{f}\right]^2}{\left[4g_d\left(R_i + R_s + R_g + \frac{\omega_c L_s}{2}\right) + 2\omega_c C_{gd}\left(R_i + R_s + 2R_g + \omega_c L_s\right)\right]} \tag{I.4}$$

is almost constant over the whole V_{gs} swing. Generally devices showing a fairly wide plateau of g_m and more precisely of f_m with V_{gs} will have a similarly behaving gain. As for class B amplifiers, in spite of the non linear nature of operation, the same

criterion can be applied due to the fact that harmonic suppression and fundamental output reconstitution can be achieved using push-pull configurations.

I.3. From MESFET's to MODFET's

The field effect transistors principle of operation depends upon the control of the current passing from the source to the drain by the potential applied on a gate controlling the channel. The type of the channel determines the type of the field effect transistor to a great extent.

The GaAs MESFET's use a doped GaAs channel for conduction (fig. I.1). The GaAs layer plays a double role; it provides the carriers necessary for the current carrying process and the transmission media for these carriers. The net current is a function of the carriers concentration and their drift velocity, the first depends upon the donors concentration, and the second is dependant on the channel material properties. In addition, the device current gain cut off frequency depends on the average drift velocity :

$$f_c \approx \frac{v_{eff}}{2\pi L_g} \tag{I.5}$$

where L_g is the gate length and v_{eff} is the effective velocity of carriers in the channel. To increase the current level, either the charge density or the drift velocity is to be increased or both if possible. The use of sub-micron gates gives rise to the possibility of having overshoot effects increasing the average carriers velocity under the gate. However the reduced gate length enforces a reduction in the channel thickness to overcome the short channel effects, this reduction in the channel thickness results in reducing the total charge density. To correct this draw back, the doping level in the channel is to be increased. The carriers drift velocity is determined by the momentum and energy randomisation process resulting from the interaction (collisions) with the ionised impurities. The increase in the doping level means an increase in the number of collision events since the current carriers and the ionised impurities share the same charge space region. Consequently there will be a drop in the average drift velocity and a reduction in both the current and the operating frequency.

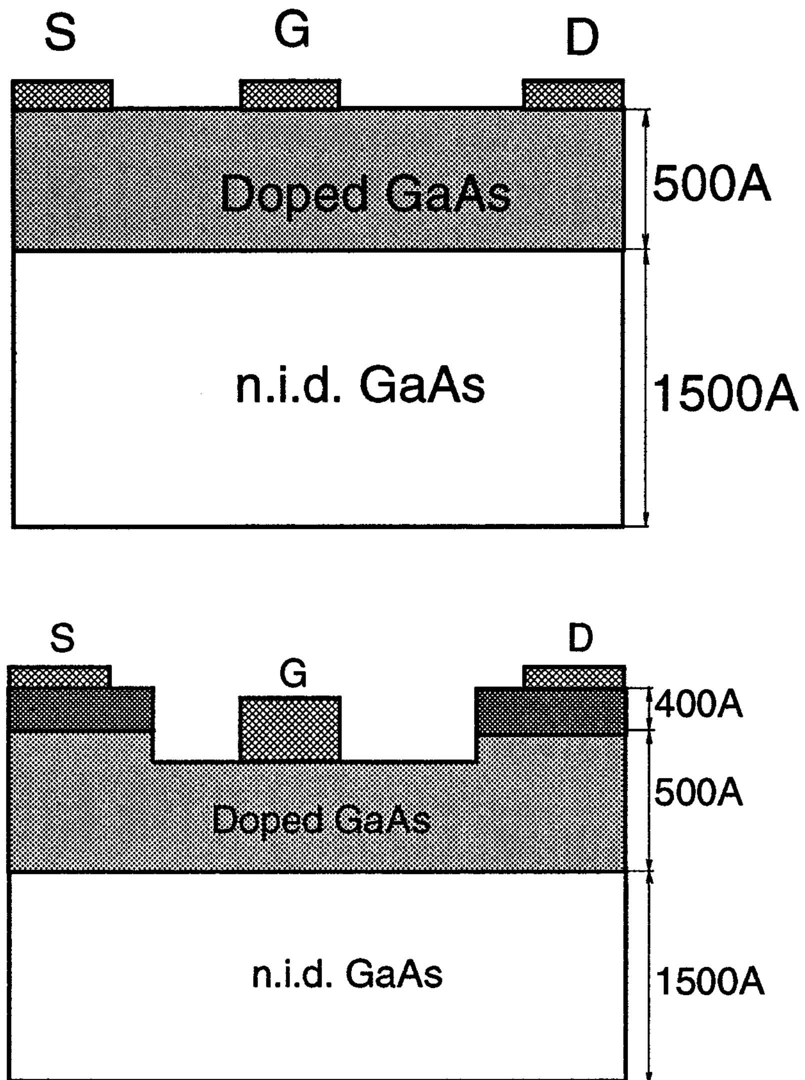


Figure I.1: GaAs MESFET; planar and recessed structures. The recessed gate structures were imposed by practical necessity.

Separating the carriers and their donor atoms represents a solution to solve the conflict between charge concentration and velocity reduction by increased scattering events. This could be achieved by introducing a heterojunction between two layers: one acting as the supply layer and the other as the channel. The resulting system is known as modulation doped layers.

I.4. Modulation doping principle of operation

The heterostructure, fig. (I.2), is lined up in such a way that the electrons placed in the smaller density of states, wide band gap material diffuse into the initially undoped small band gap material, presenting a plenty of empty states, across the heterojunction formed between the two materials. The accumulation of electrons on one side and the positive donors on the other side of the heterojunction results in an electrostatic dipole causing band bending in the conduction band. The electrons

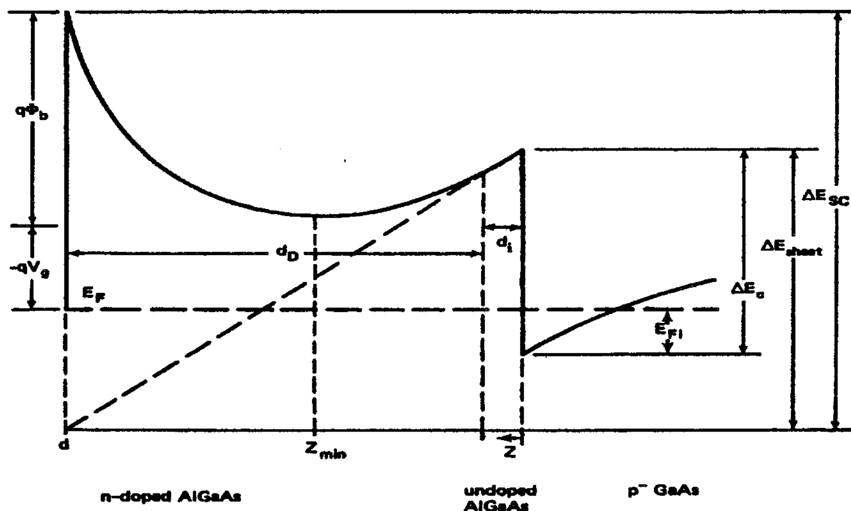


Figure I.2: Conduction band diagram of a heterojunction formed between wide band gap, heavily doped AlGaAs and non intentionally doped small band gap GaAs at room temperature.

are confined in the quantum well formed where they could only occupy discrete energy levels thus losing a degree of freedom in this sense. The electrons are now only able to move in two directions, this is behind naming the electron gas formed as Two Dimensional Electron Gas (2DEG) which gave the name TEGFET, Two dimensional Electron Gas FET to the resulting device fig. (I.3)[3].

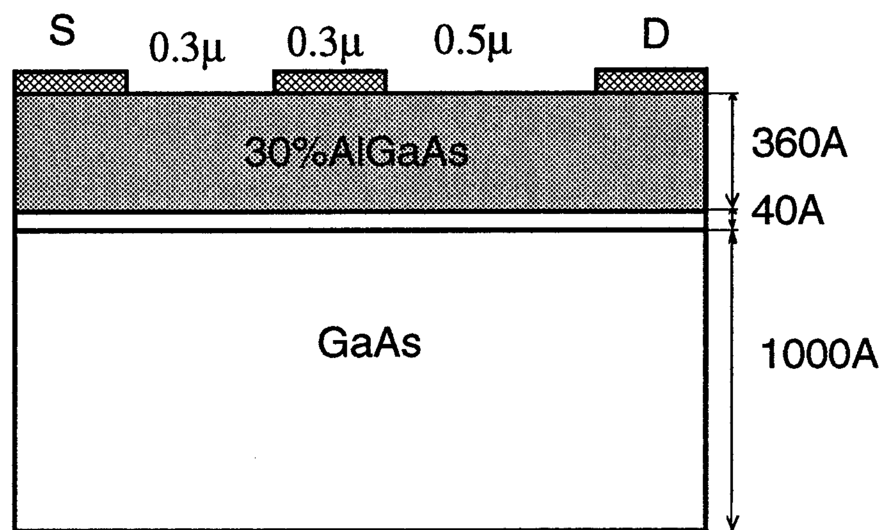


Figure I.3: A planar AlGaAs/GaAs MODFET

Of other names given to this device is the Modulation Doped FET, MODFET, High electron Mobility Transistor, HEMT, which exploits an important feature of such devices which is that the current carriers, electrons, present in the quantum well are spatially separated from the donor atoms occupying a different space charge region. They are saved the mobility reduction interactions with ionised impurities and they travel with a mobility close to that of the undoped bulk material they are in.

The FET's using the technique of modulation doping permit the introduction of high doping level and at the same time allowing the electrons to travel with high average velocity. This results in higher cut off frequency. The other thing that is achieved is a higher transconductance due to a reduced gate channel separation, (100-400$\AA$ instead of about 500-1000$\AA$ in MESFET's).

The conventional AlGaAs/GaAs HEMT's are fabricated using the state of art thin film deposition techniques. The growth techniques such as MBE and MOCVD have yielded high quality, smooth and extremely thin layers. The material system built up on GaAs substrates has been the base of high performance microwave HEMT's. The chase for higher output power and higher operation frequencies pushed device designers to look for nouvelle material systems and structures.

I.5. Alterations on HEMT structures

The three parameters that are continuously pushed up in power applications are the maximum output current, the operating frequency and the breakdown potential. For conventional HEMT's the process of increasing the current level passed by increasing the doping level and accordingly the breakdown potential is reduced[4].

It was then of importance to find replacements or alterations for the conventional HEMT for better power performance. Should we consider increasing the current level, the way to achieve this is to increase the 2DEG concentration in the first place. Increasing the conduction band discontinuity, ΔE_c, will result in an increase in the carriers interface density n_s[5]. In conventional HEMT's, this could be done by increasing the Al content in the barrier layer AlGaAs. Typical value used for

Aluminium mole fraction, 0.28, gives a ΔE_c of about 0.23 ev, increasing the Al content will result in increasing the concentration of the deep level traps responsible for several anomalies in HEMT operation specially at low temperatures. In addition, the increase in the 2DEG density tends to saturate for higher values of the Aluminium mole fraction[6],[7].

Another approach would be in adding a confining AlGaAs buffer layer, fig. (I.4), such

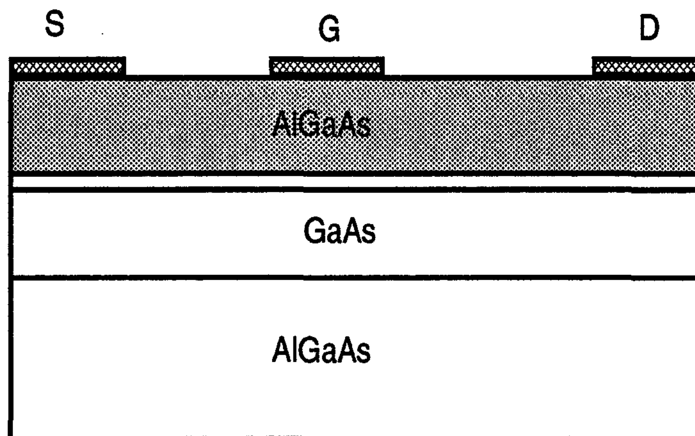


Figure I.4: A typical Single quantum well AlGaAs/GaAs/AlGaAs planar MODFET

a modification will introduce two quantum wells in the vicinity of each interface, the two being placed close to each other will merge into one and thus giving a single quantum well. This was found to improve the output conductance but to have a

slight effect upon the increase in the 2DEG concentration.

In this same perspective, the introduction of several AlGaAs/GaAs systems alternatively in the place of the n.i.d. GaAs buffer layer will result in creating several conduction channels in parallel[8][9]. This technique, although of limited n_s per channel, allows to have a total high n_s which will be more or less the sum of the different channel's charge density. The major drawback of such devices is the difficulty of conceiving an efficient access zone. The commonly used techniques (diffused contacts) for current collection may result in a high access resistance. In addition, such complex structures present a low yield which hampers their large scale production.

In the conventional HEMT frame always, there is the possibility of using a pulse doped barrier layer. This will permit the placement of a high doping density plane as close as possible to the channel while keeping the zone under the gate undoped. The advantage will be in having an improved breakdown potential but the increase in the n_s is not very impressive (15-20% of the uniformly doped n_s which is about $1.10^{12} cm^{-2}$) as there is always an upper limit for doping levels because of the trapping problem[10].

As a general observation, for all the conventional HEMT's, there exists an upper limit for the charge density that could be obtained which is about $1.2 - 1.4.10^{12}$ cm^{-2}. A radical change in the material system has to be made to achieve higher n_s.

The first step in changing the material system will be through the use of different compounds for the channel fabrication. The most appropriate is the use of InGaAs instead of GaAs. The InGaAs presents better transport characteristics than GaAs as the In component is increased[11]. However to use this material in a GaAs based system is a restrained process. The existence of a lattice mismatch between InGaAs and the GaAs or the AlGaAs atoms imposes a maximum thickness on the InGaAs layer so as the lattice mismatch could be accommodated as an elastic strain. This will be discussed in more details afterwards.

The InGaAs will provide a higher conduction band discontinuity than GaAs with AlGaAs for the same Aluminium mole fraction. An Al mole fraction of 0.28 is

necessary to have a ΔE_c of 0.23 ev with GaAs channel, whereas a mole fraction of 0.2 is sufficient to provide a ΔE_c of 0.31 ev with an InGaAs channel with 0.2 Indium mole fraction. In addition, this system has a naturally introduced confining layer which will add to the increase in n_s.

Having a barrier layer of AlGaAs gives the possibility of using the technique of planar doping for breakdown potential improvement at the same time while keeping the Al content low to avoid the deleterious effects of DX centres. It also benefits from the advanced technology in contact growth associated with AlGaAs layers. There is another possibility for increasing n_s. The first is in the introduction of a doped channel. This solution however has negative consequences on the drift velocity in the channel in the same context as the MESFET's. It might provide although an improvement in the breakdown potential specially in the near pinch off zone, the thing that is under extensive research currently. The second possibility is in introducing a doped buffer layer[12]. This was found to increase significantly the n_s, values as high as $4.10^{12} cm^{-2}$[13] were reported

The other family of devices that is based on the modulation doping technique is grown on InP. There are two types: the lattice matched InGaAs to InP which occurs for In content of 53% and the AlInAs/InGaAs pseudomorphic HEMT's which provide a very high n_s (7x10^{12} cm^{-2})[14]. The device performance figures are remarkable in the low noise applications specially in the 60-90GHz range. However, the commonly used metals for the fabrication of Schottky gates do not provide a high barrier over the AlInAs doped layers. This may result in a small breakdown potential, setting a limitation upon the usable doping levels and causing leaky gates. This latter invokes a solution based upon the use of an insulator layer under the gate to have a MIS like HEMT's or otherwise known as insulated gate HEMT. This device will be of main interest for high speed logic applications and the possibility of having complementary devices [15].

And since we talk about the power devices, the use of InP MISFET for power applications should be underlined. The InP presents a high threshold energy for the ionisation process thus it could provide higher breakdown voltages. This is also

emphasised by the use of MISFET structures. In addition the better transport dynamics of InP will allow for higher operating frequencies. However and in spite of that the InP MISFET have shown very good performance figures in the 30 GHz range, they present some major problems such as the insulator layer stability and transient time instabilities in the millimetre range[16]).

Another non field effect device which could be a potential candidate for power applications is the Heterojunction Bipolar Transistor (HBT). Yet the application range seems to be limited to the less than 30 GHz zone.

I.6. State of the art in power performance

It would be interesting by now to look comparatively on the power performance of the different mentioned devices at the time when we started this work (1992). When we look at the state of the art in power FET's, we can conclude from fig.(I.5), which gives the maximum output power, that for frequencies about 10 GHz, the InP MISFET's are the devices of choice, whereas at 30 GHz the multiple channel AlGaAs GaAs HEMT's present respectful performance. However, should we consider now

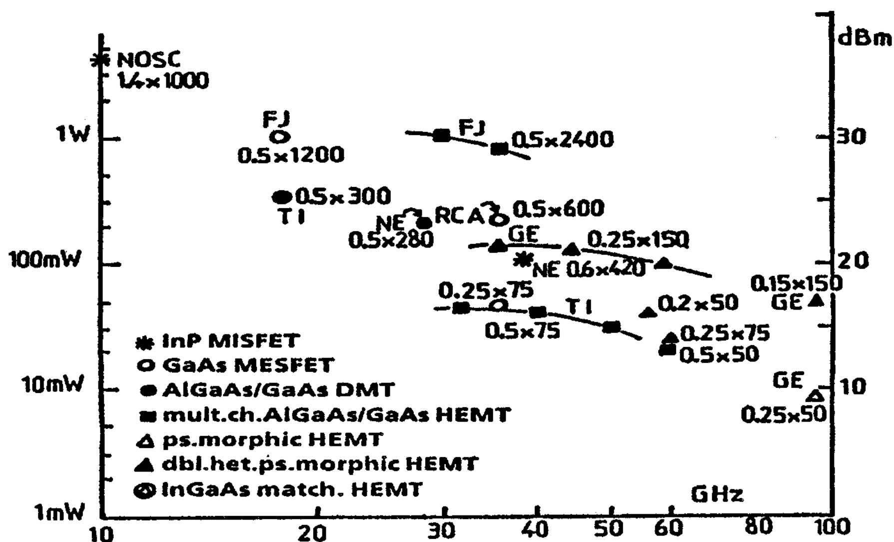


Figure I.5: Maximum output power for different field effect transistors versus operation frequencies, ref.[2]

the output power density shown in fig. (I.6), we notice that at frequencies higher

than 60 GHz all the devices fail to give a better performance. Yet, even with the general drop met in the output power density we find the InGaAs pseudomorphic

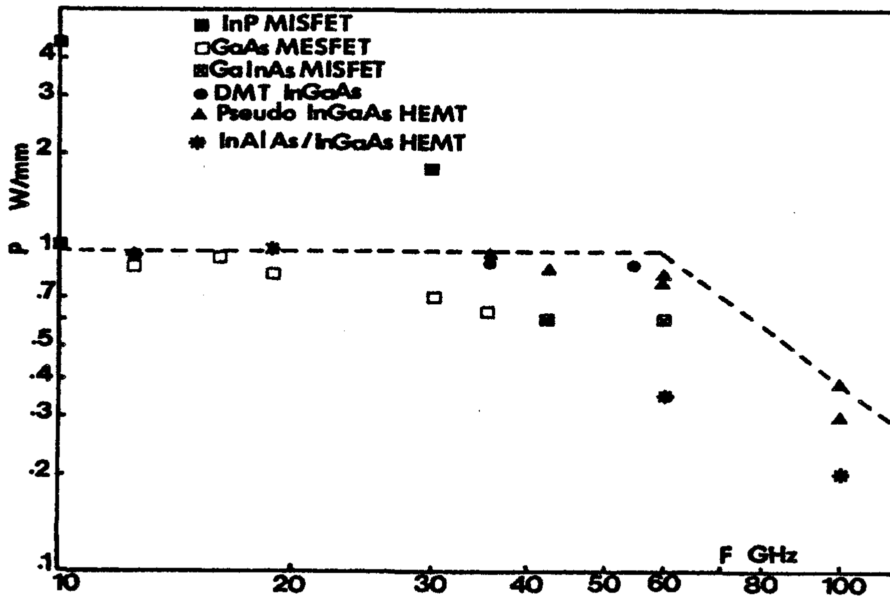


Figure I.6: Output power density with operation frequency, a general drop in the output power density for operation frequency higher than 60 GHz

HEMT's hold the pole for the most promising performance. As we said earlier, this plot was done by mid 1992, where the the most recent figures at the time were presented, it should be noted that since then impressive performance figures were published specially for the InP grown devices. On the other hand, for the different devices advancements were obtained mainly in the technological mastering of their fabrication giving the ultimate performance and increasing the yield. It could be seen that the pseudomorphic HEMT's with pulse doped barrier layer present the most suitable candidates for eventual rise in the frequency. To increase their frequency response a through optimisation of their structures is needed. This can be done only after a full understanding of the device major physical aspects.

I.7. Pseudomorphic Structures

I.7.1. Synthesis of pseudomorphic layers

In the growth of heteroepilayers systems, the different lattice constants of different layers have to match. There is an important problem that could be seen when we

examine fig. (I.7) giving the different lattice constants for compound semiconductors. In heteroepitaxial systems, lattice match could be obtained only over a limited range

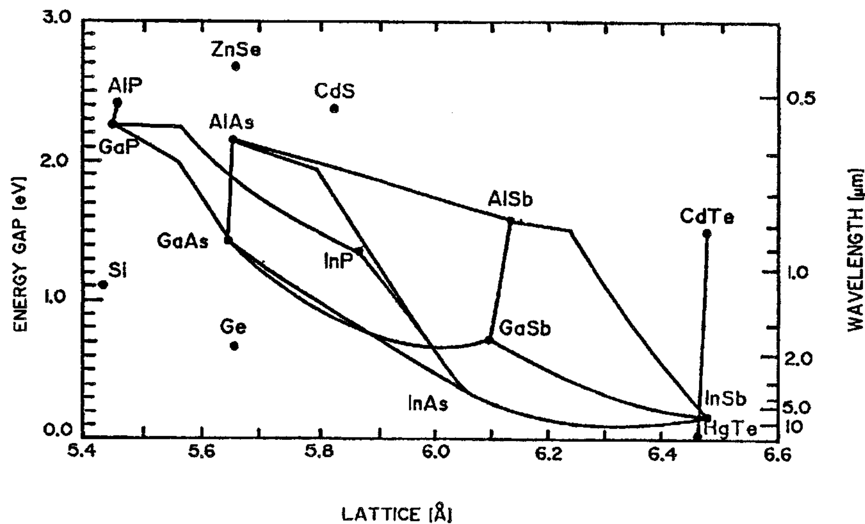


Figure I.7: Lattice constant, energy bandgap and wavelength for different semiconducting materials

of semiconductors. The degree of matching is an important factor that controls the crystal quality and compositional uniformity of the system. We can find that InGaAs will only match to InP for a high Indium content (47%) having a larger lattice constant than GaAs or AlGaAs. Attempts to grow two materials with different lattice constants will result in interface dislocation creating traps which will result in a drop in the carrier density and accordingly the current. This problem could be solved using pseudomorphic layers. It was proposed that lattice mismatch could be accommodated as an elastic strain in the intermediate layer provided that this layer stays thin, fig. (I.8)

The limited thickness of the strained layer is a function of the In mole fraction, Matthew[17] have studied the different thickness values that could be used without having a dislocation induced traps as well as a structurally stable device.

The strained layers system, or pseudomorphic structure, permit the growth of alternate layers of different lattice constants while keeping the total strain almost zero on the substrate and freeing the material from crystal faults.

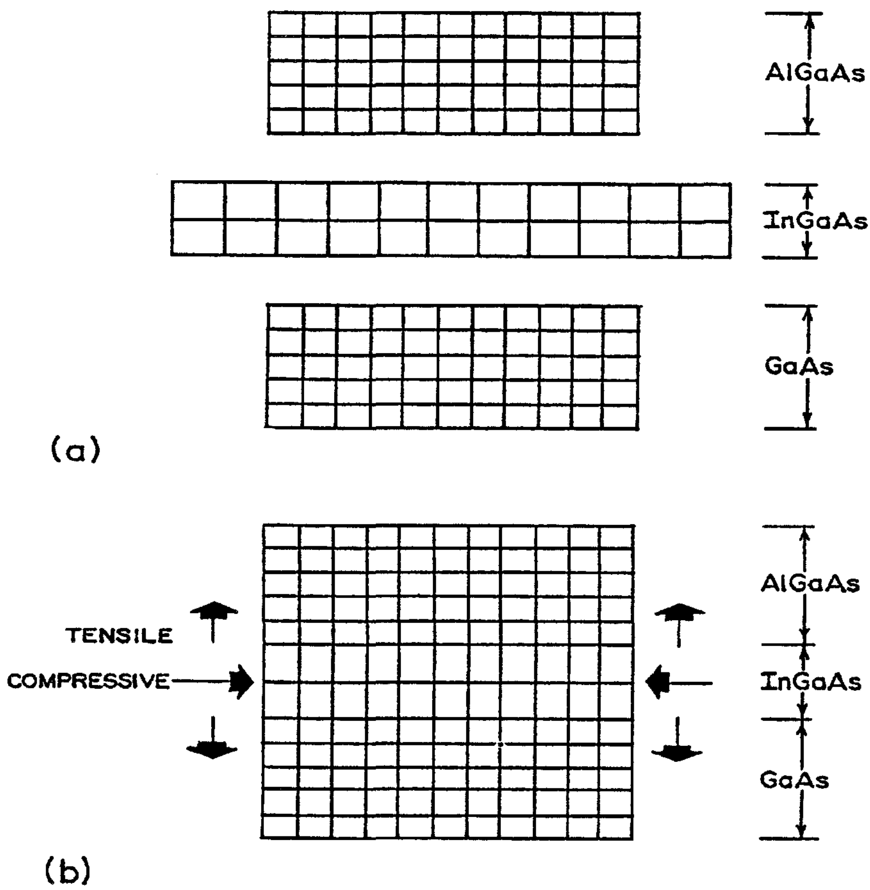


Figure I.8: Strained layer hetero epitaxi. Lattice mismatch could be accommodated as an elastic strain so that the tension upon the buffer layer is minimal

I.8. Optimisation problem of Pseudomorphic HEMT's

Optimisation of pseudomorphic InGaAs structures passes by the investigation of a large number of parameters. When we look at a ty[ical device, fig. (I.9), we will easily realise this fact. There exists a variety of parameters to investigate :

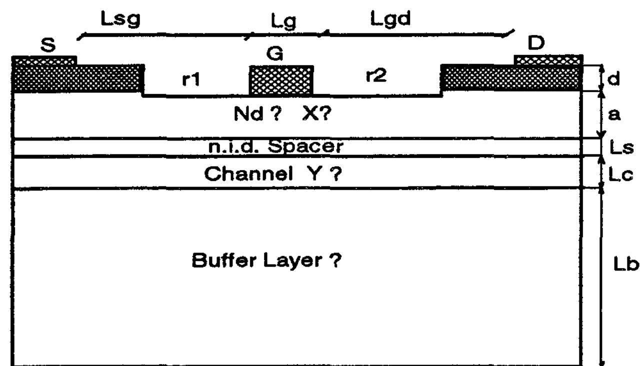


Figure I.9: A typical pseudomorphic InGaAs HEMT

- Geometrical; gate length, recess shape, epilayer thickness
- Material; mole fractions, doping profiles, doping levels, growth temperatures, etc.

As for example of the complexity of the problem, the increase in the doping level will result in increased current density yet a reduced breakdown potential will hamper the device output power. The increase in the Indium content will increase the charge density but will restrain the well's width to smaller values resulting in reduced mobility.

Another critical parameter is the recess offset distance that improves the breakdown potential but dramatically reduces both the output drain current and the current gain cut off frequency.

To perform the optimisation study in an experimental way means that a huge budget should be prepared. Add to this the uncertainty of obtained results due the interaction of different parameters and due to typical fabrication drawbacks such as the non uniformity of devices on the same wafer. Without devaluing the experimentation, a simulation tool is needed both for cost reduction and accurate approach.

I.9. Particularities in pseudomorphic HEMT's modelling

I.9.1. Carriers properties in strained layers

The InGaAs posses superior transport properties compared to GaAs which is an increasing function of the Indium mole fraction, however increasing this later will have deleterious effects upon the stability of the heterojunction as was previously stated in relation with the critical thickness criteria. However, pseudomorphic structures based on InGaAs channels grown over GaAs or AlGaAs substrates have the conduction channel made of a strained layer, this affects the band structure and transport dynamics. Not only this but the carriers effective mass and average energy are consequently affected

Should we consider for example compressive forces, the in plan atoms gets closer and consequently the conduction band and the valence band move away from one another and the band gap increases. The opposite situation occurs in the case of tensile forces where in plane atoms move away from each other and the band gap decreases.

The valence band on the other hand undergoes profound changes due to the presence of holes with large effective mass termed heavy holes, and light holes with small effective mass. In the presence of a strain, the original degeneracy in the band is removed and the two types of holes are separated with the light holes moving up or down away from the heavy holes function of the type of applied strain. For tensile strain, the light holes move up, and for compressive strain they move down. In short this characteristic of hole splitting permits the use of strained layers or the

fabrication of p doped HEMT's useful for logic applications.

Another effect of strain is related to the Indium mole fraction. The more this is increased, the narrower the channel will be, this will result in increased sub-band scattering.

These effects will lead to having lower carrier mobility in the strained layers as compared to what is found in bulk materials. The reduced mobility is then comparable to that of the bulk GaAs (7500-8200 $cm^2/V.sec$).

I.9.2. Carriers quantisation

In heterostructure, carriers are confined to a narrow well in the vicinity of the heterojunction. Accordingly instead of having a continuum of density of states, the carriers will be allowed to have only discrete energy values. The energy distribution could then be seen as being split into a series of energy subbands. This quantization could be understood when we look at fig.(I.10) where we can see how the conduction band is split. The wave vector k will have discrete values in the direction perpendicular to the interface k_z while in the other two directions k_x and k_y it will increase parabolicaly. In terms of energy, this means that the carriers will have the possibilty of occupying energy states in a continuous form in the x and y direction, while in the z direction discrete energy states will only be possible. The carriers will have a two dimensional nature of transport as long as they are confined in the well, upon certain values of the electric field, they will escape the two dimensional system into the three dimensional one. If the barrier height is smaller than the central-satellite valley separation, this transfer to the three dimensional system will occur at relatively low values of the corresponding electric field (less than the threshold field which is about 4KV/cm for GaAs), in the case of submicronic structures, to a drain voltage of about 0.5 V [19]. This might be true for conventional AlGaAs/GaAs devices where the barrier height is 0.23 ev which is smaller than the inter valley separation (≈0.34 ev). As for pseudomorphic HEMT's, such an assumption might be sort of a wild one. The barrier height in this case (0.31 ev) is comparable to the inter valley separation and it is possible that the carriers stay in the two dimensional system

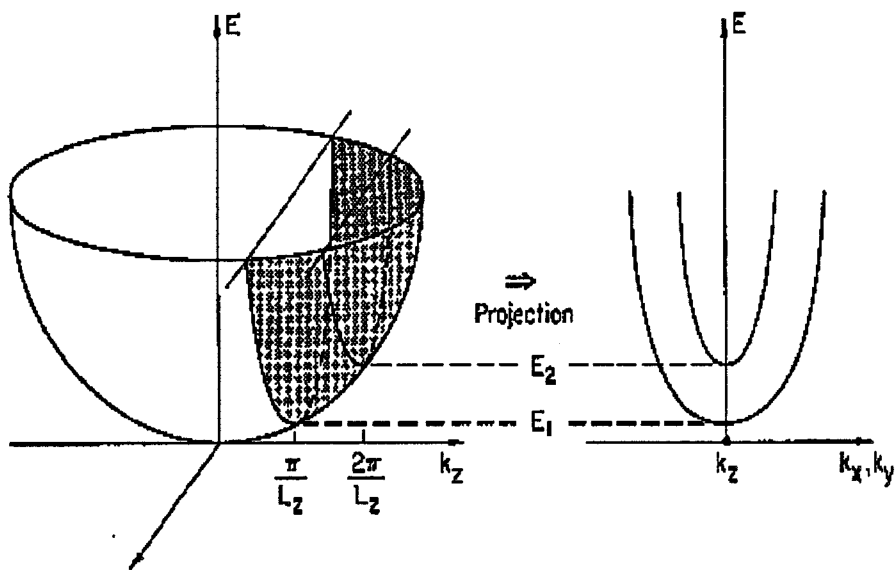


Figure I.10: Schematic presentation of the division of a conduction band into a sequence of subbands

for relatively high electric fields. Accordingly, throughout a considerable range of the carrier transport, they will have a flat band energy equal to, at least, the first sub-band rather than the bottom of the conduction band.

I.9.3. Screening effect

The experimental observations of early HEMT's have demonstrated the variation of the 2DEG mobility with the gate potential. This phenomenon was explained as a result of the increased interface charge density, the 2DEG (screens) the Coulomb interaction with the ionised impurities present in the nearby barrier layer (about 30-40 A separation) resulting in a net increase in the low field mobility.

In fig.(I.11) a thorough experimental study upon different devices have indeed put this phenomenon into evidence. Unfortunately, a satisfactory physical reasoning and quantitative explanation of this phenomenon are missing.

This phenomenon has a major effect upon device characteristics as the carriers do not travel with their maximum low field mobility except when the 2DEG is fully formed.

I.9.4. Impact ionisation

I.9.4.1. What is impact ionisation

A very important phenomenon occurring in power devices is the breakdown due to avalanche multiplication by impact ionisation. Avalanche multiplication could be induced by either type of current carriers; electrons and holes under going a series of impact ionisation events. Without any loss of generality, we will talk only about electrons, the same concepts are applicable to holes. When a free electron in the conduction band is subjected to an external electric field, it moves with a drift velocity $v = \mu(w) * E$ where w is electron's kinetic energy, and its energy starts to increase with the increased electric field. When the electrons travel in the material they undergo a certain number of interactions with the lattice in the form of energy

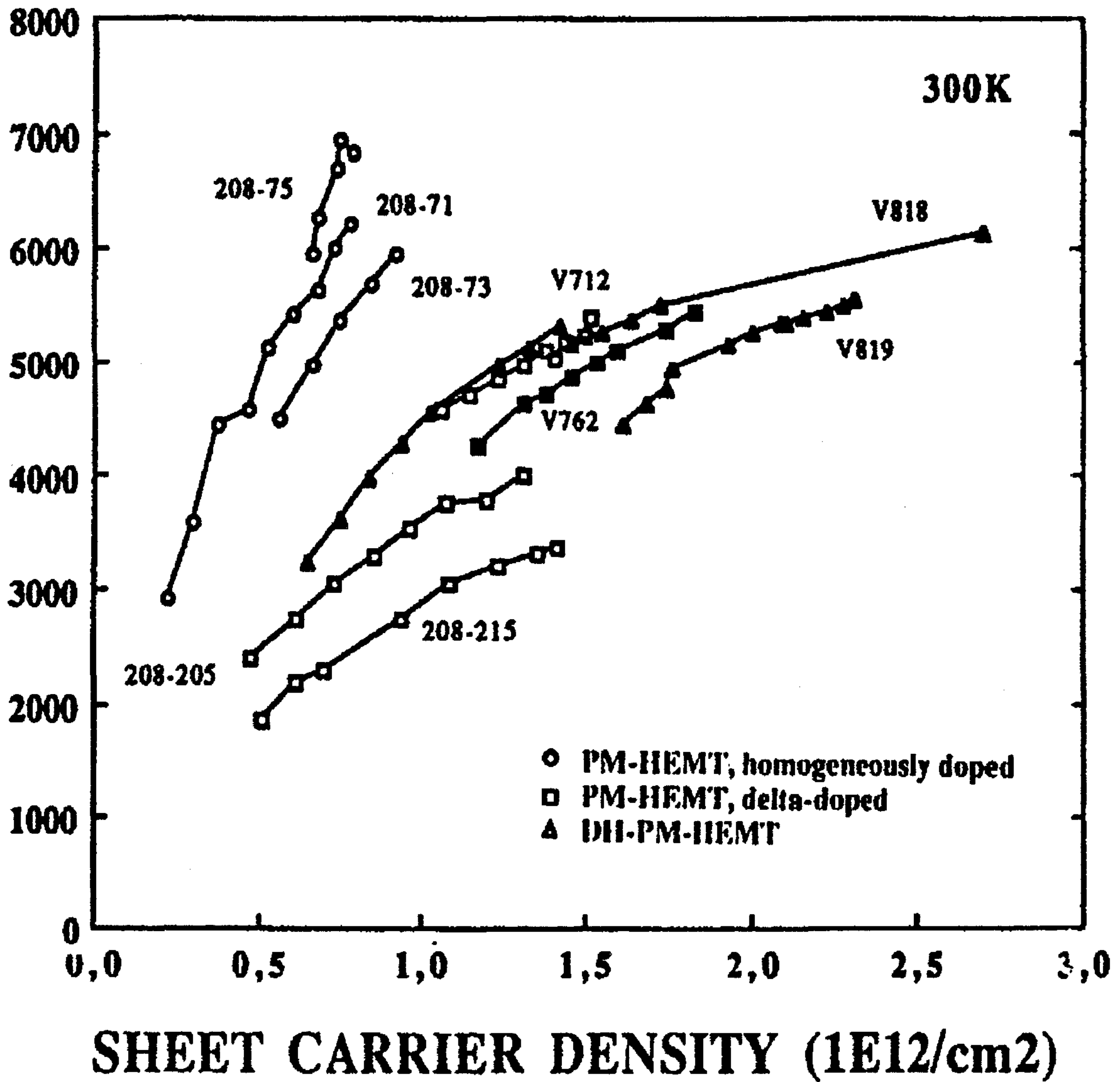


Figure I.11: Low field mobility versus 2DEG concentration for different HEMT's (courtesy of D.Pons, Thomson CSF)

dissipating and momentum randomising collisions. When the field is increased the rate of collisions increases due to increased collisions with the net result of having a constant drift velocity. This is the general figure in an arbitrary semiconductor, the details of velocity variation with electric field differs subject of the material conduction band structures and the different probabilities of scattering events that exist. What interests us here is what happens for still higher electric fields.

Always in the microscopic perspective, an electron might be lucky enough to travel a certain distance λ without undergoing any sort of energy dissipating interactions, it will have an energy known as ionisation energy, such that when they collide with an atom enough energy will be transferred to the valence electrons that they will be able to go to the conduction band leaving a hole in the valence band.

Another electron would have a different approach, this electron will encounter a series of energy dissipating collisions at which its energy is not sufficient to invoke an electron hole pair, and thus these collisions result in dissipating a certain portion of the electrons energy. However there could be a collision at which the electron energy is sufficient to liberate the valence band electron in an ionising event. Thus, in both cases, the electron has to acquire a minimum energy in order to be able to produce an electron hole pair in an ionising collision, this energy is known as the ionisation energy.

I.9.4.2. The principle of Ionisation energy

In a phenomenological approach, the ionising collision involves energy and momentum perturbation. For and initial electron in the conduction band, this electron is defined to be in state 1 with a wave vector k_1 and kinetic energy W_1 sufficient to create an electron hole pair, collides with an atom to liberate an electron which goes to the conduction occupying state 2 of wave vector k_2 and energy W_2 and leaves a hole in the valence band with wave vector k_v and energy W_v where as itself, the initial electron, occupies a new state 3 with wave vector k_3 and energy W_3. This situation could be summarised as follows:

$$k_1 = k_2 + k_3 + k_v \tag{I.6}$$

$$W(K_1) = \frac{(\hbar.K_2)^2}{2m_c} + \frac{(\hbar.K_3)^2}{2m_c} + \frac{(\hbar.K_v)^2}{2m_v} + W_g \tag{I.7}$$

Assuming that there is no energy loss to the lattice; no phonon generation, and assuming that the effective masses in the valence and conduction band are related by:

$$m_v = \gamma m_c \tag{I.8}$$

The minimum energy for ionisation will then be the one given by (I.7) assuming equal distribution of the momentum on the three output states in which case the wave vectors in the two bands are related as

$$K_v = \gamma K_c \tag{I.9}$$

Algebraic manipulation of (I.6) and (I.7) using the relation between the effective masses and wave vectors in the two bands will lead to the following expression for the ionisation energy:

$$W_i = \eta_g \left(\frac{\gamma + 2}{1 + \gamma} \right) \tag{I.10}$$

In the assumption of equal effective masses we get the following simplified relation

$$W_i = 3/2\eta_g \tag{I.11}$$

which means that an electron to be able to ionise has to have kinetic energy equal to at least one and a half the energy band gap. The exact value depends upon the

distribution of the energy with the wave vector i.e. the band structure. This for non isotropic bands will change function of the growth direction.

To sum up this paragraph, the impact ionisation is an essentially energy dependant phenomena that necessitates the existence of a minimum kinetic energy of an electron to be able to create an electron hole pair.

I.10. Conclusion

To conclude this chapter, the state of art in power performance of field effect transistors shows that the pseudomorphic HEMT's are the devices with the most promising performance. This performance however, shows a dissatisfaction concerning the output power density at frequencies higher than 60 GHz. A way to improve the output performance of pseudomorphic HEMT's is to optimise their structural design. This task requires a good understanding of the physical phenomena occurring in the device so as to be able to simulate them. Of the major phenomena manifested in heterostructure field effect transistors are the screening effect, carriers quantization and impact ionisation. These phenomena, specially the impact ionisation phenomena are highly energy dependant ones and this requires a full energy based treatment.

This page is intentionally left blank

This page is intentionally left blank

Chapter II

SIMFETII - The device simulator

Abstract

In this chapter, the physical model upon which SIMFETII is based will be presented and explained. The inclusion of the physical phenomena to enhance its performance are discussed as well as the characteristics inherited from it predecessors. The mathematical techniques are presented and the computational model is next detailed pointing out the strong points behind its unprecedented performance. Also the modifications done to render the simulator more handy are presented and discussed. Typical performance indications are presented.

II.1. Introduction

For semiconductor device simulation, a large variety of models exists. These models differ in their accuracy and their complexity, in general, the simulation models tend to satisfy as much as possible two major criteria : accuracy and rapidity.

The most accurate, most complex and the slowest accordingly, the Monte-Carlo models, are found at one end of the scale. This family of models could be divided into two groups, quantum Monte-Carlo models and those that could be named classical Monte-Carlo models. The Monte Carlo technique owes its precision to the nature of its microscopic treatment of the transport problem where an electron or an ensemble of electrons, is injected in the device subjected to different acting forces, external and internal, scattering mechanisms applied and the steady state distribution obtained [20, 21, 22]. Although quite accurate, there are still unresolved problems, or at least not in a satisfactory way, concerning some physical phenomena in the device. For example we have the particle-particle interaction in high density regions, the screening effect and interface scattering [23].

On the other extremity there are the analytical models based on close form formulas describing, under several and sometimes severe simplifications, the device characteristics [24, 25, 26]. These models are essentially very fast and far from being complicated, but on the other hand not so accurate and need continuous adjustment for each simulated device.

Between the Monte-Carlo family of models and the analytical models we have what are known as *semiclassical models* . This family of models is quite large and contains a lot of models varying in accuracy and complexity. They could all be traced back to the same origin, the Boltzmann Transport Equation. Although basically using a set of balance equations, it was shown that these balance equations are related to the BTE through the relaxation time approximation [27].

The first member in this group are the quasi-two dimensional models [28][29]. These models solve for the charge control in the device in one direction (perpendicular to

the gate) and for the current transport in the other direction. Very fast models, they represent a good tool for CAD techniques needing only a personal computer and give fast good indications on the device performance. However, these models have a major draw back in their neglecting of the two dimensional aspect of transport. This results in poor estimation of the device performance near pinch off and when the device dimensions are largely sub-micron.

The other members of this group could be classified further under another name: The Hydrodynamic Models. These models treat the carriers as a fluid (carrier gas) and thus the hydrodynamic name. The structure of such models is globally the same: Poisson equation and a set of balance equations. The drift diffusion models [30], consider only the current balance equation thus ignoring the energy balance in the device. This has a consequence of not being able to account for hot carrier transport found in short channel devices. The drift diffusion models fall flat in trying to simulate sub-micron structures.

We then find the Hydrodynamic energy models [31], in these models the energy balance is considered allowing to account for non isothermal transport and hot carrier effects. Such models present a powerful tool for the analysis of sub micron structure mainly field effect transistors.

They were used successfully in the analysis of MESFET performance [32], optimisation of sub micron MESFET's [33], performance analysis of InP MISFET's, [37], performance analysis of conventional HEMT structures at room temperature [39], and at cryogenic temperatures [40].

Recently, always belonging to the same category of models, quantum Hydrodynamic models were published [41] making it possible to model extremely small devices where the quantum transport starts to be a non negligible transport mechanism.

In this work, we have chosen the Hydrodynamic energy model because of its various advantages : A structure of moderate simplicity that permits the introduction of different physical phenomenon and a faster execution time compared to more complicated Monte-Carlo models with only slight difference in accuracy.

II.2. The Boltzmann Transport Equation

The Boltzmann Transport Equation is a semiclassical equation satisfied by the distribution function $f(k, r, t)$ such that :

$$\frac{df(k,r,t)}{dt} = 0 \tag{II.1}$$

When we expand this equation we get :

$$\frac{df}{dt} + \frac{df}{dk} \cdot \frac{dk}{dt} + \frac{df}{dr} \cdot \frac{dr}{dt} = 0 \tag{II.2}$$

This states that the sum of the effect of the net force equals to zero. The net force acting upon a system of carriers $\vec{F}$ is in fact the sum of internal and external forces:

$$\vec{F} = \vec{F_i} + \vec{F_e} \tag{II.3}$$

The basic equation of particle dynamics gives the net force acting on a system of particles could as:

$$\vec{F} = \hbar \frac{dk}{dt} \tag{II.4}$$

Now using (II.4) in (II.3) we get :

$$\frac{df}{dt} + \frac{df}{dk} \cdot \frac{F_i}{\hbar} + \frac{df}{dk} \cdot \frac{F_e}{\hbar} + \frac{df}{dr} \cdot \frac{dr}{dt} = 0 \tag{II.5}$$

Let us examine what happens in the absence of external applied forces. The last term in the left hand side gives the variation in position. This in the presence of only internal forces, which are mainly due to collisions ones, will tend to vanish. The third term, which represents external forces is non-existent, accordingly we can say that the effect of internal forces is equal to the rate of change due to collisions.

This yields the Boltzmann transport equation in its well known form :

$$\frac{df}{dt} + \frac{F_e}{\hbar} \cdot \nabla_k f + \vec{u} \cdot \nabla_r f = \left(\frac{df}{dt}\right)_{coll.} \tag{II.6}$$

The BTE is justified under the following assumptions

- Collisions involve only two particles
- The collision duration is much shorter than the inter collision time so collisions could be assumed instantaneous
- The external forces are constant throughout the collision duration and they remain so over a distance comparable to the particle wave function length

II.3. Simplifications to solve BTE

Of many assumptions done to solve BTE we present the two that are used in the current formulation : he relaxation time approximation and the displaced maxwellian approximation.

II.3.1. Relaxation Time Approximation

Whereas the left-hand-side of (II.6) is obtained in a rigorous way, the collision terms are too complicated to be treated in the same manner but rather in a quantum mechanical framework. A phenomenological approach which is of satisfactory physical reasoning could be used in this case.

In the absence of external forces, the collisions tend to randomise the energy and momentum to push them to the equilibrium state. It is of interest to suppose in a linear response after a perturbation, that the relaxation is proportional to the deviation from the original state. Hence, the decay of such a deviation will occur in an exponential manner. Thus, at any instant t, the function while relaxing to the original status would be

$$f(t) = f_0 + (f - f_0)e^{\frac{-t}{\tau}} \tag{II.7}$$

where f_0 is the equilibrium value. This permits to define the decay time τ as the system relaxation time.

So, in a general manner with all the assumptions from BTE holding, the scattering rate due to collisions could be approximated to

$$\left(\frac{\partial f}{\partial t}\right) = -\frac{f - f_0}{\tau} \tag{II.8}$$

The relaxation time τ is then to be calculated from the knowledge of different scattering rates through a Monte Carlo technique.

II.3.2. Displaced Maxwellian approximation

In the case of a homogeneous (independent of r), non-degenerate material, it seems reasonable to assume that the solution of the steady state BTE when the carriers are submitted to an electric field, could be approximated by a displaced Maxwellian distribution function given as [42]

$$f(k) = Aexp\left(-\frac{\hbar^2(k - k_0)^2}{2mK_BT_e}\right) \tag{II.9}$$

where A is a normalising factor, and T_e is the electronic temperature $=T_e(E)$ and $k_0 = k_0(E)$ where E is the electric field. The drift velocity will then be $v(E)$ and the average electron energy is

$$\xi_{av} = \langle 3/2K_BT_e \rangle \tag{II.10}$$

This approximation is very convenient as it is simple to use, however the electronic temperature is non measurable in contrary to the noise temperature. Yet it has been shown [43] that for a displaced maxwellian distribution function, the electronic

temperature is equal to the noise temperature and at the same time [same reference], such an approximation is difficult to satisfy unless the relaxation time is constant and independent of k which is non occurring in the situations of physical interest. In spite of that, the displaced maxwellian approximation could still be used in a rectifying frame which will be discussed later. As a general approximation, what is true for the displaced maxwellian distribution could also be considered true for any type of symmetrical distribution where the second order derivatives would vanish.

II.4. The Non Stationary model

II.4.1. The general model

Generally speaking, non stationary models are derived from the Boltzmann transport equation subjected to relaxation time approximation. However, being a general framework describing carriers distribution evolution in the phase space, and since a global picture of the band structure of the material is hardly feasible, a set of coupled BTE each of an equivalent minima has to be solved.

To get the set of transport equations, in each of the valleys, the method of moments is applied where the first three moments of BTE are retained. These moments present, in ascending order, the carrier density, average energy (or average temperature) and the average momentum.[1]

The BTE reads in terms of space variables (r, k, t) :

$$\frac{df_i}{dt} + \frac{\vec{F_i}}{\hbar} \cdot \nabla_k f_i + \vec{u_i} \cdot \nabla_r f_i = \left(\frac{df_i}{dt}\right)_{coll} \tag{II.11}$$

In each of the semiconductor conduction band valleys, if only three valleys are considered, the index i will be then referring to Γ, L or X.

Taking the first moment means multiplying this equation by $k = 1$ and integrating

[1]The model has been extensively detailed elsewhere [44],[36], we will only present the broad outlines of the derivation procedure

over the whole k space and for each distinct valley. This would yield the particle conservation equation:

$$\frac{\partial n}{\partial t} + \nabla(n\vec{v}) = \left(\frac{\partial n}{\partial t}\right)_{coll.} \tag{II.12}$$

Where the collision term involves both inter and intra valley scattering events and the index i to be implicitly understood.

Now if we multiply the BTE by $(m^*(\vec{r})\vec{v}(\vec{k},\vec{r}))$ we will get the second moment resulting in the momentum conservation equation which reads :

$$\frac{\partial nm^*(\vec{r})\vec{v}}{\partial t} + \nabla_r \cdot nm^*(\vec{r})\vec{v}\vec{v} + \nabla_r \cdot nK_BT_e + n\nabla_r E_c \cdot \frac{1}{M^*} - n\nabla_r ln(m^*)E = \left(\frac{\partial nm^*(\vec{r})\vec{v}}{\partial t}\right)_{coll.} \tag{II.13}$$

The physical interpretation of this equation shows that the first two terms in the LHS give the rate of change and the outflow of the momentum density. The next two terms represent the force exerted by the electron pressure and the force arising due to the position variation of the effective mass, the effective mass tensor contains all the information concerning the non-parabolicity. The last term represents the force induced by a heterostructure or by the band deformation resulting from high doping or induced strain. The RHS gives the rate of change of momentum due to collisions.

Multiplying BTE by W_i, the electrons kinetic energy in the valley i given as:

$$W_i = 1/2m^*v^2 \tag{II.14}$$

and then integrating as before will yield this time the third moment of BTE giving the energy conservation equation (the subscript i is to be implicitly understood):

$$\frac{\partial(n\bar{W})}{\partial t} + \nabla_r n\bar{W}\vec{v} + \nabla_r(\vec{v}\cdot nK_BT_e) + \nabla_r nQ + \nabla_r En\vec{v} = \left(\frac{\partial(n\bar{W})}{\partial t}\right)_{coll.} \tag{II.15}$$

The first two terms represent the rate of change and the out flow of the kinetic energy density. The next three terms represent the work done by the electron pressure, the divergence of the heat flow and the energy supplied by the effective electric field. The RHS gives the change in the energy density due to collisions.

II.4.2. Collision Terms

As we have seen, the transport terms resulting from external forces are derived rigorously through the moment method, unfortunately this is not possible for the collision terms. The relaxation time approximation was validated as a way to find the collision terms [45], the generalised expressions were given in [46] as:

$$\left(\frac{\partial n}{\partial t}\right)_{coll.} = -\frac{n_i}{\tau_{nij}(\xi)} + \frac{n_j}{\tau_{nji}(\xi)} - R_{ni} \tag{II.16}$$

$$\left(\frac{\partial P_{xi}}{\partial t}\right)_{coll.} = -n_i m_i^* \vec{v_{xi}} \left(\frac{1}{\tau_{pii}(\xi_i)} + \frac{1}{\tau_{pij}(\xi_i)} + \right) = -\frac{q n_i v_{xi}}{\mu_i(\xi_i)} \tag{II.17}$$

$$\left(\frac{\partial n_i W_i}{\partial t}\right)_{coll.} = -\frac{n_i(W_i - W_0)}{\tau_{Wii}(W_i)} - \frac{n_i W_i}{\tau_{nij}(W_i)} - \frac{n_i(W_j + E_{cj} - E_{ci})}{\tau_{nji}(W_j)} \tag{II.18}$$

For the particle conservation, the first two terms represent the particle transfer from the ith valley to the jth valley and vice versa. The third term gives the net generation-recombination rate.

As for the momentum relaxation time, the inter-valley and the intra-valley scattering rates are included. Since the momentum is assumed to completely randomised, there is no momentum transfer from the satellite to the central valley.

The energy relaxation time includes three terms; the first accounts for the relaxation towards lattice temperature in intra-valley scattering, the second for the energy leaving the ith valley in inter-valley scattering and the third for the energy supplied to the ith valley by scattering from the jth valley.

We stress upon the major assumptions used in driving these terms :

- The scattering rate is independent of the distribution function in the destination valley
- The scattering probability is only dependant on the electron kinetic energy.

II.4.3. Single equivalent gas concept

II.4.3.1. Physical reasoning

The set of equations given above is characteristic of each valley, Γ, L and X. This means that a complete set of equations is formed of nine conservation equations plus Poisson's equation. This presents a heavy task even for the most smart programming techniques and powerful machines.

To find an appropriate way out, the single equivalent gas concept is used. This is based upon the existence of a single equivalent gas instead of three gases in the three valleys. The properties of this gas is a weighted average of the three different gasses. This averaging represents a lot of difficulty to be performed in an exact way from a mathematical point of view, thus a physical reasoning or framing is required.

Using this physical framing, the resulting properties give a satisfactory description of the equivalent gas. For example let us consider the velocity field characteristics, where the negative differential mobility phenomenon in GaAs is a result of intervalley transfer

In a three valley frame, the carriers initially in the high mobility central valley are heated up by an applied electric field. Later they are scattered into the empty, low mobility satellite valleys. This process continues with increased field and results in the reduction of fast carriers density and the increase of heavy slow carriers density, thus a decrease in the average carrier velocity.

Let us see now what is the case in the single equivalent gas frame; there is no

inter valley scattering but as the carriers energy increase they become heavier and accordingly they slow down giving to a reduced velocity with increased field. But since we talk here of an ensemble of carriers in the conduction band as a whole, the energy considered has to be the average total energy of carriers.

II.4.3.2. The average total energy concept

Special attention should be paid concerning the energy balance equation when averaged over the different valleys. This equation considers so far the kinetic energy of electrons. This is particularly cumbersome to evaluate since it needs information concerning the upper valley population. The following transformation is done:

$$\xi_i = W_j + E_{cj} - E_{ci} \tag{II.19}$$

where ξ the total energy, may be considered as the sum of the kinetic and the potential energies in a valley i. When the averaging is performed we obtain $< \xi >$; the average total energy of the electrons in the conduction band.

Thus we conclude that the momentum and the energy relaxation times, effective mass and mobility, become a function of this average total energy of carriers to present similar variation to that occurring in the three valley frame.

The information related to the intra-valley transfer as well as the different scattering events are used by Monte-Carlo material simulators to get the single equivalent gas dependence of different parameters on the average total energy of the particles.

II.5. The generalised single gas model

The four equations single gas model gives the possibility of precisely describing the behaviour of the major quantities defining carrier transport in the device; potential, energy and charge concentration. The three transport equations, or known as conservation equations, are derived as we said before by the moments method applied to BTE, giving the continuity of the carriers density, average momentum and total energy, the fourth equation is Poisson's equation.

The set of equations read for an electron gas:

1- Poisson's equation :

$$\nabla \cdot \epsilon \nabla V = q(n - N_d - p + N_a) \tag{II.20}$$

2- Carrier's density conservation equation :

$$\frac{\partial n}{\partial t} + \nabla \cdot (nv) = G - R \tag{II.21}$$

3- Carriers average momentum conservation equation :

$$\frac{\partial (m^* v)}{\partial t} + m^* v \nabla v - \frac{1}{n} \nabla (nKT_e(\xi) = q\vec{E} - \frac{m^* v}{\tau_m(\xi)} \tag{II.22}$$

In each of the three directional axis x,y and z

4- Carriers average total energy conservation equation:

$$\frac{\partial \xi}{\partial t} + \frac{1}{n} \nabla \cdot Q + v \nabla \xi + \nabla \cdot (vKT_e(\xi)) = qv \cdot \vec{E} - \frac{\xi - \xi_0}{\tau_\xi(\xi)} \tag{II.23}$$

where the electrons average total energy ξ includes thermal, kinetic and potential energies of the carriers present in the satellite valleys.

$$\xi = 3/2 K_B T_e(\xi) + 1/2 m^* v^2 + U \tag{II.24}$$

where $T_e(\xi)$ is the electronic temperature and U is the average potential energy.

This model when used as a whole without simplifications is known as the momentum-energy conservation model. Though highly accurate, such a model is sometimes unnecessarily complicated and has to be simplified to be of practical use.

II.6. The hydrodynamic energy model

The hydrodynamic energy model is the base of the model used in this work. It is a single equivalent electron gas model subjected to a certain number of justified assumptions in order to simplify its structure for faster performance and easier handling. We will discuss these assumptions for each equation on its own.

1- Poisson's equation :

In the adopted equation the dielectric permitivity is considered as a scalar and not a tensor. The resulting Poisson equation is then (if only electrons are considered) as follows,

$$\nabla^2 V = \frac{q}{\epsilon}(n - N_d) \tag{II.25}$$

2- Electron's density conservation equation:

The electrons generation rate as well as the recombination rate are neglected provided that the electric fields are of moderate values. The resulting equation would then be

$$\frac{\partial n}{\partial t} + \nabla \cdot (nv) = 0 \tag{II.26}$$

3- Average momentum conservation equation :

The general equation reads :

$$\frac{\partial (m^* v)}{\partial t} + mv \cdot \nabla v - \frac{1}{n}\nabla(nKT) = q\vec{E} - \frac{m^* v}{\tau_{mv}} \tag{II.27}$$

The first approximation is not to consider the temporal derivatives, this assumption is backed up by the fact that the momentum relaxation time is much more shorter

than the energy relaxation time. Afterwards the algebraic manipulation readily yields the generalised current density constitutive equation as

$$J_n + \tau_{mv} J_n \cdot \nabla(J_n/qn) = \mu_n(qn\vec{E} + \nabla(nKT)) \qquad \text{(II.28)}$$

where the parameter

$$\mu_n = \frac{q\tau_{mv}(\xi)}{m^*(\xi)} \qquad \text{(II.29)}$$

is the generalised energy dependant mobility.

The second term in the LHS of (II.28) is a convection term which accounts for spatial momentum variation due to spatial fields non homogeneity. This term was found to cause numerical instabilities [2] [36] and neglecting it is of little effect unless very high field gradients exist. This would finally give the following form for equation (II.28)

$$J_n = \mu_n[qn\vec{E} + \nabla(nKT)] \qquad \text{(II.30)}$$

At a first glance, this equation resembles the one used in drift-diffusion models. However there are two basic differences; the first is that the mobility is not a function of the electric field but rather the average total energy. The second difference is in the diffusion term which when expanded will account for the gradients in the electronic temperature;

$$J_n = \mu_n(\xi)qn\vec{E} + \mu_n(\xi)KT(\xi)\nabla n + \mu_n(\xi)nK\nabla T(\xi) \qquad \text{(II.31)}$$

Knowing $\mu(E)$ and $E(\xi)$ permits the derivation of the generalised $\mu(\xi)$ curve. We will assume that these relations could be obtained from characterisation or Monte Carlo

[2] It might be possible however to introduce this term in a corrective recurrent form

models in steady state and bulk materials which means zero spatial and temporal derivatives.

4- The energy conservation equation :

The generalised conservation equation for the electrons average total energy gives :

$$\frac{\partial \xi}{\partial t} + \frac{1}{n}\nabla \cdot Q + v\nabla\xi + \nabla \cdot (vKT_e(\xi)) = qv \cdot \vec{E} - \frac{\xi - \xi_0}{\tau_\xi(\xi)} \tag{II.32}$$

The second term in the LHS gives the gradient of heat flow Q. This later could be calculated as :

$$\nabla Q = \kappa \nabla T(\xi) \tag{II.33}$$

where κ is the electrons' thermal conductivity .

This term could be neglected based on two approximations; the first is that the electrons' thermal conductivity is very small and the second is that ∇Q is proportional to the third moment of the distribution function. Accordingly for symmetrical distribution functions, this term would vanish.

II.7. Discussion

The model just presented was the one efficiently used and validated for the simulation of conventional HEMT's [36], whereas simpler version[3] were used for the simulation of MESFET's[47] and InP - MISFET's[37]

There is a certain number of points to be discussed in the structure of this model.

- The different parameters are assumed to be function of the average total energy independent of the field orientation. Their functional dependence could be

[3]The simplifications involved mainly the energy balance equation

derived from the steady-state Monte-Carlo simulations, this implies assuming that the distribution function is only a function of the average total energy.

- The distribution function in each valley could supposedly be specified by its local parameters n, v, ξ the validity of this approach are discussed in [48]
- While taking the second and third moments of BTE, the effective mass and electronic temperature appear originally as tensors. Assuming a slightly deviated from symmetry distribution function, these tensors are reduced to scalars. In the same context, the rate of heat flux which appears as a third moment of the distribution function reduces to zero. These approximation are acceptable under non-equilibrium conditions if the distribution function is quasi-displaced maxwellian or Fermi-Dirac shaped.
- No formation of subbands is considered. In the case of simple conventional HEMT's, this was verified by Monte-Carlo steady-state simulations [49] which demonstrated that even at low V_{ds} values , the carriers are found mainly in a three dimensional system(c.f. the discussion in chapter I). This might cause problems for structures of high energy barrier or single quantum well
- Gradual compositional variation is assumed across the heterojunction. Reducing the grading distance over which the variation occurs approaches us to the abrupt limit although with a lower bound enforced by the instabilities and numerical noise limit.

Due to the neglecting of carrier generation, the electric fields has to remain low so as to avoid heating the carriers to the impact ionisation limit. This would limit the use of the model to low noise applications discarding possible studies applied to power devices.

In addition, further physical phenomena has to be accounted for to improve the model performance such as screening effects.

In the following section we will discuss the enhancements introduced by the author to the model that will widen the range of application of the model to more complex and nouvelle structures for power application.

II.8. Enhancements of the model

II.8.1. Simulation of carrier's quantization

As we saw in chapter I, the heterojunction results in the formation of a quantum well in the vicinity of the interface. A general approach would be to solve the Schrodinger wave equation coupled to the system of equations constituting our model. This is a heavy mathematical task which would further complicate the model. Another approach would be based upon the physical reasoning related to the specific structure treated. This approach relies on the fact that in most HEMT's at least the first subband is completely filled when the heterojunction is at equilibrium. Accordingly, we assume that the electrons have a flat band energy equal to that of the first subband and not that of the bottom of the conduction band as before. This would result in causing the heating of electrons to occur earlier with eventual consequences upon their transfer and accordingly the transport dynamics as well as the real space transfer phenomenon.

Calculation of the equivalent energy[50]

In most practical applications, the quantum well formed at the hetero-interface could be presented in one of two approximations: the triangular well approximation and the square well approximation fig.(II.1). The first type could be applied in the case of conventional AlGaAs/GaAs HEMT's where no energy barrier at the buffer side is introduced. The solution of Schrodinger wave equation gives for this type:

$$\varepsilon_l = \left(\frac{\hbar^2}{2m}\right)^{1/3} \left(\frac{3\pi}{2} qE_s\right)^{2/3} (l + 3/4)^{2/3} where\, l = 0, 1, 2, 3, ... \tag{II.34}$$

and l stands for the sub-band.

Now if we consider the first subband $l = 0$ we have

$$\varepsilon_0 = 1.83 * 10^{-6} E_s^{2/3} \tag{II.35}$$

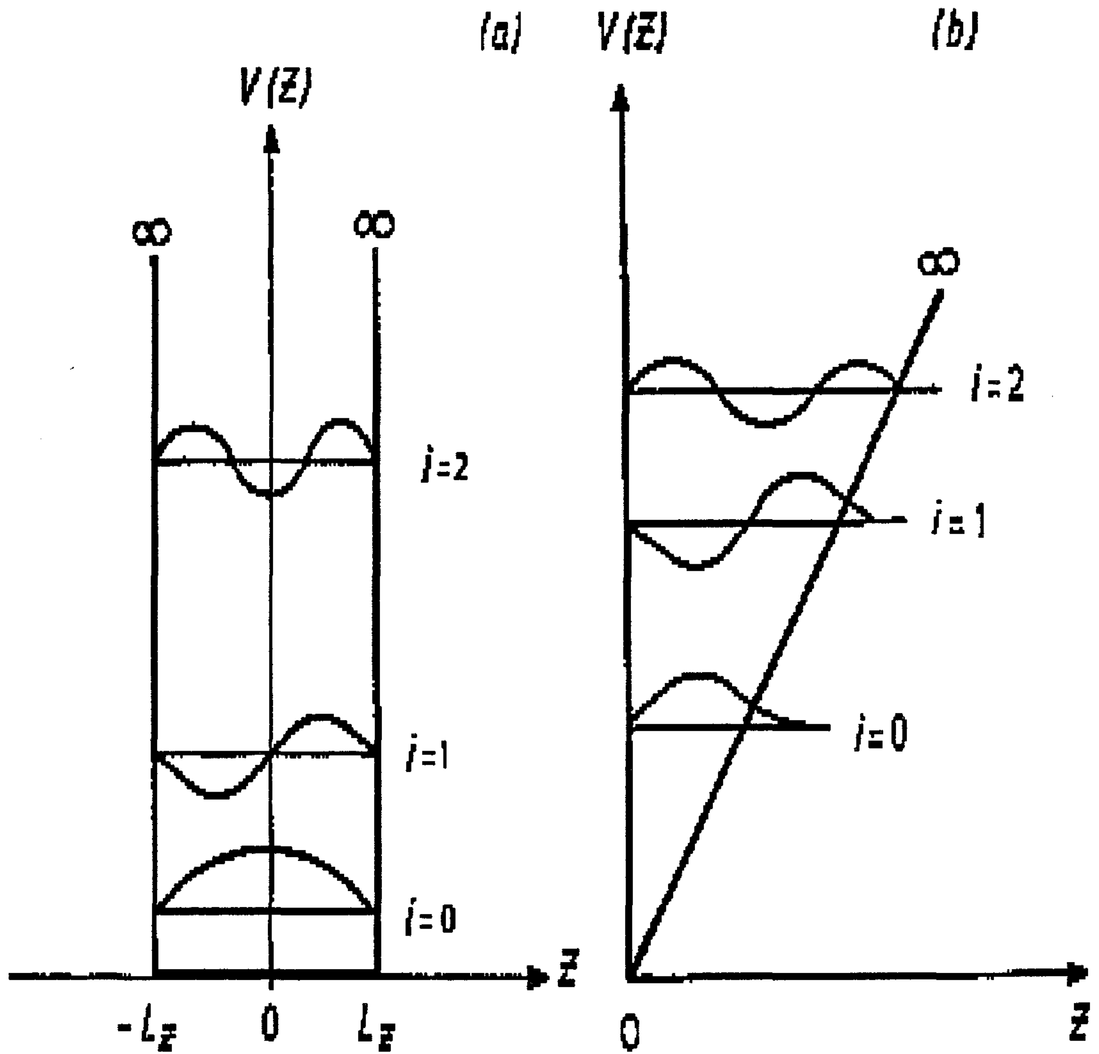


Figure II.1: Approximations used for the simulation of quantum wells: Triangular well approximation and the square well approximation

Where E_s is the electric field at the interface including the junction equivalent field.

As for the square well approximation, this could be applied in the cases where we have a single quantum well structure. Typical example would be the pseudomorphic AlGaAs/InGaAs with AlGaAs buffer. For this type we have the subband energy as

$$\varepsilon_l = \frac{\hbar^2\pi^2}{8m^*L^2}(l+1)^2 where\, l = 0,1,2,3,... \tag{II.36}$$

which for the first subband $i = 0$ gives:

$$\varepsilon_0 = \frac{(\hbar\pi)^2}{8m^*L^2} \tag{II.37}$$

And L is the width of the well.

We then modify the second term in the RHS of (II.32) to be:

$$\frac{\xi - \xi_0 + \delta\xi}{\tau_\xi(\xi)} \tag{II.38}$$

where $\delta\xi$ is calculated as in (II.35) or in (II.37). As a matter of fact, this means modifying the potential energy of carriers by the amount equivalent to the first subband.

We tested the quantisation correction on a 0.3μm gate pulse doped pseudomorphic AlGaAs/InGaAs/GaAs HEMT with a doping level of $3 * 10^{12} cm^{-2}$. The well width is $140\AA$ with 20% Indium; the barrier layer width was $150\AA$ with 20% Aluminium. We observe from fig.(II.2) a reduction in the output current. The difference exists over the whole range of V_{gs}, it is mainly pronounced for positive gate bias. This is related to the increase in the average total energy through the added potential energy resulting in reduced average velocity.

II.8.2. Simulation of the screening effect

As mentioned in chapter I, the screening results in establishing a dependence of the low field mobility of the 2DEG and the charge density in the channel. Previously, it

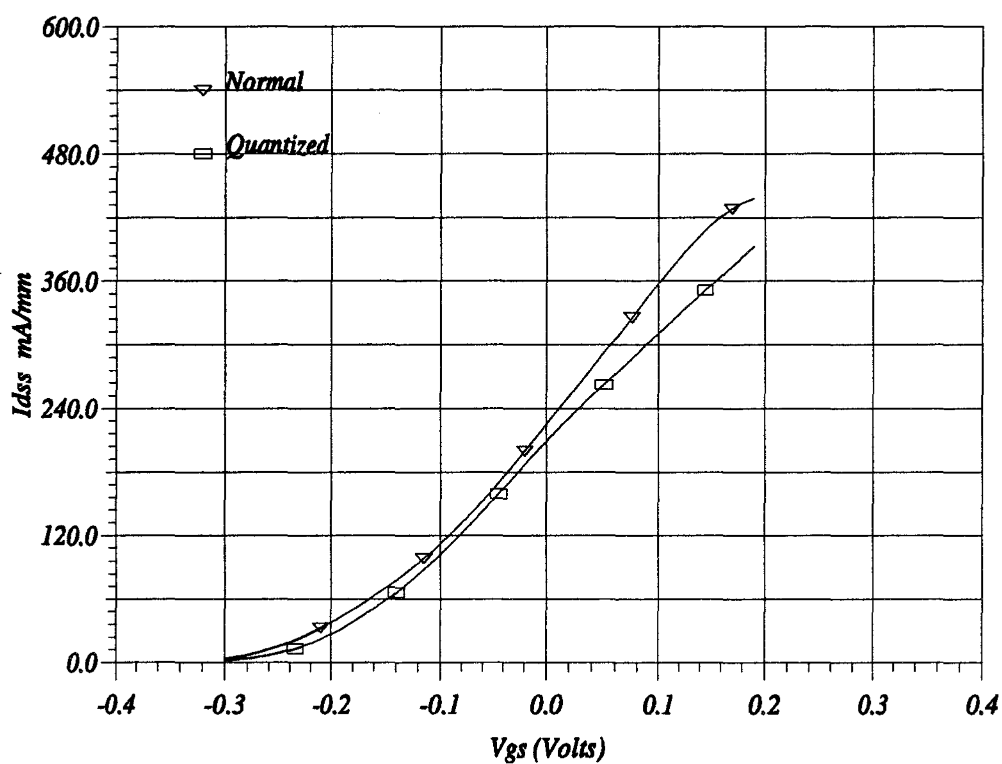


Figure II.2: Effect of energy quantification on the drain current in a pulse doped pseudomorphic AlGaAs/InGaAs device with $3 * 10^{12} cm-2$ *doping level and* $0.3 \mu m$ *gate*

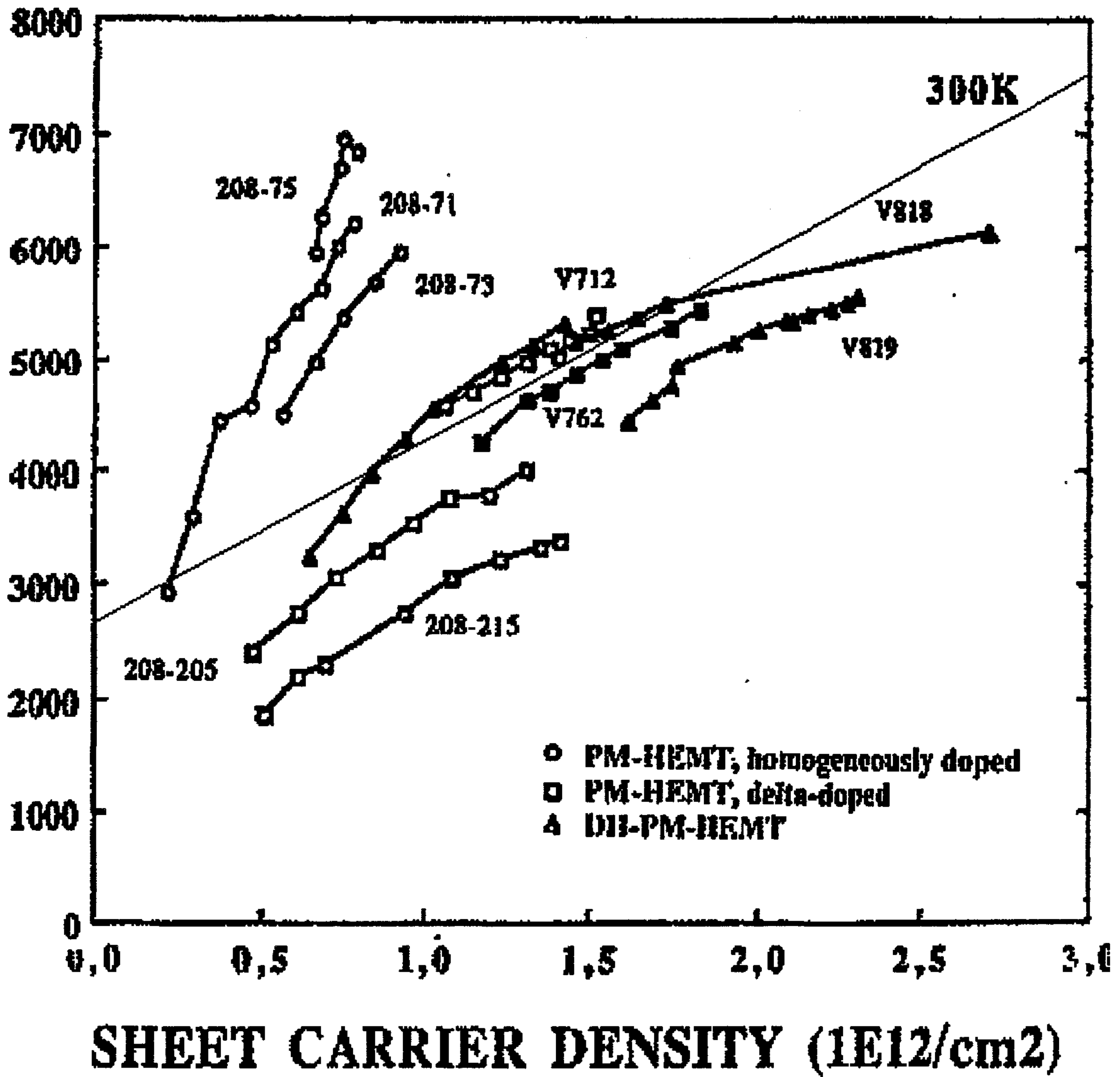


Figure II.3: The variation of n_s with the low field mobility, the fine dotted line gives the fitting used to simulate the phenomena in a first approach.

was assumed that the 2DEG low field mobility is that of the undoped bulk layer in which it is situated; this is only true upon the complete formation of 2DEG, i.e. $n_s = n_s max$ as was demonstrated by experimental measurements presented in chapter one. Based on these experimental measurements we introduced an approximate variation of $\mu_{l.f.}$ with n_s that would fit to most of the devices characterised. This fit is shown by the fine dotted line in fig.(II.3)

The dependence of $\mu_{l.f}$ on n_s could then be expressed as

$$\mu_{fl} = An_s + B \tag{II.39}$$

for the fit of fig.(II.3), $A = 1.6x10^{-9}$, $B = 2700$.

The n_s is calculated as:

$$n_s(x) = \int_L ndl \tag{II.40}$$

where n is the volumetric charge density and L is the width of the quantum well.

Accordingly, in the physical model, the mobility is not only a function of the average total energy of the electrons, but in the quantum well, it will be a function of 2DEG density n_s as well.

$$\mu(\xi, n_s) \tag{II.41}$$

Elsewhere, μ_{lf} has indeed nothing to do with n_s.

The effect of such a correction is of great impact, fig.(II.4), on the predictions of the simulator. The intrinsic current gain cut-off frequency, which is a sensitive parameter to calculate as it combines all the other parameters, shows a drop of about 25% to remarkably fit the experimental results found.

II.8.3. Simulation of avalanche multiplication

The avalanche multiplication phenomenon has an important effect specially in high fields, the onset of the avalanche multiplication results in sudden increase in the device current limiting it use due to the occurrence of breakdown.

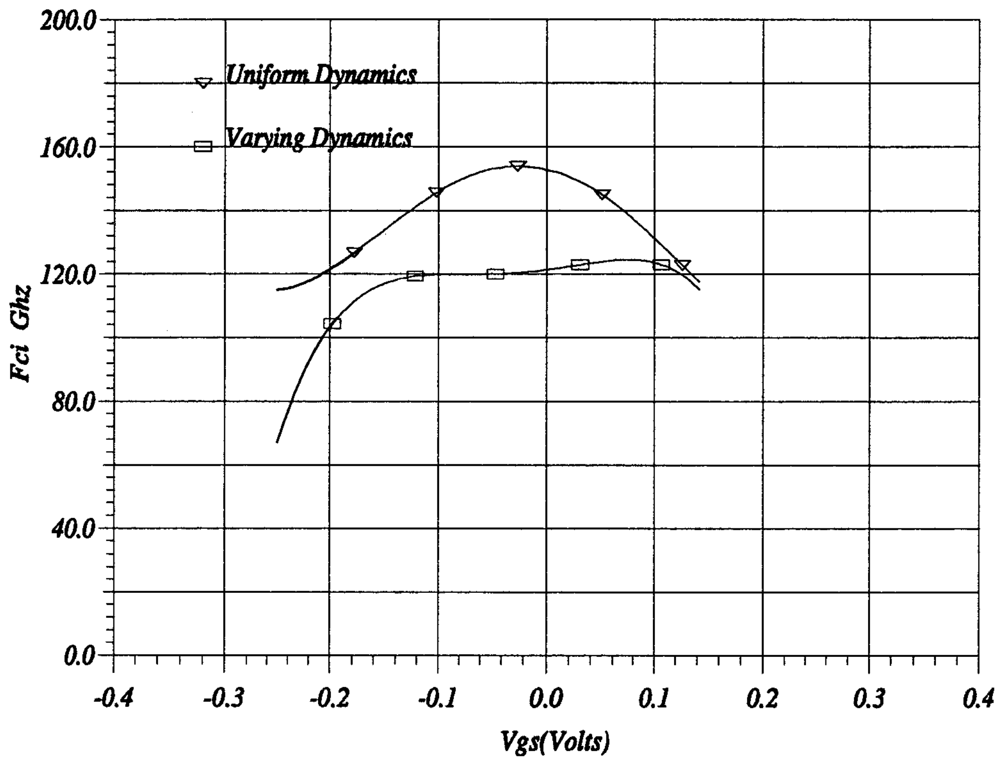


Figure II.4: Effect of 2D electron gas screening on f_c in a pulse doped pseudomorphic AlGaAs/InGaAs device with $3 * 10^{12} cm{-2}$ doping level and $0.3 \mu m$ gate

As a first approach to include the effect of this phenomenon, we retained the generation term G, which was neglected in the continuity equation. This quantity corresponds to the number of electron hole pairs generated per unit time. It is obvious that the net generation rate will be the sum of the pairs generated by an electron initiated ionisation and those generated by a hole initiated ionisation. Thus the net generation rate could be calculated as :

$$G = G_n + G_p \tag{II.42}$$

which could be resolved to :

$$G = \alpha_n \cdot \left|\frac{J_n}{q}\right| + \alpha_p \cdot \left|\frac{J_p}{q}\right| \tag{II.43}$$

where α_n and α_p are the electrons and holes ionisation coefficients respectively and J_n and J_p are the electrons and holes current density respectively. The ionisation coefficient is a quantity giving the number of electron-hole pairs generated in a unit distance (cm^{-1}). This quantity depends mainly upon the distribution function of electrons (holes) in the conduction (valence) band which will obviously be the most complicated task in the precise calculation of the ionisation coefficient.

In calculating the ionisation coefficient, several theories were developed which attempted to find a form to express the impact ionisation coefficient as a function of the electric field so as to satisfy the physical reasoning and to fit the experimental data as well.

The general form for the ionisation coefficient is given as:

$$\alpha = Aexp\left(\frac{B}{E}\right)^m \tag{II.44}$$

where A, B and m are material dependant and E is the local electric field.

The elementary theory for breakdown given by the precedent formula was obtained under uniform electric field conditions to which the sample was subjected. In the submicronic structures, this could be a wild assumption. The electric field is non

	A	B	m
GaAs	$2.994 * 10^5$	$6.848 * 10^5$	1.6
AlGaAs	$2.21 * 10^5$	$7.64 * 10^5$	2

Table II.1: Material parameters for calculating ionisation coefficient in GaAs and AlGaAs with 30% Al

uniform. In fact, an electron found in an electric field, although very high, will not be able to ionise unless its energy is superior to a certain minimum, (c.f. Chapter I). Thus, the electron has to travel a certain distance, known as the black space[51], in the high field region, before it is able to ionise. This is taken into consideration by assuming the ionisation coefficient as a function of the average total energy of the electrons population. To get this dependence, two steps of calculations are carried:

1. The dependence of the electric field on the average total energy is calculated through the tables presented by Monte-Carlo simulations. In our laboratory, such simulations were made to account for high field transport in the conduction band[52]

2. Once the energy dependant electric field is known, the ionisation coefficient is calculated using equation(II.44)

For the ionisation coefficient of GaAs and AlGaAs, we used the parameters presented in table (II.1). The variation of α with ξ is presented in fig.(II.5), we could notice that the onset of ionisation is at about 0.6ev. For energies higher than 1.ev, the ionisation coefficient in the GaAs becomes higher than in the AlGaAs, this value of average energy occurs at drain voltage as low as 3V, so in the average range of Vds, the ionisation in GaAs will be dominant. This confirms with the experimental observations published[53].

The electrons conservation equation is then modified to retain its original form as:

$$\frac{\partial n}{\partial t} + \nabla(nv) = \alpha_n(nv_n) + \alpha_p(pv_p) \tag{II.45}$$

As for the total energy variation, we assume that the impact ionisation does not change the distribution of the average total energy and we neglect a reduction term in the form of $Gdx\xi$.

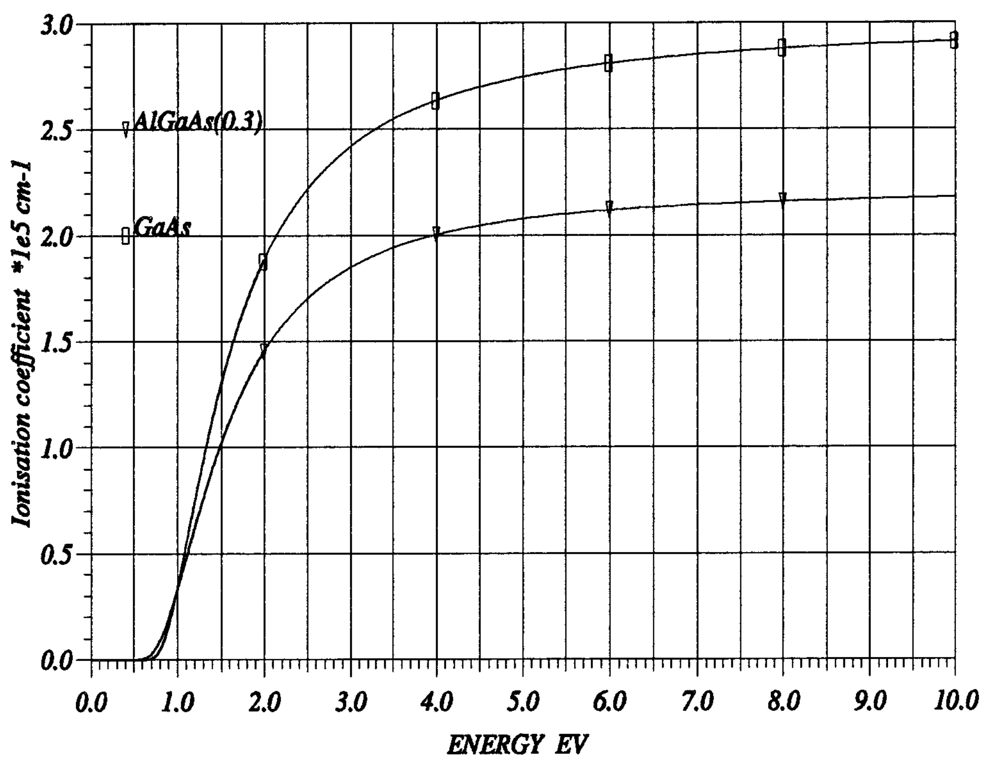


Figure II.5: Variation of the ionisation coefficient with energy

An indispensable information concerning the minority carriers charge concentration and velocity would indicate the inclusion of these latest in the model. This is discussed in the following section.

II.8.4. Simulation of minority carriers transport

The concept of Boltzmann Transport Equation holds for both types of carriers, electrons and holes. In the case of unipolar devices, the minority carriers contribution to the total current is negligible under normal operating conditions and thus the minority current could be calculated in an approximate manner. The set of conservation equations could be used to describe minority charges provided that the appropriate set of physical parameters is used. There exist however, a number of points to consider in simulating the minority charges specially in HEMT's where they are holes in most of the cases.

The effective mass of holes is much greater than that of electrons and hence their mobility is far below that of the electrons. In addition we do not expect to see phenomena like negative differential mobility. The form of the velocity field characteristics would then be of the silicon like type. Being heavier than electrons, the hot holes dynamics would not be of the same weight in the carrier transport as the hot electrons dynamics. We find it appropriate then to consider the holes problem as that of a drift-diffusion type. This means that the average total energy will be considered as an instantaneous function of the electric field, this would lead us to consider the physical parameters to be varying as a function of this electric field. This goes for the mobility in the first place. The second consequence is to assume that the diffusion coefficient is independent of the energy implying that for the diffusion current term, the carriers are cold ($T = T_0$). The diffusion term is thus a function of the electric field on its turn.

II.8.5. The complete physical model

The physical model used is thus an enhanced version of old one. This model is suitable for simulating heterostructure field effect transistors under avalanche multiplication condition as well as structures with pseudomorphic channels. The equations constituting this model assuming positive electrons are :

1. **Poisson's Equation:**

$$\nabla^2 V = \frac{q}{\epsilon}(n - N_d - p + N_a) \tag{II.46}$$

2. **Electrons charge density conservation equation:**

$$\frac{\partial n}{\partial t} = -\nabla(nv_n) + G_n + G_p \tag{II.47}$$

3. **Holes charge density conservation equation:**

$$\frac{\partial p}{\partial t} = -\nabla(pv_p) + G_n + G_p \tag{II.48}$$

4. **Electrons current density:**

$$J_n = \mu_n(\xi)[n\vec{E} + \nabla(nK_B T_e(\xi)] \tag{II.49}$$

5. **Holes current density:**

$$J_p = \mu_p(E)[p\vec{E} + \nabla(pK_B T_0)] \tag{II.50}$$

6. **Electrons average energy conservation equation:**

$$\frac{\partial \xi}{\partial t} + v\nabla\xi + \nabla\cdot(vKT_e(\xi)) = qv\cdot\vec{E} - \frac{\xi_q}{\tau_\xi(\xi)} \tag{II.51}$$

Where ξ_q is the modified rest energy accounting for carriers quantization and given as:

$$\xi_q = \bar{\xi} - \xi_0 + \delta\xi \tag{II.52}$$

Physical parameters functional dependence

The different parameters, as was mentioned before, are assumed to be dependant on the average total energy of majority carriers (electrons).

These parameters are : The electrons mobility, electrons ionisation coefficient, electrons average energy relaxation time and the electronic temperature. This later taken as a function of the average total electrons energy represents a slight correction to the assumption of having a displaced Maxwellian distribution which, as we have seen before, is sort of a hard assumption.

For the holes, the mobility and the ionisation coefficient are taken a function of the local electric field.

II.8.6. Main Characteristics of the physical model

The major characteristics of the presented physical model could be summarised as :

1. It is a transient time model. This permits to have information concerning the device transient response as well as the device dynamic parameters, e.g. τ. This is done by applying an FFT to the output drain current function of time after neglecting the first few femto seconds to avoid the zone here the inertial effects might be effective.
2. The model accounts for the hot electron dynamics by including the energy balance equation and then taking the physical parameters as a function of the energy.
3. The minority carriers effect is taken into account through the introduction of a drift diffusion transport model.
4. The model accounts for carriers multiplication through impact ionisation in an accurate way by taking the ionisation coefficient for hot electrons as a function of energy.
5. The model accounts for a number of physical phenomena occurring in HEMT's:

- Deep-level traps in heavily doped high aluminium content layers. This phenomena is behind the loss in current level at cryogenic operation of conventional HEMT's. It is simulated by introducing an equivalent deep DX-centre traps in the form:

$$N_D^+ = \frac{1}{1 + 2^{(E_{fn} - E_d)/K_B T_e \xi}} \tag{II.53}$$

  with E_d at 160mev below the conduction-band L minimum.

- Surface degradation leading to the presence of traps. This is simulated by introducing an equivalent depletion surface field in the form:

$$E_s = \left(\frac{2qN_D V_s}{\epsilon}\right)^{1/2} \tag{II.54}$$

  with V_s as the surface potential.

- Carriers quantization in the quantum well which is a phenomena occurring in heterostructure FET's responsible for electron heating.

- Variation of low field mobility with interface charge density in the well (screening effect)

II.9. Numerical approach

The set of partial differential equations presented and which constitute the hydrodynamic energy model is, sort of, impossible to solve in a closed form. A numerical technique is needed to render the problem solvable.

In general, the numerical techniques aim at transforming the domain of the original problem to a set of smaller domains as a first step. Next, the set of partial differential equations is transformed to a system of difference equations through the discretisation technique over the sub domains. The problem is then transformed to a mathematical one consisting in solving systems of non-linear algebraic difference equations.

There exist three major techniques for the process of domain division. The most famous and the most widely used in the field of device simulation is the *finite difference* . The second technique is the finite boxes scheme which a modification of

the finite difference one. The third is the finite element scheme which is based on triangular elemental sub domains. This technique is the most suitable for irregular structures, such as passive microwave devices, wave guides etc.. The use of finite element method for semiconductor device analysis[54] has been atepted since long time. The finite element method is of particular interest for irregular sahpes simulation. This could be of use in treating devices where the gate for example, is included thus permitting the study of the effect of gate shape. This permits also the efficient use of adaptive mesh techniques and thorough analysis of the phenomena occurring in recess corners.

In our work we adopted the finite difference technique. This choice is based upon the simplicity of structure for the set of differential equations discretised in finite difference, also the flexibility of modifying the lateral device structure with minimum complications in the mesh design. A major advantage was also the possibility of using the numerical techniques previously adapted for finite difference method.

The domain of the problem is partitioned into a set of rectangular sub domains through the introduction of a non uniform mapping mesh or grid. The technique is applicable for sure in three dimensional space analysis, however we restrict our discussion here to the two dimensional space problem, the generalisation is straight forward.

First of all we had to define two sets of variables:

1. Scalar variables include: Charge density, potential, average energy, mobility (holes and electrons), effective mass, electronic temperature, ionisation coefficients.
2. Direction dependant quantities include: Electric field, Electrons velocity, holes velocity, electrons current and holes current. For these quantities we have two sets per variable defined in the two directions x and y.

In the finite difference scheme, the scalar quantities are defined on the main mesh point, fig. (II.6). The direction dependant quantities are defined on the interme-

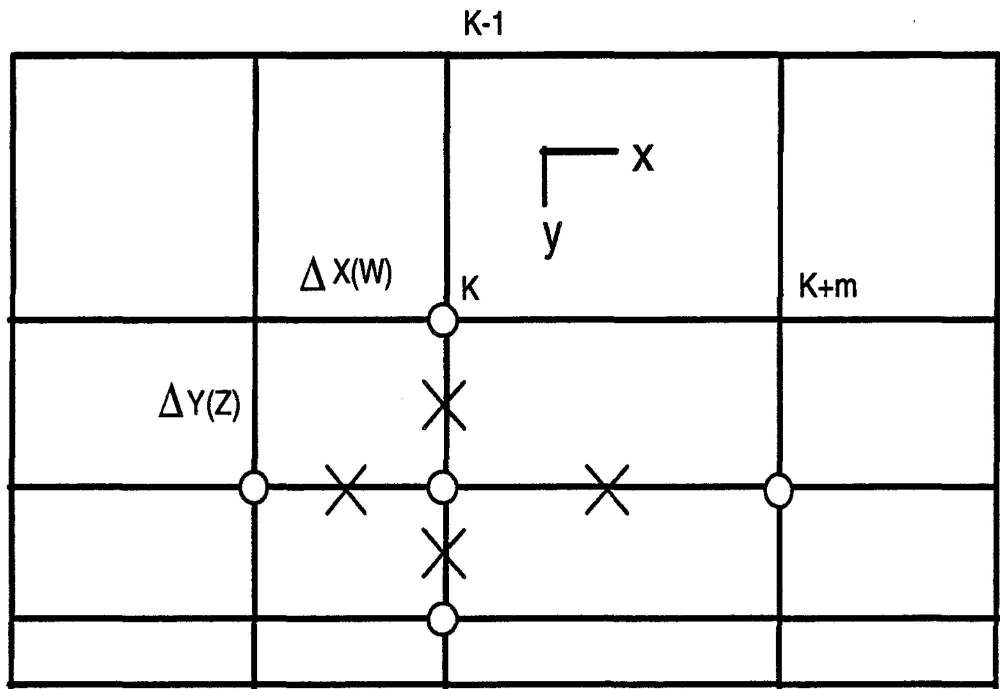


Figure II.6: The non-uniform mesh nomenclature adopted in SIMFET II. Scalar quantities are defined on the main node points (O), vectorial quantities are defined on the secondary points (X)

diate points. A five point formulation then follows to produce the set of non-linear difference equations.

II.9.1. Spatial discretisation

In the finite difference scheme employed, a highly non-uniform mesh is used. The minimum mesh spacing in both direction has to be in the order of the Debye length given as :

$$L_D = \sqrt{\frac{\epsilon KT}{q^2 N_m}} \tag{II.55}$$

where N_m is the maximum local charge concentration.

At the same time, the mesh has to be of fine resolution while keeping the number of points as small as possible. These conditions are not easy to satisfy so we put a set of design rules for the mesh based on the gradient of the charge density. This meant imposing the minimum mesh size in the direction perpendicular to the heterojunction and in the quantum well to be strictly equal to L_D since the maximum charge density gradient is to be found there, with a maximum of not more than twice this quantity. In the other direction, and at the gate exit, the zone of eventual accumulation, the minimum step is to be twice L_D and goes as high as 4 times this quantity at the source side. Obviously this imposes the use of a non uniform mesh.

II.9.2. Temporal Discretisation

The temporal discretisation of time derivatives is done using an Euler fully implicit backward formula of the first order. This formulation guarantees stability without going down with the time step. The discretised function at a time instant t^k reads :

$$\vec{F}(u^k) = \frac{u^k - u^{k-1}}{t^k - t^{k-1}} \tag{II.56}$$

A non uniform dynamic time step is used which allows for an automatic setting of the appropriate value. This is going to be discussed later in more details.

II.9.3. Boundary Conditions

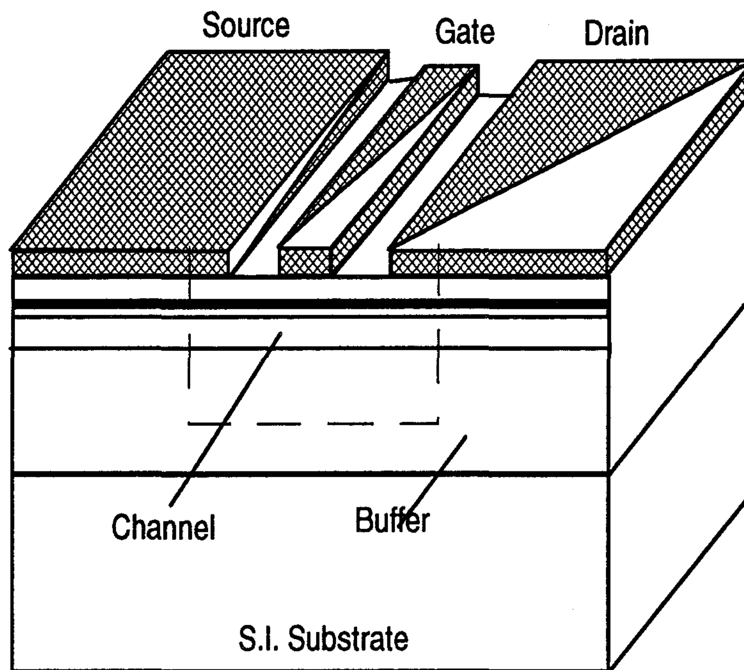


Figure II.7: A schematic presentation of an actual device, the considered portion for the simulation is the one bounded by the dashed line. The rest of the electrodes is considered as a parasitic resistance, negligible current is assumed to be passing through the imaginary (dashed) boundaries

When we consider real devices, we find that the active region of the device could be extracted as a small portion, fig.(II.7) of the original structure. This region could be seen as bounded by artificial (imaginary) boundaries in a way to guarantee the simulation of the device in a self contained manner. These boundaries are also

to guarantee that no current leaves the device through them. As for the ohmic contacts, a small segment is considered while the rest could be represented as an access resistance. This is of particular interest since the typical sizes of the electrodes are too large with respect to the active zone of the device.

This sums up the problem of boundaries into three types: Ohmic boundaries where the carriers should essentially be cold and the electric field is sufficiently low, Schottky contacts, and artificial boundaries.

Ohmic contacts

There we apply Dirichlet boundary conditions. For the electrostatic potential, this means that on the source we have:

$$\psi(t) = 0 \tag{II.57}$$

where we take the source as reference.

For the drain we have :

$$\psi(t) = \psi_D(t) \tag{II.58}$$

The charge densities also submitted to the Dirichlet boundary conditions on the source and drain we set the charge densities n, p as:

$$n = N_d \tag{II.59}$$

for electrons and

$$p = N_a \tag{II.60}$$

for holes.

For the average electrons energy at the source electrode we apply a similar boundary condition:

$$\xi = \xi_0 \tag{II.61}$$

where ξ_0 is given as :

$$\xi_0 = \frac{3}{2}\frac{K_B T_0}{q} \tag{II.62}$$

This is totally acceptable as there are no hot carriers at the source side of the device.

At the drain electrode the situation is different, the energy of electrons is left floating and updated normally with the energy balance equation. Normally the average total energy of electrons under the drain is low since the carriers are cooled down through two mechanisms; the first while traversing the energy barrier into the barrier layer and the second while traversing this layer itself in the process of being collected by the drain. Throughout the simulations, we verify that the carriers are cool at the drain electrode.

The Schottky Contact

The problem of the Schottky contact is a crucial one since the physics of Schottky metal-semiconductor contacts are rather complicated. In a general way, the charge density on a Schottky barrier depends upon the amount of current passing in it which in turn depends on the external applied bias. Thus we can say that there exist a certain charge concentration given as:

$$n_0 = N_c exp(\frac{E_{fn} - E_c}{KT_0}) \tag{II.63}$$

where N_c is the effective density of states in the conduction band, E_{fn} is the quasi Fermi level.

The thermionic diffusion theory gives the current across the junction as:

$$\vec{J_n} \bullet \vec{n} = -qv_r(n - n_0) \tag{II.64}$$

where $\vec{n}$ is the unit vector perpendicular to the interface and v_r is the surface recombination velocity.

In most practical cases, specially for pulse doped devices, it is acceptable to assume that electrons current across the gate is almost zero within a specfied range of gate bias. This, in fact is an approximative hypothesis that represents a limitation to the use of this model, currently this point is subject to extensive work to overcome such limitation. So for the electrons we have:

$$\vec{J_n} \bullet \vec{n} = 0 \tag{II.65}$$

The question for the holes is a little bit different; in order to allow for the holes to flow, we impose a condition on the charge concentration in a way to neglect the diffusion portion of the current. This is acceptable since at the gate the electric field is extremely high.

As for the potential boundary condition, again a Dirichlet B.C. is applied :

$$\psi(t) - \psi_b = const \tag{II.66}$$

where here ψb is a value combining the Schottky barrier high and the built in potential.

The average electrons energy at the gate is set to ξ_0 since the electrons are practically cool in the vicinity of the gate.

Artificial Boundaries

The artificial boundaries must be in such a way to guarantee that no current passes through them. Also the confinement of the carriers inside the device is to be ensured. To do this we apply the Neumann boundary conditions :

$$\frac{\partial \phi}{\partial \vec{n}} = 0,\ \frac{\partial \xi}{\partial \vec{n}} = 0\ J_n \bullet \vec{n} = 0 \tag{II.67}$$

Generally speaking, the choice of the position of the artificial boundaries must be based on approximate knowledge about the carrier distribution. This is particularly important in the case of the buffer layer thickness which should have a boundary far enough from the main transport region. For devices with strong buffer injection, the choice of the buffer thickness played a role in the estimation of the output conductance [36].

II.10. Computational Model

The resulting set of non-linear difference equations is partitioned over four distinct groups representing Poisson's, electrons continuity, holes continuity, and average electrons energy. To find a possible solution, the set have to be recasted into a

linear system of equations, accordingly a linearisation technique is applied using a modified Newton method [36].

Each set when arranged in a matrix or a vector form gives a standard equation in the form :

$$AX = B \tag{II.68}$$

where A is the coefficient matrix of dimensions (m * n) where m is the number of mesh points in the y direction and n is the number of mesh points in the x direction, and X is the unknowns vector of length n*m.

The system is again casted in a vectorial form so as to benefit from the inherent speed of vectorial processors, this requires special attention in the formulation of difference terms and boundary conditions. The linear system could then be solved by one of the following methods:

- Direct solution by elimination, this is very exact and unconditionally stable for all values of the time step, however it is very slow. It would be of interest then to use it to solve a fairly simple structure equation such as Poisson's equation. This is solved by Gauss elimination. The five diagonal matrix resulting is decomposed into L-U and they are calculated only once which saves a lot of CPU time. We use this technique to solve Poisson's Equation.
- The relaxation Method, used for solving the linearised conservation equations. The major inconvenience is that the convergence of the solution is not always guaranteed in contrast to the direct solves in addition to its sensitivity towards the time step used. On the other hand it is a simple structure method and quite well adapted for vector processing. We used this technique to solve the electrons conservation equation, holes conservation equation and the energy conservation equation.

II.10.1. The algorithm

The data flowchart is given in fig.(II.8). The decoupling of the equations is based upon the relaxation time approximation and the order of magnitude difference in the energy relaxation time and the momentum relaxation time.

This difference, particularly announced for electrons and holes permits to solve the continuity equation for holes *in parallel* with the equations of charge and energy conservation for electrons.

An interesting feature that we added to the computational model is the dynamic time step. The maximum time step that could be used in the case of iterative solution is determined by Reiser stability criterion and is given as :

$$\delta t < \frac{2\epsilon}{q\mu N_m} \tag{II.69}$$

The time step that could be actually used has shown to be far below this value specially at the start of the simulation. This was explained as due to the displacement currents existing and the theoretical time step is reachable in the absence of such currents. We thus set up a convergence monitoring procedure which will test if a diverging solution builds up in the iterative solver of conservation equations. Once a divergence is detected, the process is interrupted, time step is reduced and then the solution is attempted once more. In case of a successful solution, and provided that the displacement currents are below a determined value, the time step will be increased to accelerate the solution. The increase in the time step is in the scope of reaching the theoretical one calculated function of the doping level, temperature and low field mobility. In practical cases, the time step varied from 0.1 to 5 fsec according to the selection criteria mentioned.

Typical performance figures of the program are given in table (II.2), where a record execution time of 1cpu minute for 1 picosecond simulation time is found for conventional HEMT's. This time, normally, varies as a function of the structure complexity. It should be noted too that the use of a super computer, the IBM-3090-600E, is the main factor behind this performance. The execution times drop dramatically as

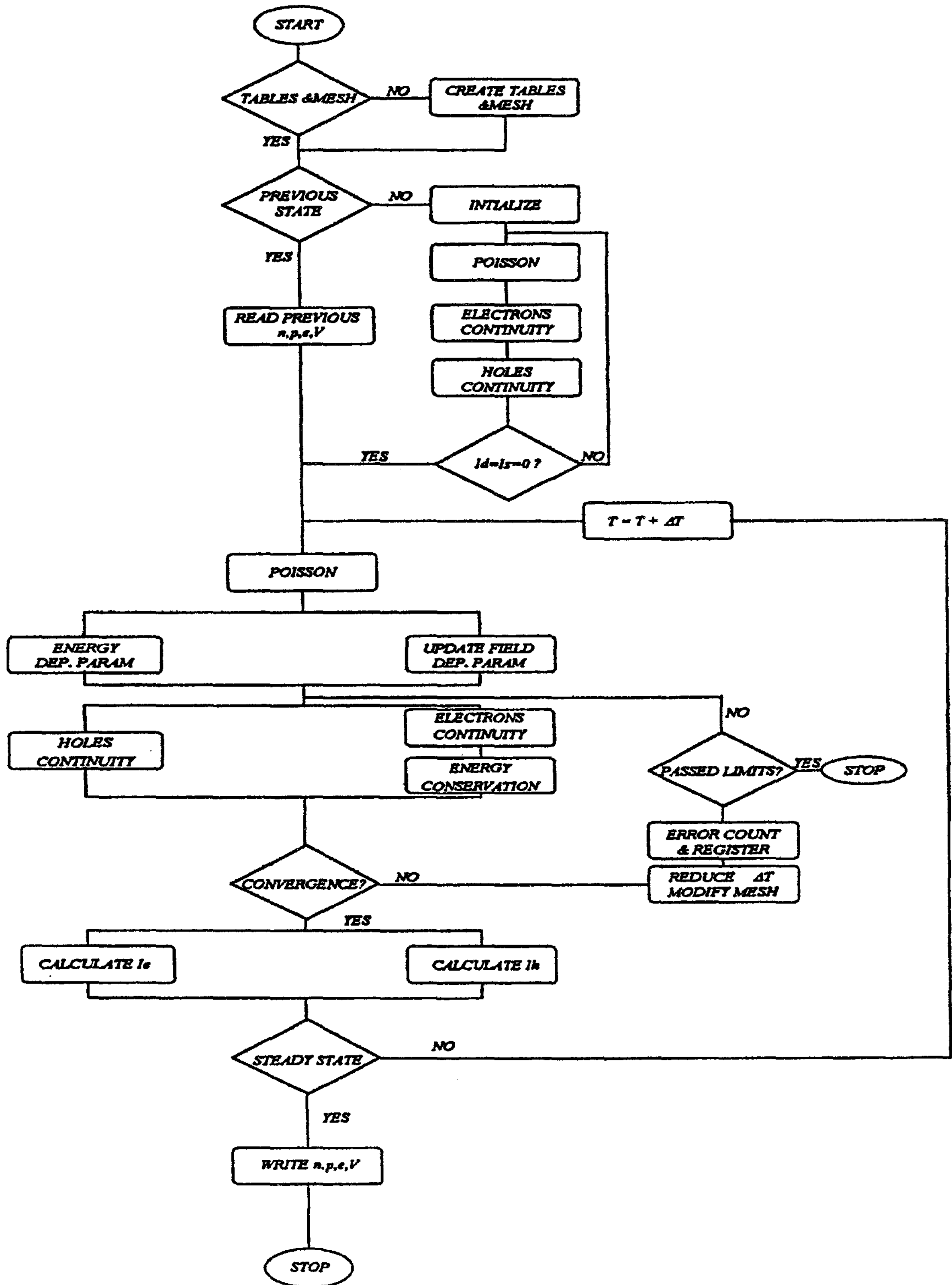


Figure II.8: SIMFET II data flow chart (version 3.0)

Device	Time (cpu.min/1psec)	Machine	Remarks
HEMT	1	IBM-3090	conventional, LN
HEMT	4	IBM-3090	Pseudomorphic UD
HEMT	6	IBM-3090	Pseudomorphic PD
HEMT	1-3	IBM-3090	Conventional BD
HEMT	40	HP workstation	Pseudomorphic UD

Table II.2: Typical performance figures of the simulator for different devices. LN=Low noise conditions, Vds < 3 V. UD=Uniform Doping. PD=Pulse Doping. BD=Breakdown conditions Vds > 5 V.

other machines are used and further effort should be given to improve the performance of the simulator on smaller machines.

This page is intentionally left blank

This page is intentionally left blank

This page is intentionally left blank

Chapter III

A study of Pseudomorphic HEMT's

Abstract

In this chapter the main features of pseudomorphic HEMT's operation are discussed. The distribution of physical parameters is presented under different operating conditions and then the circuit parameters are compared to that of the lattice matched conventional HEMT's. The effect of changing the doping profile to a pulse doped is discussed as well as the effect of the buffer type on terminal characteristics. We terminate this chapter by a discussion on the transient time instabilities manifested in HEMT's and more pronounced in pseudomorphic devices.

III.1. Introduction

We have seen in the first chapter that the pseudomorphic transistors represent the most promising devices for higher frequency in power applications. A possible way to improve their frequency response is the structural optimisation, however this is only possible after a full understanding of the device behaviour.

A systematic use of our program permits a thorough understanding of the device behaviour, this allows to compare the device performance to other structures and test the influence of certain material or structural parameters

III.2. Device Structure

A typical device structure is shown in fig. III.1. The number of parameters that have to be considered is extremely large, so in choosing some of them we stem from previous experimental and theoretical findings.

The choice of the Aluminium and Indium mole fractions is mainly governed by the desire to increase two major parameters: the n_s and the mobility. The increase in the n_s is achieved through the increase in the conduction band discontinuity which could be calculated by[55]

$$\Delta E_c = 0.9\left(\Delta E_{g1} - 2a\frac{C_{11} - C_{12}}{C_{11}}E - r\right) + 0.65\Delta E_{g0} \qquad \text{(III.1)}$$

where the first term represents the discontinuity between GaAs and strained InGaAs and the second term gives the conduction band discontinuity between AlGaAs and GaAs. In this formula a is the deformation potential, C_{11} and c_{12} are the elasticity coefficients, E_r is the lattice constant deformation coefficient, and ΔE_{g1} and ΔE_{g0} are the bandgap discontinuities.

<u>On the choice of In and Al contents</u>

Increasing the Al content in the AlGaAs will increase the band discontinuity with the InGaAs, but for Al content higher than 20 or 22%, deleterious effects related to DX-centres could become more evident. Accordingly most published work has taken this

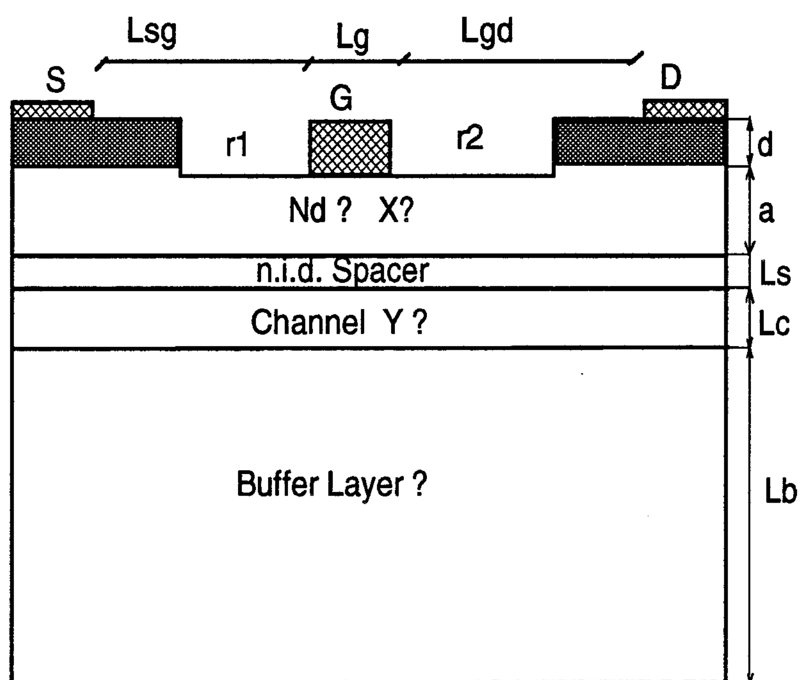


Figure III.1: A pseudomorphic AlGaAs/InGaAs/GaAs uniformly doped HEMT; there exist a plenty of adjustable paramteres

content to vary between 15 and 20%. Since, as we said before, we aim at maximising the n_s we took 20% Aluminium. The same argument is true for the Indium content, the higher the Indium content the higher conduction band discontinuity we get. It also permits to have higher low field mobility which is an increasing function of the In content, although this is questionable since we must not forget the role of the strain. And finally it will result in a reduced layer thickness, (fig.III.2). A good compromise for these contradicting demands is to take the In content as 20%, in the literature we found that the In ratio also varied typically from 15 to 20%[56].

Our first objective was to compare the performance of $0.3\mu m$ gate pseudomorphic AlGaAs/InGaAs/GaAs HEMT to a similar conventional lattice matched AlGaAs/GaAs, so we took the same structure that was considered optimal for this later device and we used the same doping level and barrier layer thickness.. This led to having a barrier layer thickness of $360\AA$ and a doping level of $1 * 10^{18} cm^{-3}$.

The stability criterion of Mathews [17], fig. III.2 imposes a thickness of not more than $150\AA$ for the InGaAs channel; in our case we took a layer thickness of $140\AA$ since n_s is an increasing function of the well thickness[57].

Mesh description

The Debye length in this case would be about $40\AA$, however the local charge concentration would be higher than the doping density in some regions, e.g. in the quantum well at the gate exit. We found that the local Debye length could be as small as $35\AA$ in some transients. The use of a mesh step in the order of the Debye length referred to the doping density could result is anomalies such as the non conservation of current in the device. This was a problem that we faced in the early stages of this work and we managed to solve it by using a minimum mesh step in the Y-direction (perpendicular to the heterojunction) of $10\AA$ and a maximum of $36\AA$. In the X-direction (direction parallel to the interface), we had the minimum of $20\AA$ at the gate exit which varied smoothly towards the drain with multiplication factor of 1.035. At the source side of the device we had the maximum mesh step of $80\AA$. The resulting mesh point number was finally 95 in the X, and 49 in the Y.

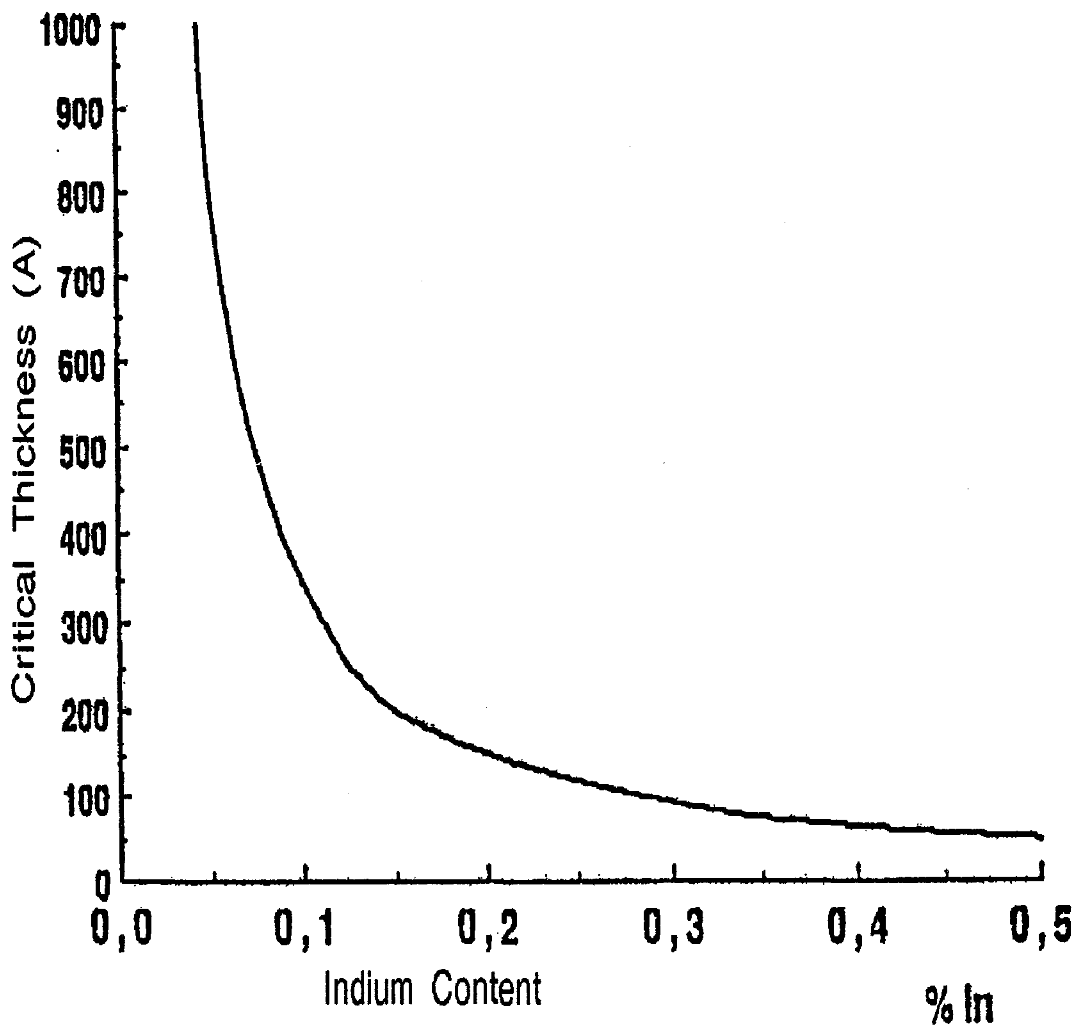


Figure III.2: The critical InGaAs layer thickness function of Indium content in a pseudomorphic AlGaAs/InGaAs/GaAs pseudomorphic HEMT.

Remark

It is important to mention that some of the simulations were carried out at the beginning of our work without the enhancements mentioned in chapter two. This gave the effect of overestimated values for the transconductance and the cut off frequency. However, since the objective was a comparative one the relative behaviour of the device is totally holding specially that the comparison was done under the same conditions.

III.3. The device physics: Distribution of main quantities

The device status could be totally described through the determination of three basic quantities : charge, average total energy and electrostatic potential distribution. A major advantage of hydrodynamic modelling is that it permits to have a good insight of the distribution of these quantities under different biasing conditions. The related physical quantities such as the electric field, electronic velocity, mobility, etc.. could be as well extracted based on the knowledge of these quantities.

We have plotted these quantities for different gate and drain bias; the figures are summarised below:

- V_{DS} = 2.5 V
  1. V_{gs}= 0 V...fig III.3
  2. V_{gs}= -0.3 V...fig III.5
  3. V_{gs}= 0.3 V...fig III.6
- V_{DS} =2.0 V and V_{gs}=0.3 V...fig. III.7

III.3.1. Charge Distribution

From fig. III.3, we could see the distribution of electrons charge density in the device simulated (first plot). The first observation is that there is much better confinement

of the carriers in the quantum well when compared to the conventional device case, fig. III.4.

The application of a positive gate bias results in adding a parallel conduction path in the AlGaAs layer as could be seen from fig. III.6 where the depleted zone under the gate has been reduced. This effect known as the parasitic MESFET is responsible for the compression in the g_m, the intrinsic transconductance, encountered in HEMT's at large positive gate bias. For power applications where linearity is needed (section I.7), it will be of importance to delay the appearance of such an effect.

III.3.2. Energy Distribution

Close examination of the second plot in fig. III.3 giving the average total energy contours reveals the existence of a high energy domain, which in this particular case is stationary at the gate exit. The formation of this domain could be viewed in two contexts. In a multi valley frame work, the electrons quantized in the sub bands of the central valley are heated up by the applied electric field. When the energy passes 0.3 e.v., which is approximately the separation energy between central and satellite valleys, the electrons will transfer to the heavier effective mass higher density of states satellite valleys if they encounter a favourable scattering event. This k space transfer results in having a larger number of slower carriers in the region of high field gradients at the gate exit with the consequence of having an accumulation region.

In the single equivalent gas picture, the variation of velocity with electric field will give the following reasoning. In the low field region under the gate, carriers travel with high drift velocity. Arriving with a high velocity at the gate exit, the carriers encounter a sharp rise in the electric field which results in a sudden drop in their velocity as they get very hot. The appearance of the accumulation region is thus a normal consequence. This accumulation-depletion layers are shown in fig. III.8 giving the distribution of the charge concentration along the channel (in the middle of the well) at $V_{gs} = 0.3V$ and $V_{DS} = 2.5V$.

Existence of hot carriers (0.65 ev) in the buffer layer is noticed when we examine

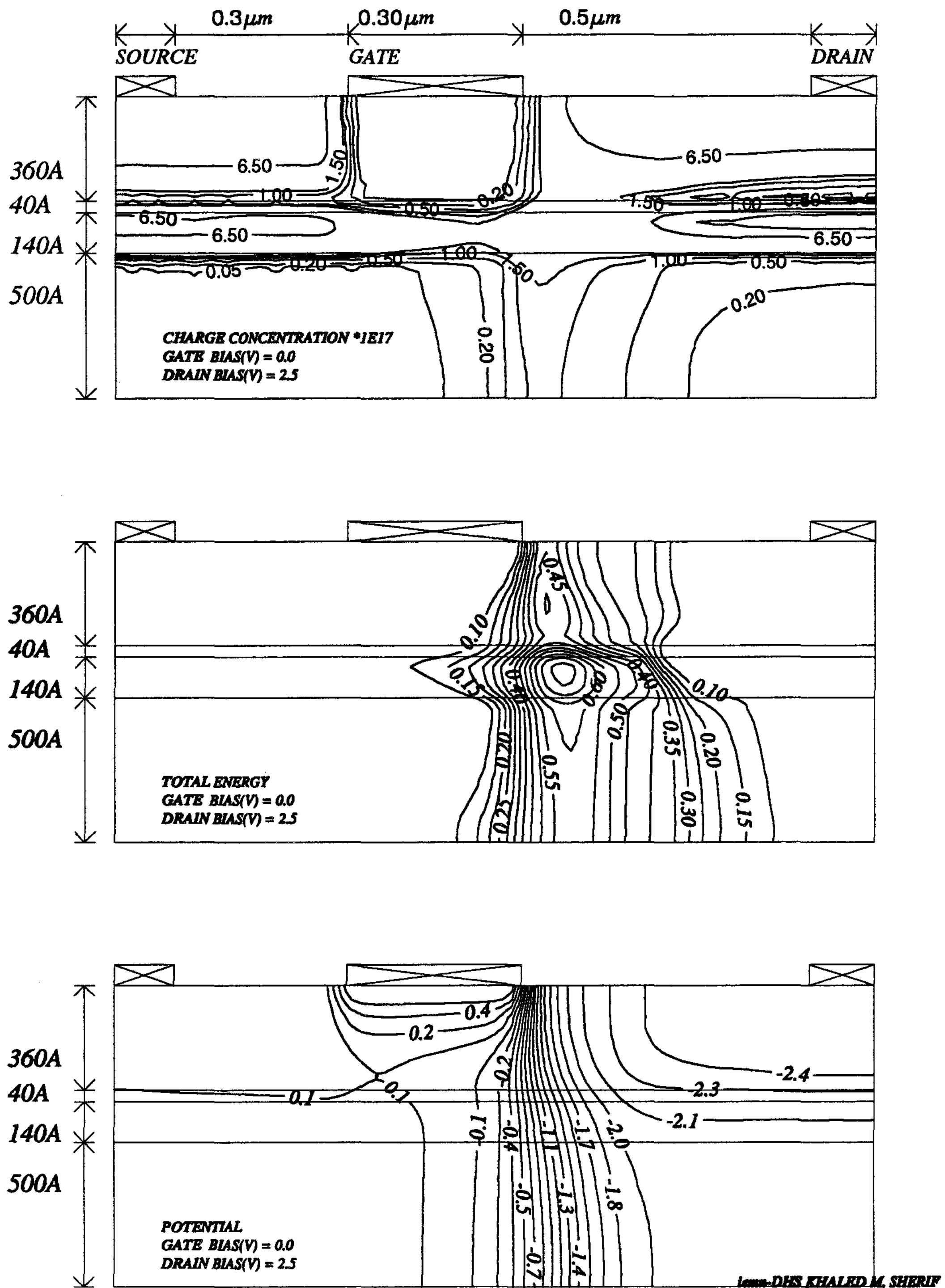


Figure III.3: Distribution of physical quantities: Electrons charge density, Average total energy, and electrostatic potential in a $0.3\mu m$ *gate pseudomorphic AlGaAs/InGaAs HEMT with uniform doping* $1 * 10^{18} cm^{-3}$ *over* $360\AA$.

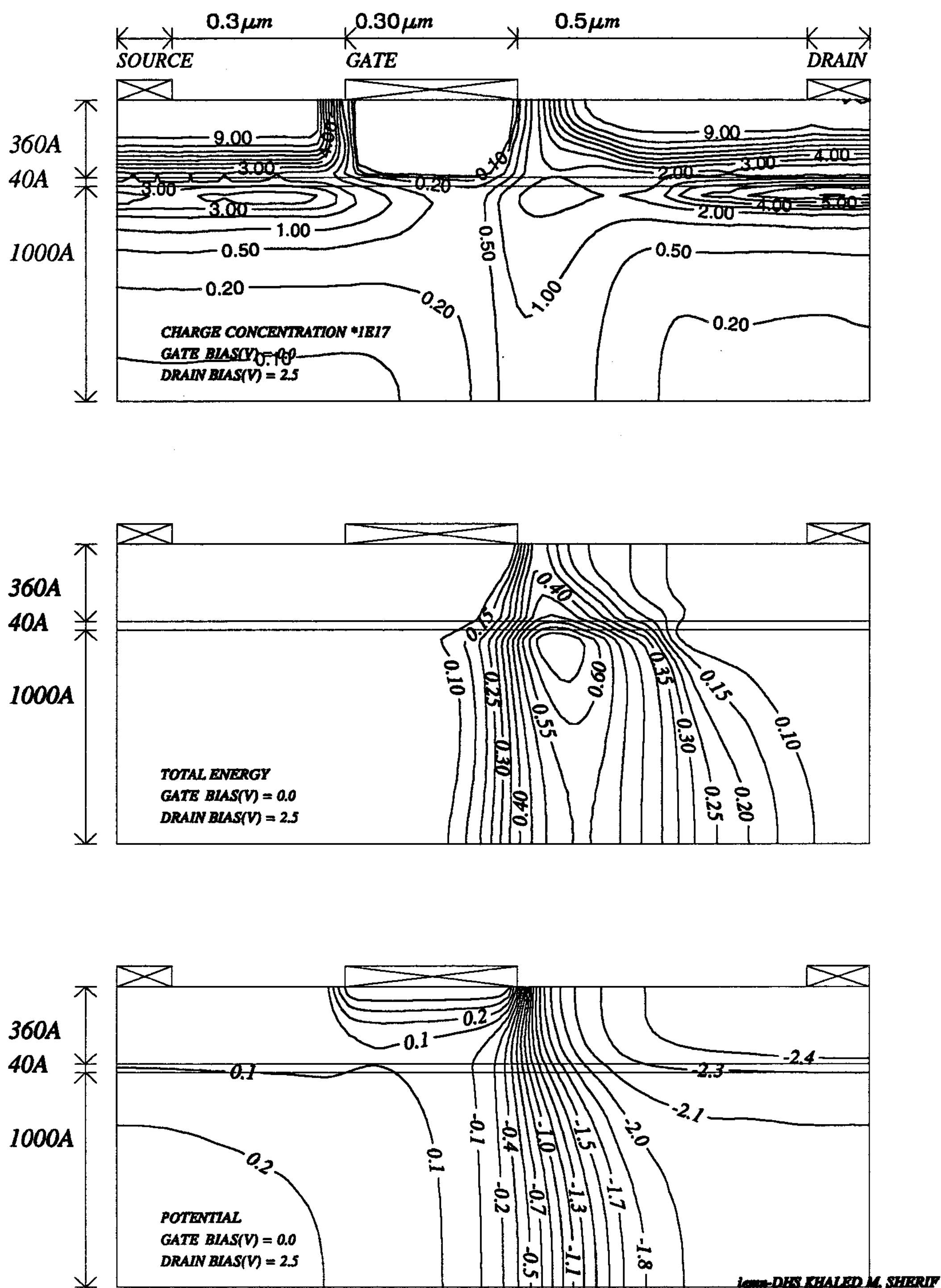


Figure III.4: Distribution of physical quantities: Electrons charge density, Average total energy, and electrostatic potential in a $0.3\mu m$ gate AlGaAs/GaAs conventional HEMT with uniform doping $1 * 10^{18} cm^{-3}$ over $360 \AA$

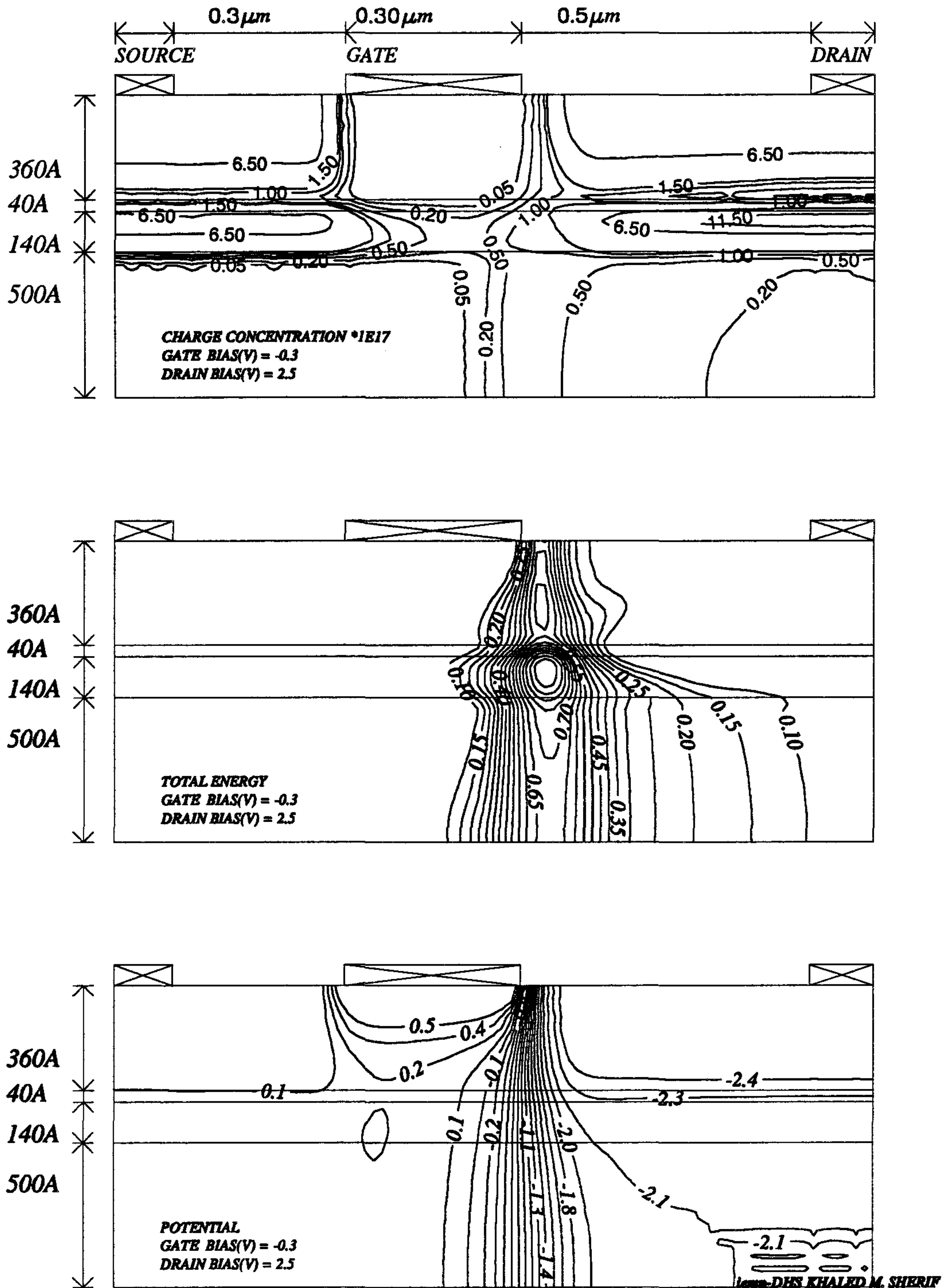


Figure III.5: Distribution of physical quantities: Electrons charge density, Average total energy, and electrostatic potential in a $0.3\mu m$ gate pseudomorphic $AlGaAs/InGaAs$ HEMT with uniform doping $1 * 10^{18} cm^{-3}$ over $360\AA$ near pinch off

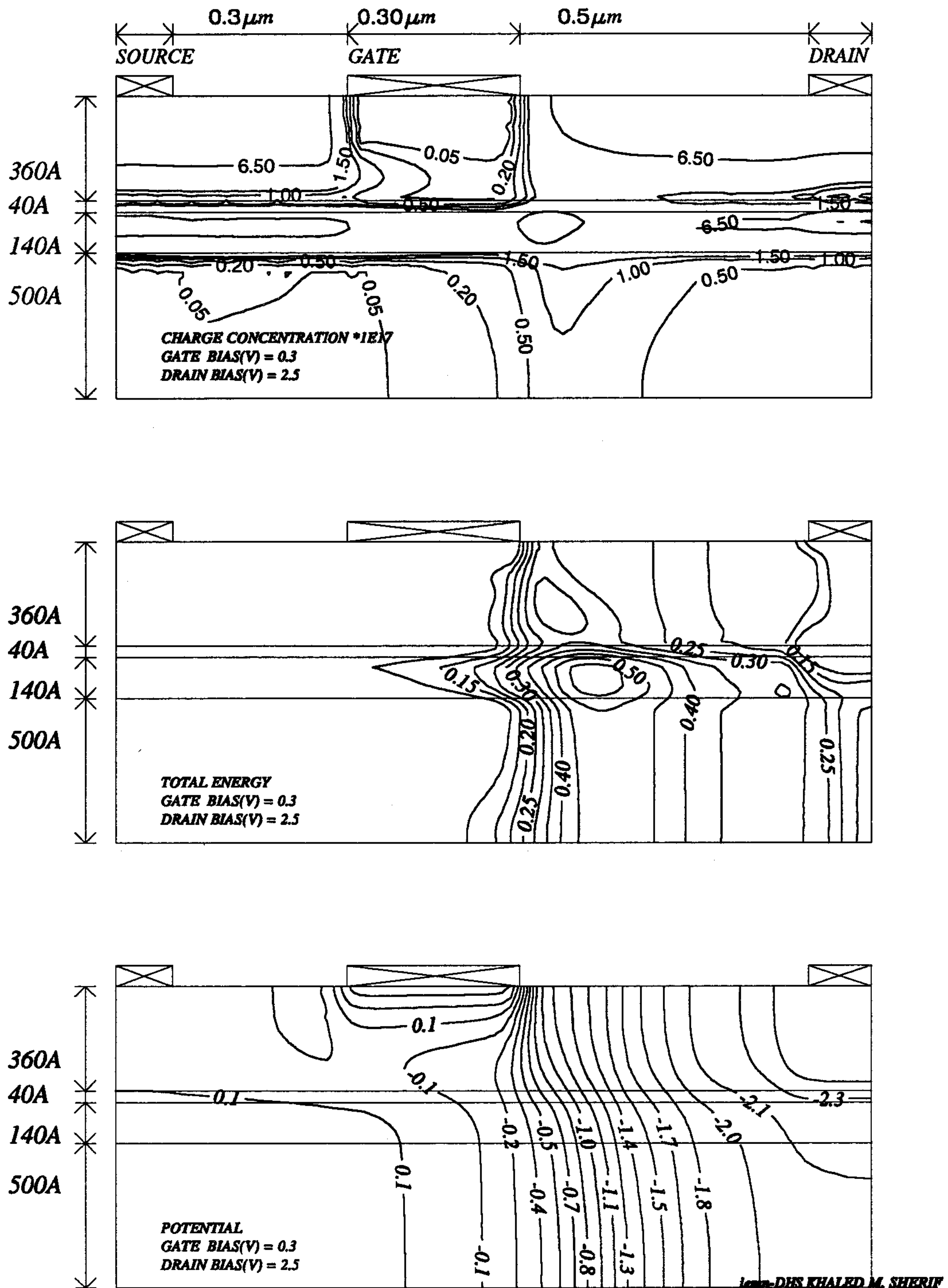


Figure III.6: Distribution of physical quantities: Electrons charge density, Average total energy, and electrostatic potential in a $0.3\mu m$ *gate pseudomorphic AlGaAs/InGaAs HEMT with uniform doping* $1 * 10^{18} cm^{-3}$ *over* $360\AA$ *in open channel operation*

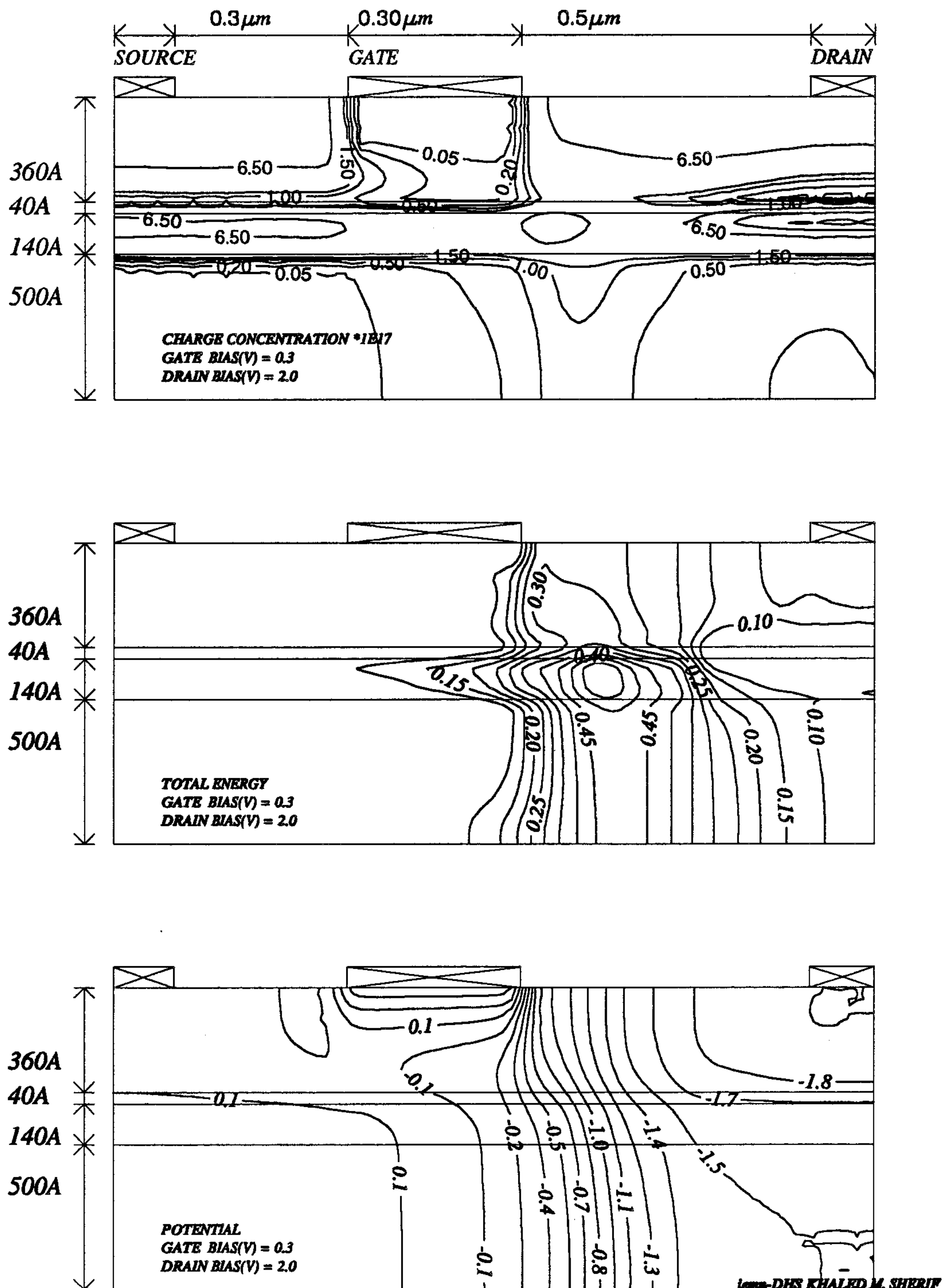


Figure III.7: Distribution of physical quantities: Electrons charge density, Average total energy, and electrostatic potential in a $0.3\mu m$ *gate pseudomorphic AlGaAs/InGaAs HEMT with uniform doping* $1 * 10^{18} cm^{-3}$ *over* $360\AA$. V_{gs}*=0.3 V and* V_{ds}*=2V*

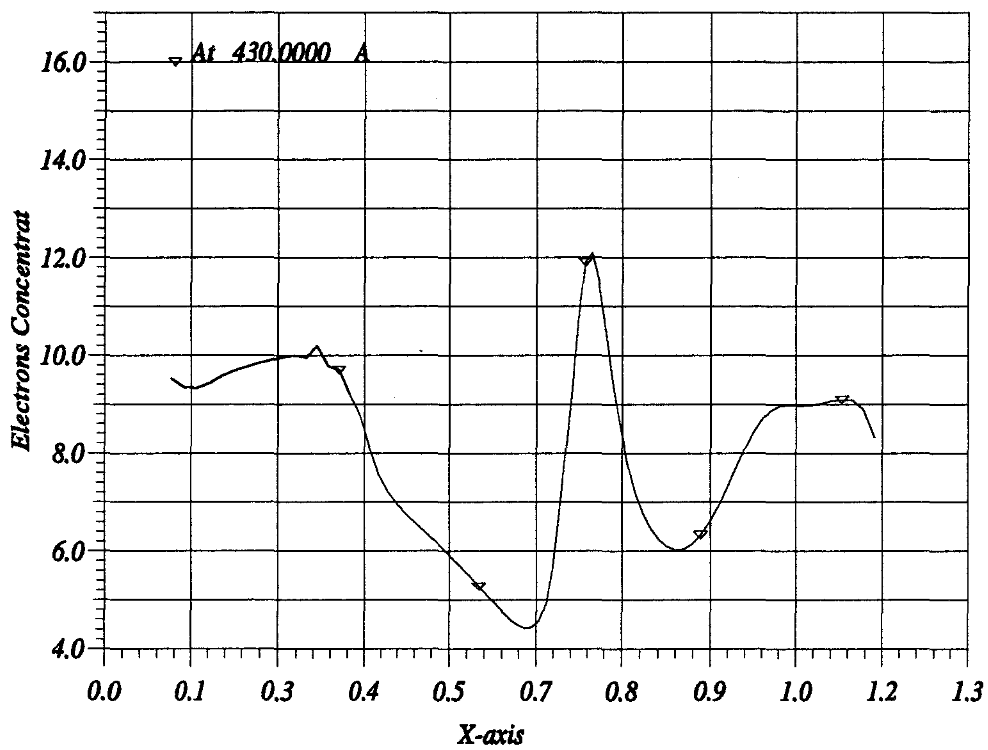


Figure III.8: Distribution of the charge density along the channel in the middle of the well (470Å from the gate) in a pseudomorphic AlGaAs/InGaAs $0.3\mu m$ *gate HEMT uniformly doped at* $1 * 10^{18} cm^{-2}$ *at* V_{gs}*=0.3 V, and* V_{ds}*=2.5 V*

the energy distribution in fig.(III.5) for $V_{gs} = -0.3$. Due to the relatively narrow well (140 A), carriers at the bottom of the well have average energy (0.9 ev) which is high with respect to the energy barrier (0.15 e.v.). It is normal then that they will crossover this barrier into the buffer layer. For positive gate bias (V_{gs} 0.3 V) the carriers heating is less as shown in fig. III.6 and the same could be noticed for smaller drain bias, fig. III.7, accordingly this phenomenon is less pronounced.

For positive gate bias, fig. III.6 we observe several phenomenon. First, there is the early electrons heating occurring under the gate which results in slowing down the electrons. A direct consequence will be the drop in the current gain cut off frequency which depends directly upon the average carrier velocity under the gate. Then we have the InGaAs high energy domain which extends over a wider distance than for a closed channel. there, the peak energy is less (about 0.5 e.v.) and the position is far from the gate exit and more towards the drain, this form of energy domain is deterministic in the break down potential of the device as we will see in chapter five. A third zone of high energy appears in the AlGaAs layer with much less peak values (0.3 ev.). This is due to the carriers traversing in the AlGaAs layer (parasitic MESFET) so not only they are in a material with mediocre transport properties (doped AlGaAs) but also they are relatively hot and accordingly slower.

III.3.3. Electrostatic Potential

The third plot in the figures give the distribution of the electrostatic potential in the device at the various operating conditions. We notice that the potential lines are almost perpendicular to the channel in the zone of the major current transfer, this characteristic is of importance for the validation of simpler models such as the quasi two dimensional ones.

For negative gate bias, the electrostatic potential lines are concentrated at the gate exit and for more negative gate bias, fig. III.5, the majority of the barrier layer is equi potential which means that the carriers collection by the drain could occur relatively early in the channel. For more positive gate bias, fig. III.6, the potential lines are distributed all along the channel and the electrode zone is restricted to the

vicinity of the contact.

III.4. Analysis of Terminal Characteristics

III.4.1. Transconductance g_m

The transconductance is a measure of gate control on the current. It is a major parameter contributing to the microwave device performance through the current gain cut-off frequency and to the digital devices through the term g_m/C_l where C_l is the load . An increase desired in these parameters implies an increase in the g_m. The transconductance could be expressed as the rate of change of drain current when the gate bias is changed at a fixed drain bias. It could be decomposed according to the different parts contributing to the total drain current. In a section under the gate we would have, for pseudomorphic HEMT's

$$g_m = \left(\frac{\partial I}{\partial V_{gs}}_{V_{ds}const}\right)_{AlGaAs} + \left(\frac{\partial I}{\partial V_{gs}}_{V_{ds}const}\right)_{InGaAs} + \left(\frac{\partial I}{\partial V_{gs}}_{V_{ds}const}\right)_{GaAs} \quad \text{(III.2)}$$

The first term is the contribution of the barrier layer, it gets more important when the channel is open and a conduction path exists in this layer. The second term is the channel share, it is mainly dependant upon the charge concentration (n_s) and the channel characteristics. The third term is the buffer contribution, it is more pronounced when there is an important buffer injection, its relative effect depends upon the characteristics of the buffer layer and the amount of injected charge.

For better HEMT operation, the first term has to be minimised since the barrier layer has mediocre transport dynamics and will dramatically affect the total transconductance.. When we examine the variation of g_m with V_{gs} compared to that of conventional HEMT's, in fig. III.9 we get the following :

- An improvement is found in the maximum value for g_m mainly due to better carrier confinement which allowed for improved gate control.

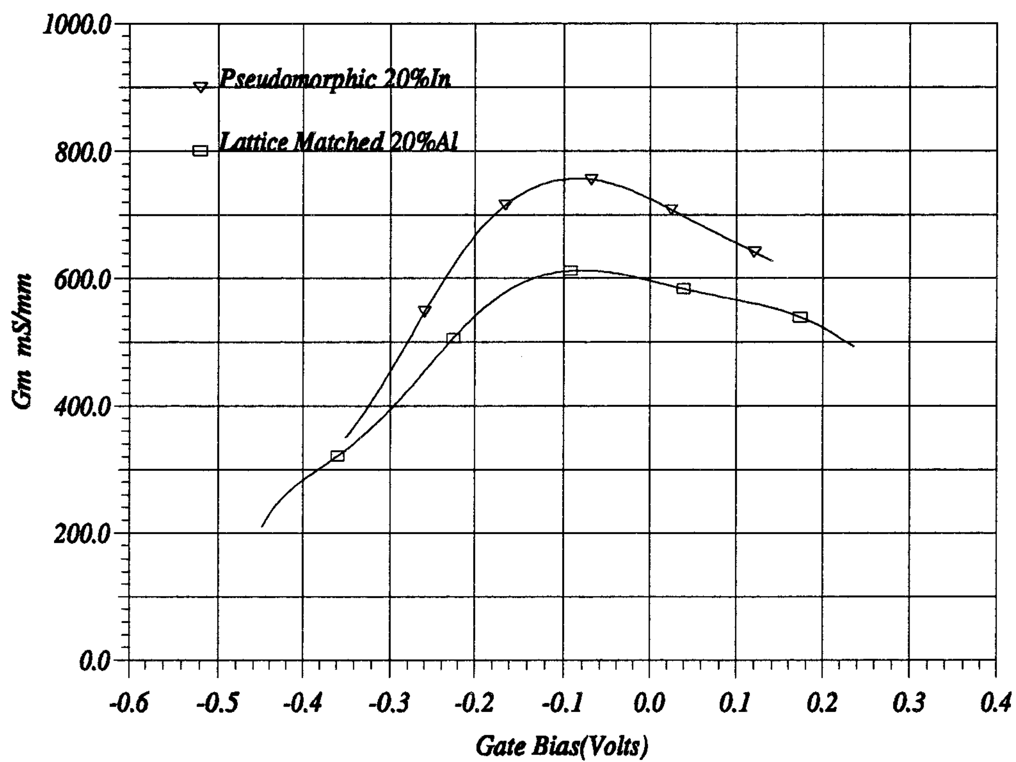


Figure III.9: Comparison of the intrinsic transconductance of conventional lattice matched and pseudomorphic uniformly doped HEMT, the devices had similar gate length of $0.3\mu m$ *and a doping level of* $10^{18}cm^{-2}$

- For large negative values of V_{gs}, the compression in g_m s due to the depletion of the 2DEG. In comparison to lattice matched device, we find a faster drop (sharper pinchoff) due to the reduction of carriers buffer injection mainly at the source side of the gate.

- For positive values of V_{gs} the conduction in the parallel MESFET starts. This has the effect of reducing the g_m (compression) in the same manner as for the conventional HEMT's.

III.4.2. The gate Capacitance C_{gs}

The second parameter that determines the microwave performance is the gate capacitance. The gate source capacitance is the rate of change of charge with gate potential is expressed as

$$C_{gs} = \left(\frac{\partial Q}{\partial V_{gs}}_{V_{ds}const}\right)_{AlGaAs} + \left(\frac{\partial Q}{\partial V_{gs}}_{V_{ds}const}\right)_{InGaAs} + \left(\frac{\partial Q}{\partial V_{gs}}_{V_{ds}const}\right)_{GaAs} \qquad \text{(III.3)}$$

In fig. III.10, we see the variation of Cgs with gate voltage for the two devices compared : a conventional lattice matched and pseudomorphic HEMT's. The first notice is that generally the pseudomorphic HEMT represents a higher capacitance. This could be explained as due to carriers confinement, the charge is kept nearer to the gate and thus the variation of the gate potential is more pronounced upon the charge variation leading to higher capacitance. The main difference in the capacitance built up could be explained according to the following :

- When V_{gs} is small the AlGaAs and the InGaAs channels are completely depleted, the 2DEG hardly formed almost does not contribute to the total capacitance and the Cgs is very small. This is true for both devices.

- When the gate bias increases the 2DEG starts to build up and the Cgs increases with V_{gs}. The capacitance variation is found to be linearly increasing

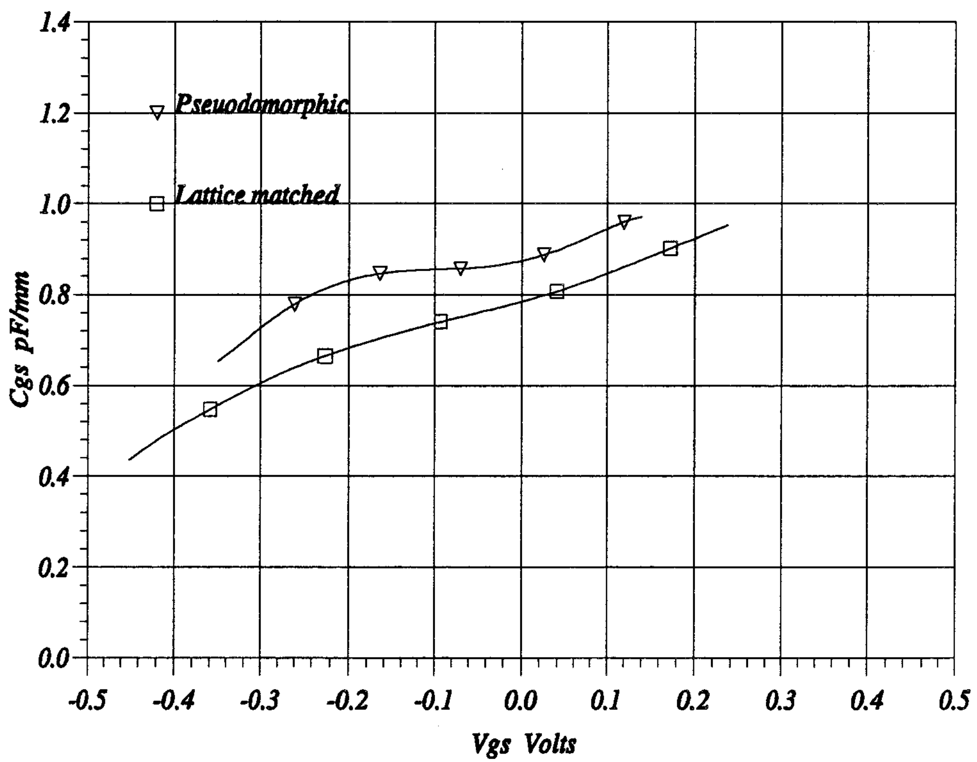


Figure III.10: Comparison of the gate source capacitance of conventional lattice matched and pseudomorphic uniformly doped HEMT, the devices had similar gate length of $0.3\mu m$ *and a doping level of* $10^{18}cm^{-2}$

with V_{gs} for conventional structures whereas for pseudomorphic ones, the variation is constant due to better carrier confinement in the channel, accordingly the separation between the gate and the 2DEG is relatively constant. This confirms with the experimental observations [58] This process continues till the full formation of the 2DEG.

- Upon the onset of parallel conduction in the barrier layer, the overall capacitance starts to increase with V_{gs} since the gate controls no more the 2DEG only but the barrier layer in addition. This is true for both devices.

The major difference is then that the pseudomorphic structures present a flatter plateau for C_{gs} as compared to the conventional devices. This is due to that the carriers are at a rather constant distance from the gate as a result of better confinement.

III.4.3. Intrinsic Current Gain cut Off Frequency f_c

The intrinsic current gain output frequency is a microwave figure of merit as it gives a direct indication of the device gain and operating frequency. When we compare conventional and pseudomorphic devices, fig. III.11, we find that the pseudomorphic structures present a better performance of f_c with V_{gs}.

This superiority could be explained when we examine the characteristics of the current gain cut off frequency. The f_c is calculated as:

$$f_{ci} = \frac{g_m}{2\pi C_{gs}} \tag{III.4}$$

Due to better carrier confinement, the pseudomorphic devices have shown superior transconductance and better linearity in the behaviour of gate-source capacitance, this led to the show superiority in the f_c

In this same context, we find that the approximation:

$$f_{ci} \approx \frac{v_{eff}}{2\pi L_g} \tag{III.5}$$

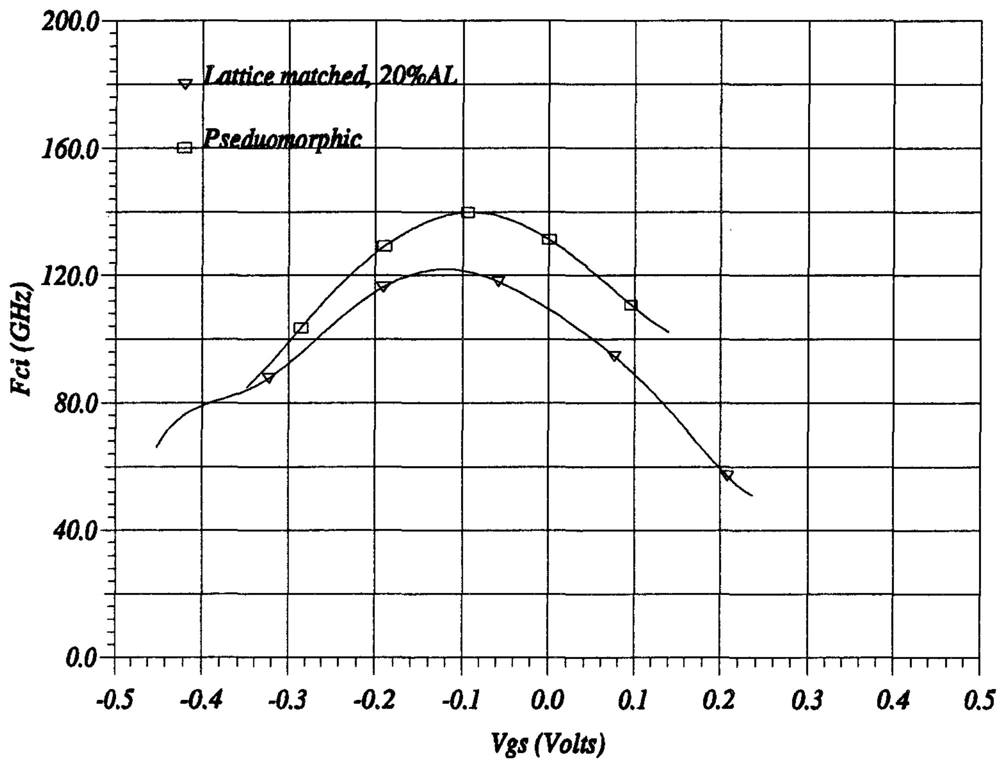


Figure III.11: Comparison of the intrinsic current gain cut off frequency of conventional lattice matched and pseudomorphic uniformly doped HEMT

where v_{eff} and L_g are the effective (average) velocity in the channel and the gate length,is not quite accurate. Since we have applied the same transport dynamics in GaAs and InGaAs, we ought to find the same frequency at the same gate length. This is not the case. In fact, the gate length in the case of the conventional device will be effectively longer due to the lack of confinement and thus having a longer path for the carriers.

III.4.4. Drain conductance g_d

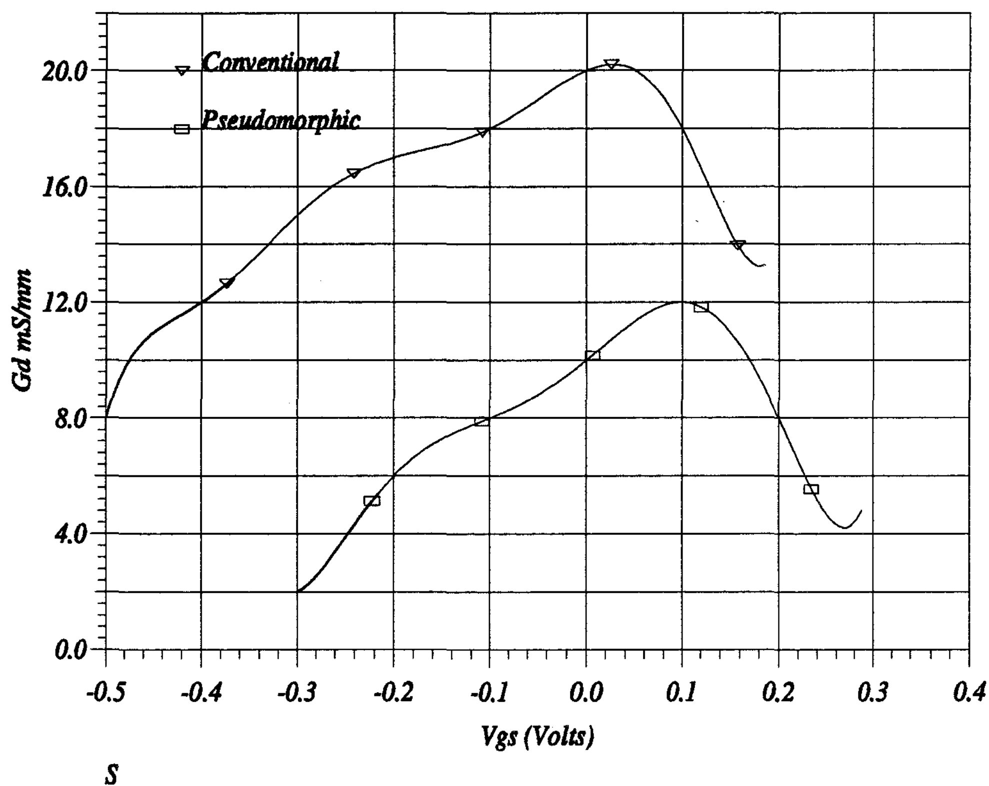


Figure III.12: Comparison of the output conductance in two identical pseudomorphic AlGaAs/InGaAs and conventional AlGaAs/GaAs HEMT's. Gate length is $0.3\mu m$, doping level is $1 * 10^{18} cm^{-3}$, drain potential is $2.5V$

A major advantage in using pseudomorphic structures would be in having reduced drain conductance g_d. This is manifested in fig.(III.12) where we find lower g_d as compared to the conventional device. The reduced g_d is essentially due to better confinement resulting in reduced path for the carriers in the transfer from source to drain.

III.5. Pulse doped HEMT's

III.5.1. Introduction

The pulse doping of the barrier layer is of interest because it allows to introduce a high doping density resulting in increasing the 2DEG charge density while avoiding the problems related to high uniform doping such as the DX centres and leaky Schottky gates. Another advantage of the pulse doped barrier layers is the reduced doping density in the vicinity of the gate and the gate drain region, this results in improved breakdown performance due to reduced gate tunnelling currents and reduced impact ionisation.

III.5.2. Distribution of physical quantities

We simulated a pulse doped HEMT which would have the same theoretical pinch off voltage as the previously simulated uniformly doped one. This device contains a superior undoped 20% AlGaAs layer of thickness $150\AA$, a doped plan of doping $1.5^{12} cm^{-2}$ over $30\AA$, a set back (spacer) layer of $40\AA$, a 20% InGaAs layer of thickness $140\AA$, then a buffer layer of $500\AA$.

The distribution of the major physical quantities of this device is shown in fig. III.13 where we could easily remark the higher concentration in the channel. As a matter of fact the n_s increased form $1.34 * 10^{12} cm^{-2}$ in the uniformly doped case to $2.1 * 10^{12} cm^{-2}$ for the pulse doped. Apart from the access zones under the electrodes, the concentration of the electrons is low in the barrier layer.

Because of the use of a GaAs buffer the phenomenon of hot electrons injection into the buffer layer at the gate exit still manifests. While the average elections energy is of the same order as for the uniformly doped device.

The potential contours show that there exists a high electric field gradient at the gate exit and the access zone for the drain electrode is over two third of the gate drain spacing.

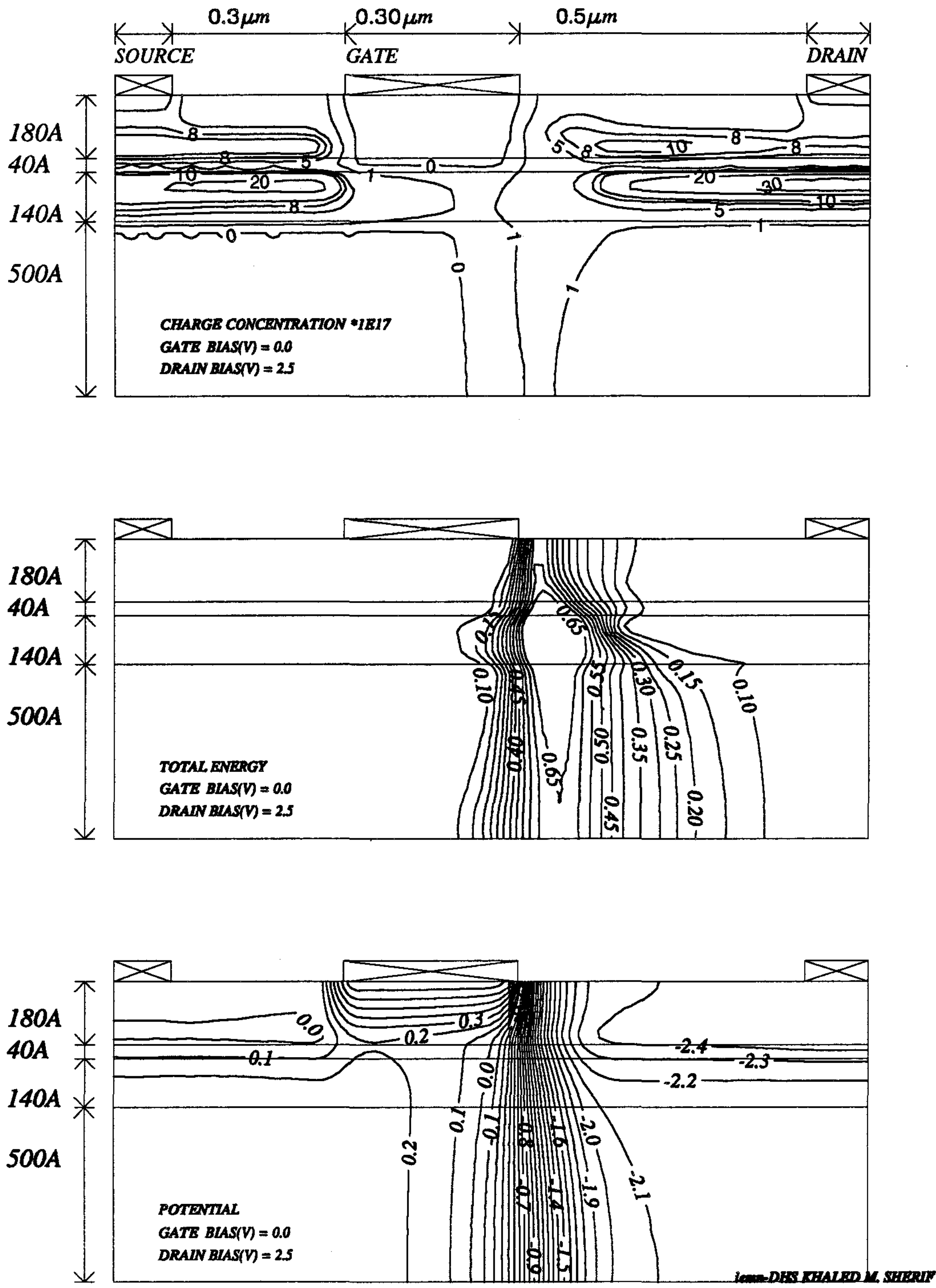


Figure III.13: Distribution of physical quantities in a pulse doped AlGaAs/InGaAs pseudomorphic HEMT

III.5.3. Comparison with uniform doping

III.5.3.1. Drain current

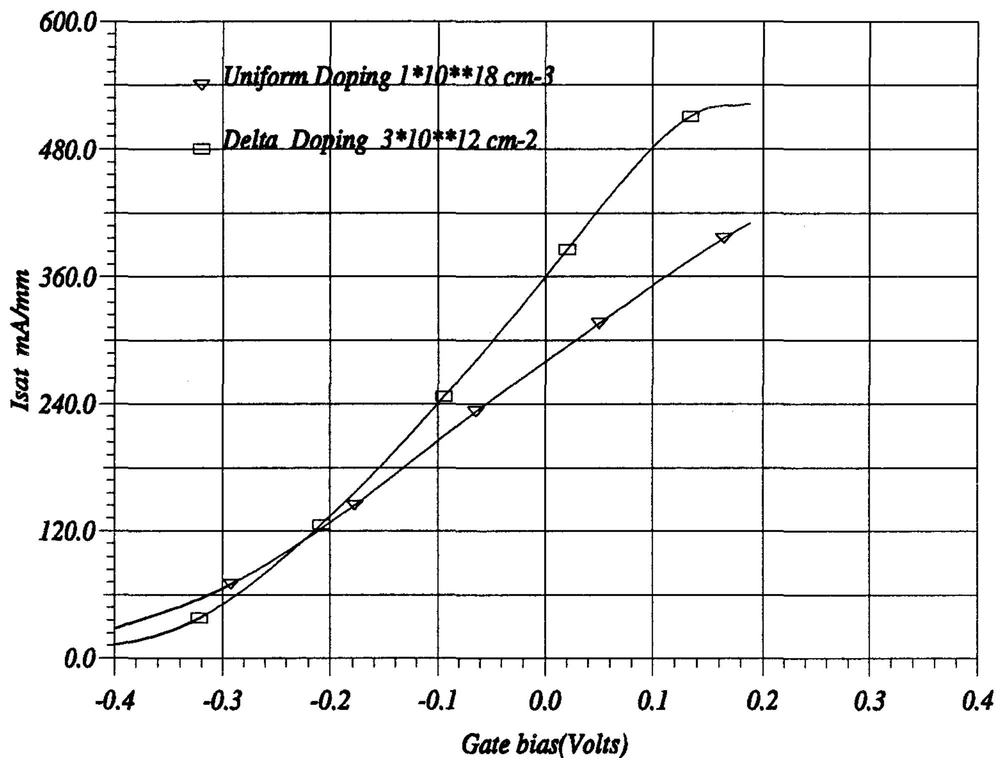


Figure III.14: Output currents of 0.3μm *gate devices; comparison between uniform and pulse doping at* V_{ds}=2.5 V

It would be expected that the major gain in using pulse doping is in an increased output current. Accordingly we plot the drain currents in the two devices in fig.(III.14. The current variation could be divided into two zones. The first zone for V_{gs} is less than (-0.2V), there we find that the pulse doped device presents a lower drain current, this is mainly due to the shorter gate-channel separation. The second zone, for gate potential higher than 0.2V, the contribution of the high n_s 2DEG is significant

and we have an increase in the current by about 25%.

III.5.3.2. The Gate capacitance

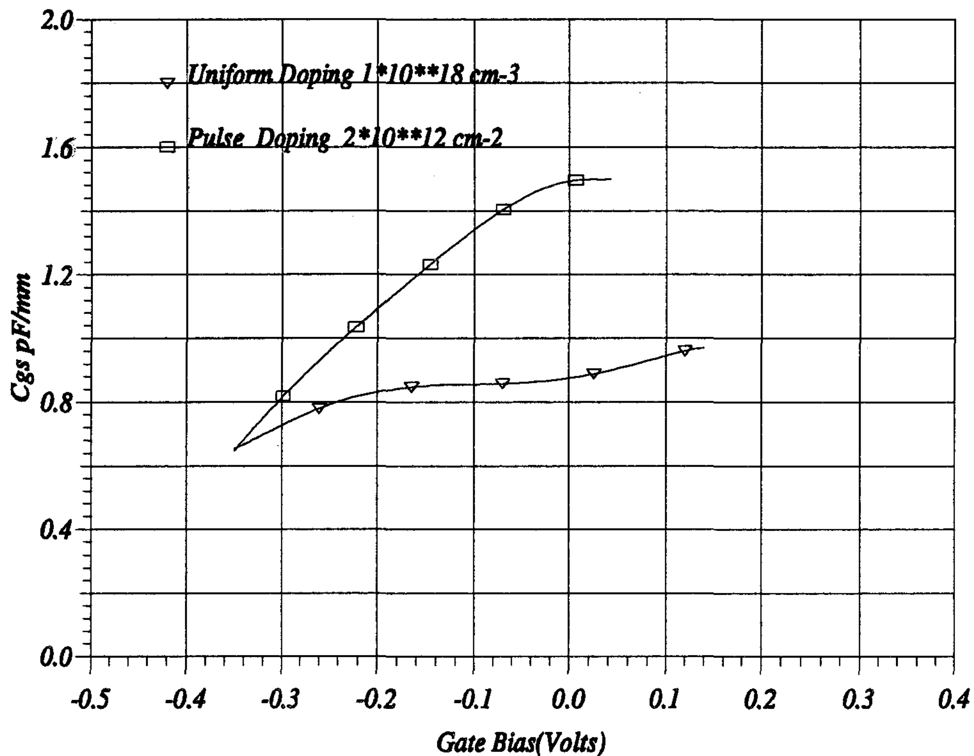


Figure III.15: Comparison of the gate to source capacitance of pulse doped and uniformly doped pseudomorphic AlGaAs/InGaAs HEMT with $0.3\mu m$ *gate. The drain bias is 2.5V*

The gate capacitance of the pulse doped device presents an overall higher value as compared to the case of the uniform doping, also it presents a faster variation with V_{gs}, fig. III.15. To understand this, let us consider the position of the 2DEG relative to the gate. First, the average distance of the 2DEG in the case of pulse doped is $290\mathring{A}$ (assuming the 2DEG is in the middle of the well), whereas for the uniform doping this distance is $470\mathring{A}$, this explains the increase in the capacitance.

As for the variation of C_{gs} with V_{gs}, we present the following explanation. Assuming that the 2DEG moves from the top of the well to the bottom when the gate bias is changed from positive to negative, in the case of uniform doping the minimum separation would be 400$\AA$ and the maximum would be 540$\AA$ giving about 26%. For the pulse doped device the minimum separation is 220$\AA$ and the maximum is 360$\AA$ giving about 63% variation. So the gate capacitance for the pulse doped will increase approximately as twice as fast as for the uniformly doped device. In addition, for the pulse doped device, upon open channel operation and the onset of parallel conduction in the barrier layer, the charge of the doped plane contributes to the increase in the capacitance.

III.5.3.3. Effect on the intrinsic Transconductance

The variation of g_m could be seen in the same perspective as the gate capacitance. Since the average separation between the gate and the 2DEG is smaller in the pulse doped device, this will result in an increased transconductance as could be seen from fig. III.16. Due to the larger relative variation in the position of the 2DEG, the g_m rate of change with V_{gs} is higher in the case of pulse doping. The sharp compression in the g_m is due to the contribution of the doped plane upon the onset of parallel conduction in the barrier layer.

III.5.3.4. Effect on f_c

The microwave figure of merit represents a rather peculiar behaviour. As a result of the increase in the aspect ratio (13.6 for the pulse doped and 7.5 for the uniformly doped) we have a peak value that is about 20% higher, fig. III.17.

For more positive gate bias, the parallel conduction (parasitic MESFET) contributes in the characteristics of the device. This conduction takes place in the heavily doped region where the transport dynamics are terribly bad. The f_c suffers then a dramatic compression to the extent that it becomes lower than that of the uniformly doped device at gate bias more positive than -0.05V.

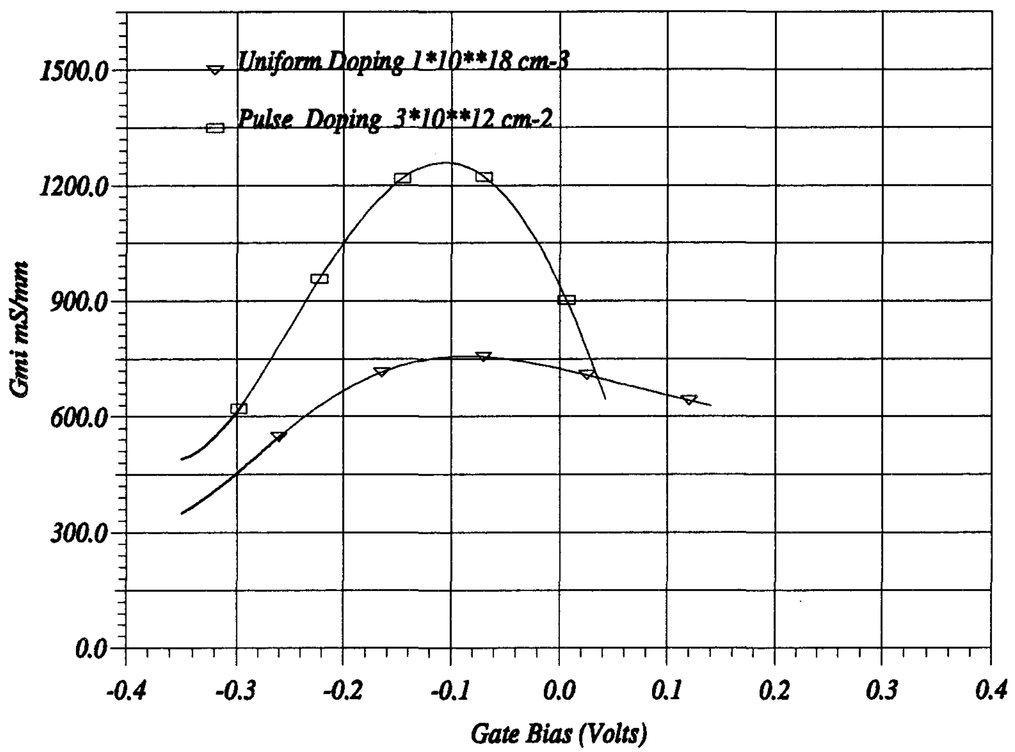


Figure III.16: Comparison of the intrinsic transconductance of a pulse doped and uniformly doped pseudomorphic AlGaAs/InGaAs HEMT with $0.3\mu m$ *gate, drain bias is 2.5V*

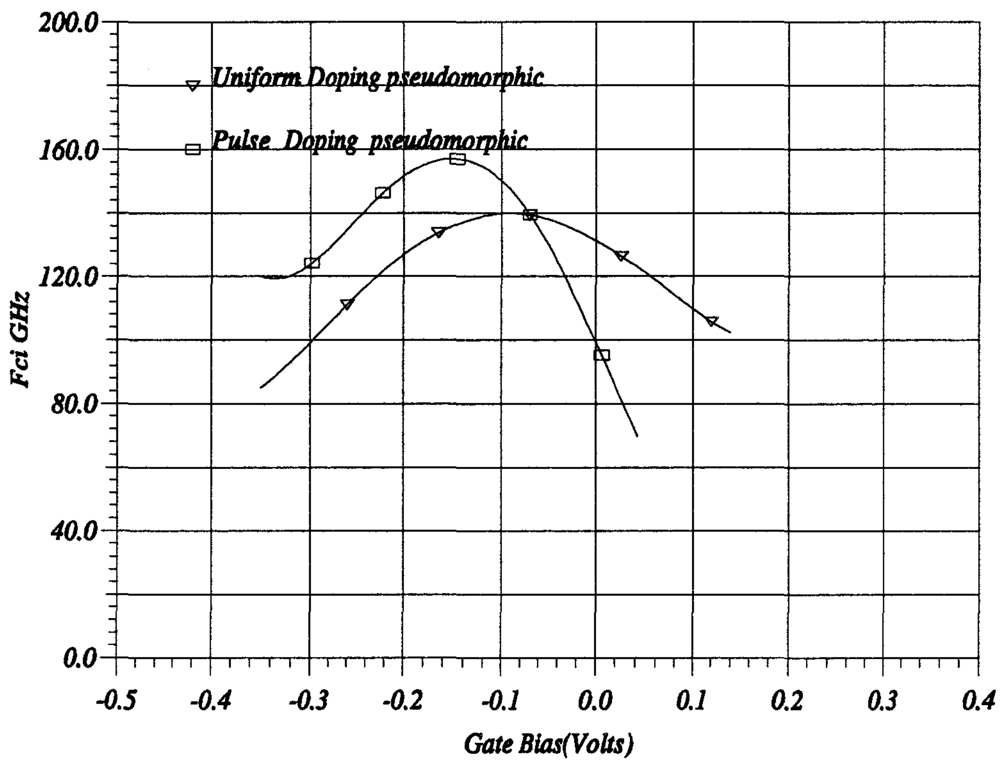


Figure III.17: Comparison of the intrinsic current gain cut off frequency of conventional lattice matched and pseudomorphic uniformly doped pseudomorphic AlGaAs/InGaAs HEMT with 0.3μm *gate, drain bias is 2.5V*

BU LILLE 1

For microwave applications it is of importance to know the variation of the cut off frequency with the drain current, it is important to have high f_c for low current (low noise) and high current (large signal) conditions. The comparison between the two devices in terms of $f_{ci}(I_{ds})$, fig. III.18 shows that the pulse doped device presents better performance over the whole range of operating currents. This is explained by the improved current levels as a result of higher n_s and the improved f_c due to better aspect ratio.

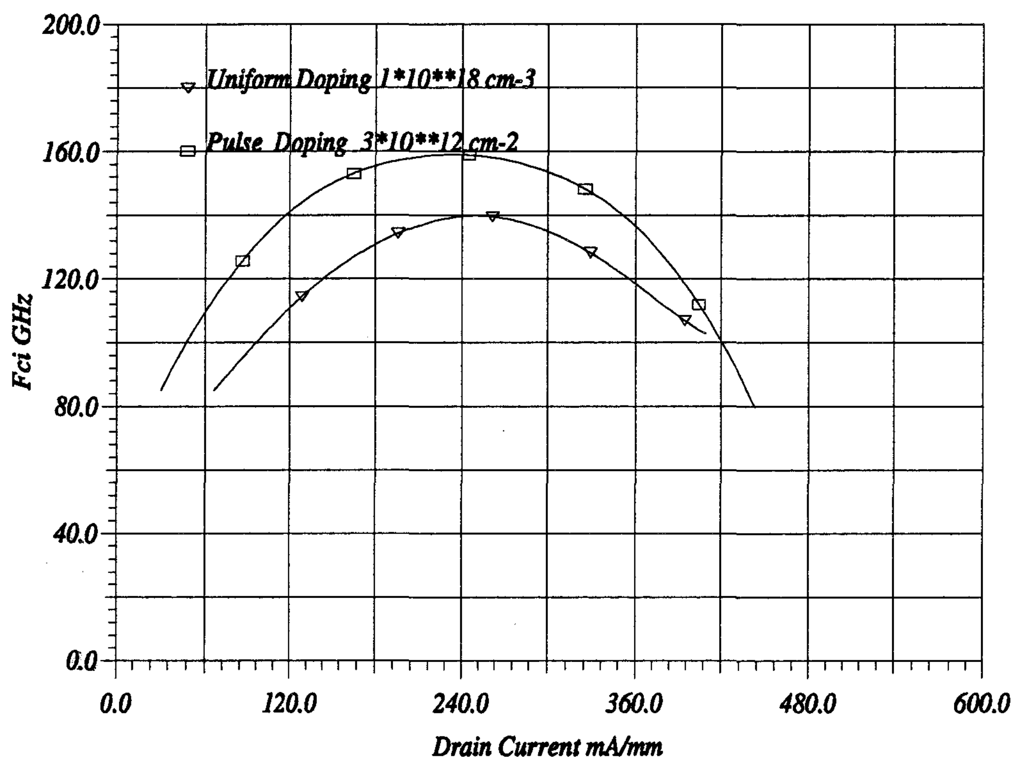


Figure III.18: Variation of the current gain cut off frequency with steady state drain current for a pulse doped and a uniformly doped pseudomorphic AlGaAs/InGaAs HEMT with a $0.3\mu m$ *gate and at a drain bias* $2.5V$

III.5.4. Recapitulation

To conclude this part, the simulation of pulse doped pseudomorphic transistors shows that they present better performance in the current gain cut off frequency both at low noise conditions and high power ones. The increase in the gate source capacitance is attributed to the relative position of the 2DEG with respect to the gate.

III.6. Effect of buffer type on device performance

To terminate the study of pseudomorphic HEMT's it would be interesting to investigate the effect of the buffer material upon the device performance[1]. The interest in changing the buffer type is motivated by the will of increasing the n_s through increased carrier confinement as an AlGaAs buffer will supply conduction band discontinuity of 0.31 ev.

For this purpose, we simulated two devices of similar structure while changing only the buffer type using AlGaAs with 20% Aluminium in one device and GaAs in the other. The devices had a pulse doped barrier layer of doping $3 * 10^{12} cm^{-2}$ and a gate length of $0.3 \mu m$.

III.6.1. Distribution of physical quantities

Fig. III.19 gives the distribution of the charge concentration average total energy and the potential for the device with an AlGaAs buffer. The equivalent distribution in a device with GaAs buffer layer are shown in fig. III.20. It is obvious that the use of AlGaAs buffer reduces to a great extent the injection of hot carriers into the buffer layer by introducing a higher energy barrier (0.31 ev).

III.6.2. Terminal characteristics.

III.6.2.1. Transconductance and current gain cut off frequency

These two parameters presented in fig. III.21 and fig. III.22, give tendency to conclude that hardly no effect occurs upon the use of a AlGaAs buffer layer. In fact, we have noticed very small increase in the confinement where the n_s increased from $2.3 * 10^{12} cm^{-2}$ to $2.4 * 10^{12} cm^{-2}$. The maximum f_c is still 120 GHz for the two devices.

[1]The simulations in this section were made using the modified model

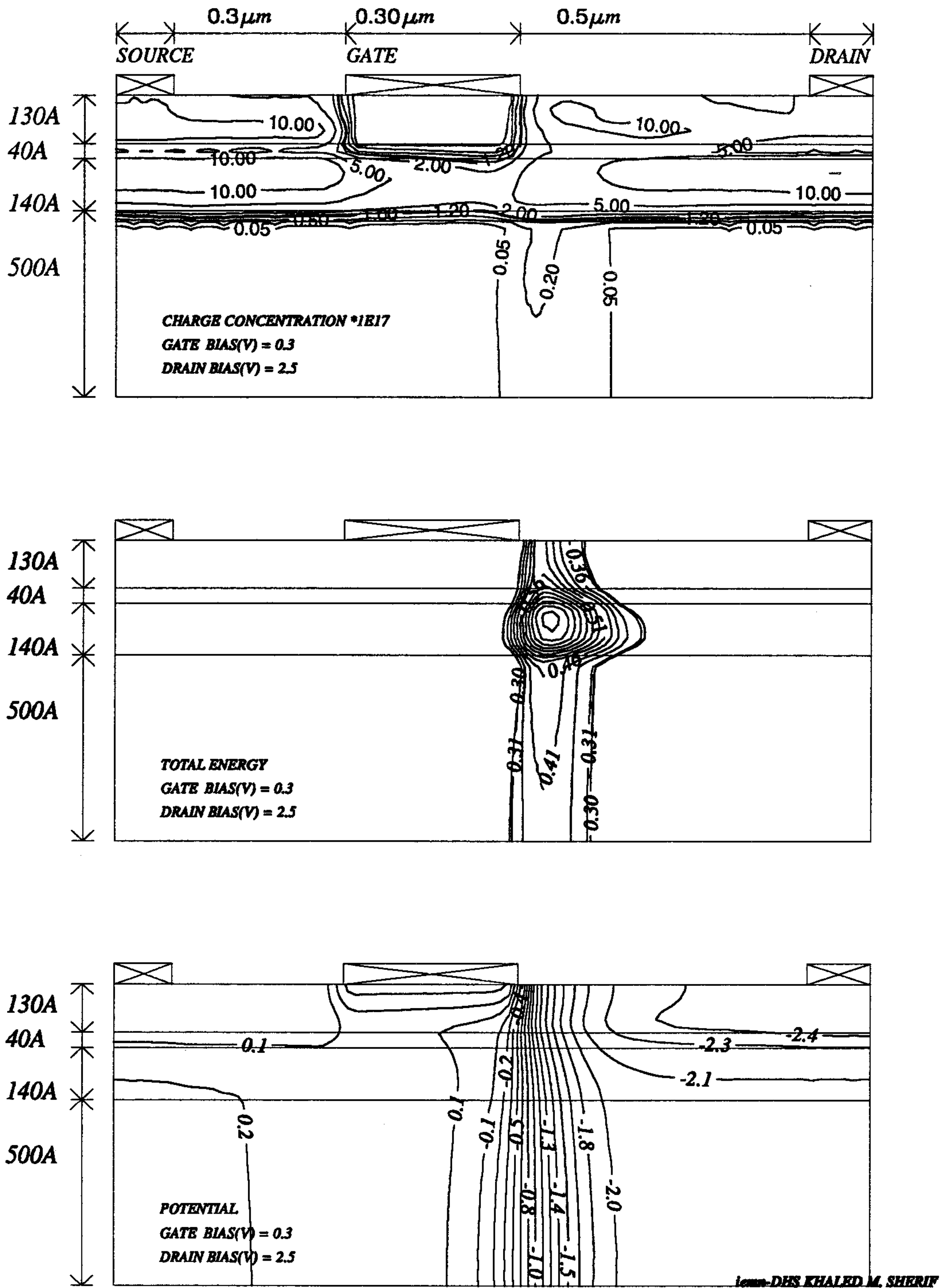
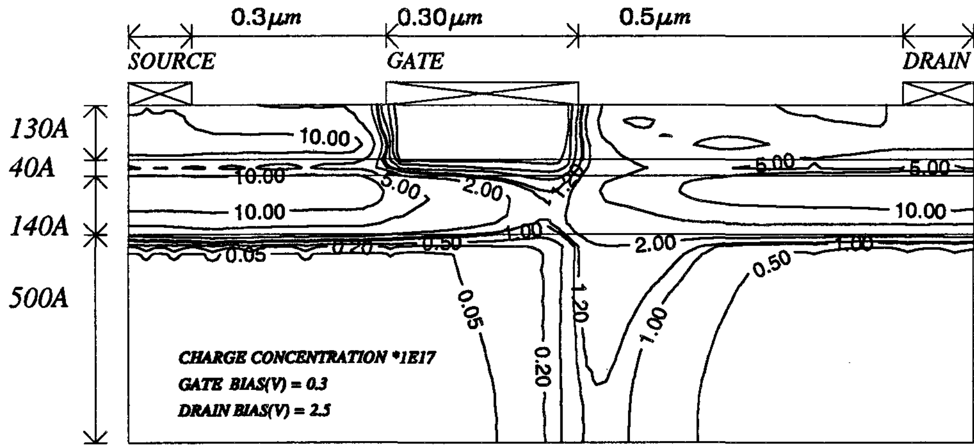
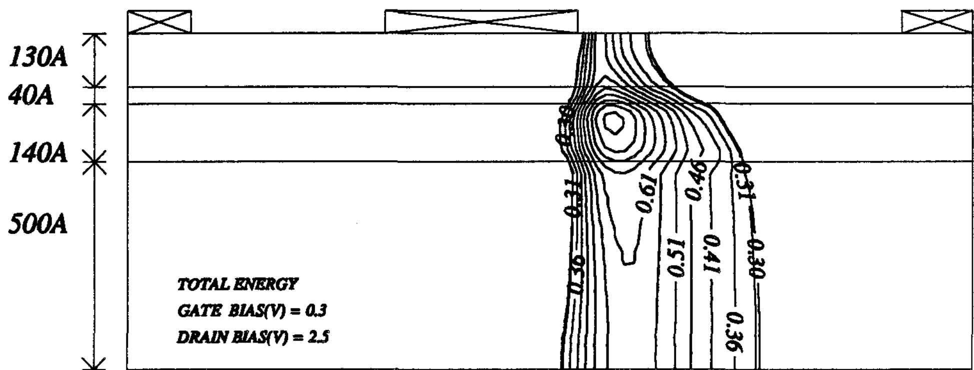


Figure III.19: Distribution of physical quantities in a pulse doped AlGaAs/InGaAs pseudomorphic HEMT with a n.i.d AlGaAs buffer layer





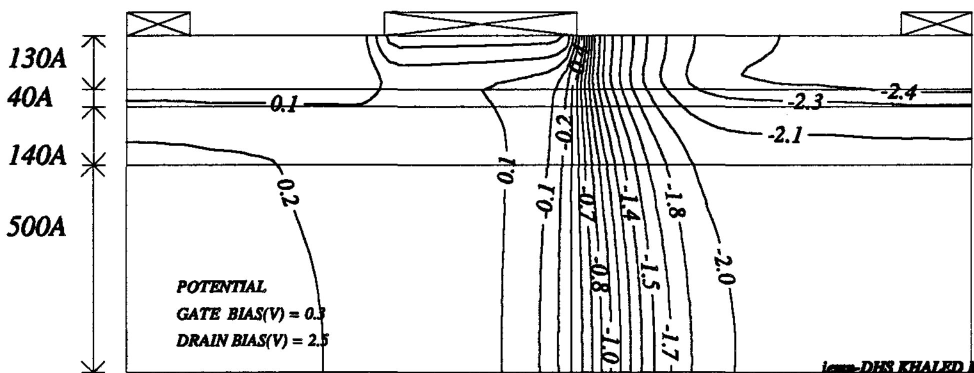


Figure III.20: Distribution of physical quantities in a pulse doped AlGaAs/InGaAs pseudomorphic HEMT with a n.i.d. GaAs buffer layer

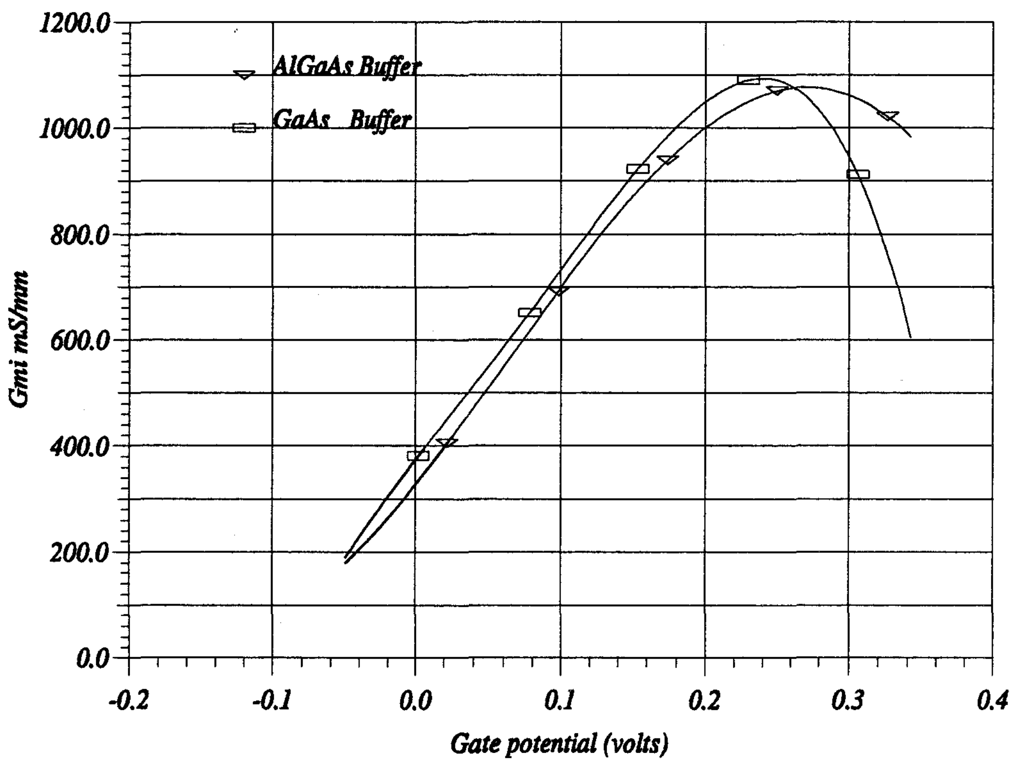


Figure III.21: Intrinsic transconductance versus gate voltage: Comparison of buffer type in a pseudomorphic AlGaAs/InGaAs HEMT with $0.3\mu m$ *gate length and planar doped at* $3 * 10^{12} cm^{-2}$.

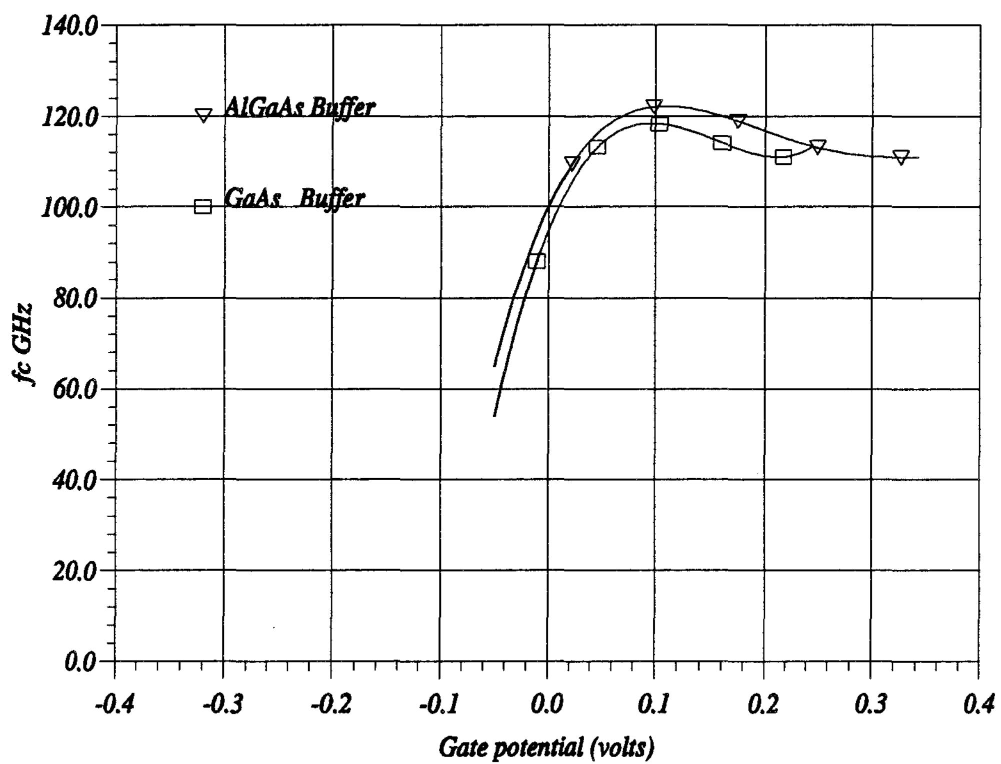


Figure III.22: Intrinsic current gain cut off frequency versus gate voltage: Comparison of buffer type in a 0.3μm gate pseudomorphic AlGaAs/InGaAs HEMT pulse doped at $3 * 10^{12} cm^{-2}$

III.6.3. Effect on feedback parameters

A possible merit for using AlGaAs buffer layers could be in having improved feedback parameters; output conductance and capacitance. These two are plotted in fig. III.23 and fig. III.24. We find that for the device with AlGaAs, lower g_d is obtained in

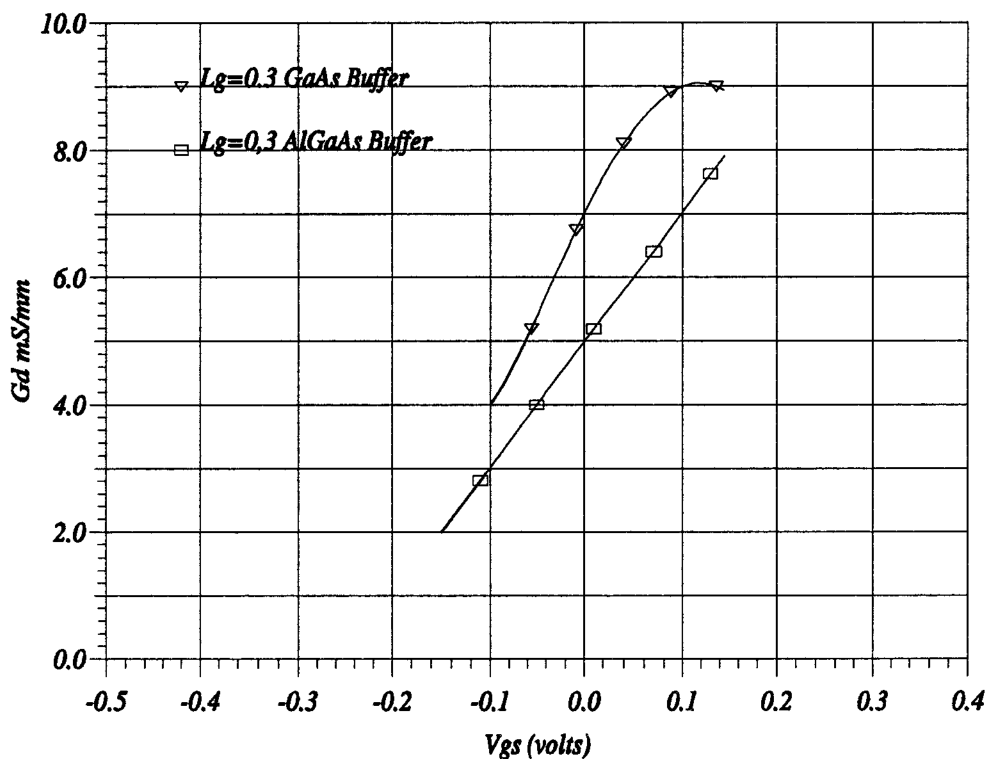


Figure III.23: Output drain conductance versus gate voltage: Comparison of buffer type in a pseudomorphic AlGaAs/InGaAs HEMT with $0.3\mu m$ *gate length and planar doped at* $3 * 10^{12} cm^{-2}$.

the region of low noise operation, although the maximum is comparable for the two devices. As for Cgd, the two devices again give comparable behaviour with slightly lower values for the AlGaAs buffer.

When we calculate the maximum available gain cut off frequency F_{MAG} and the

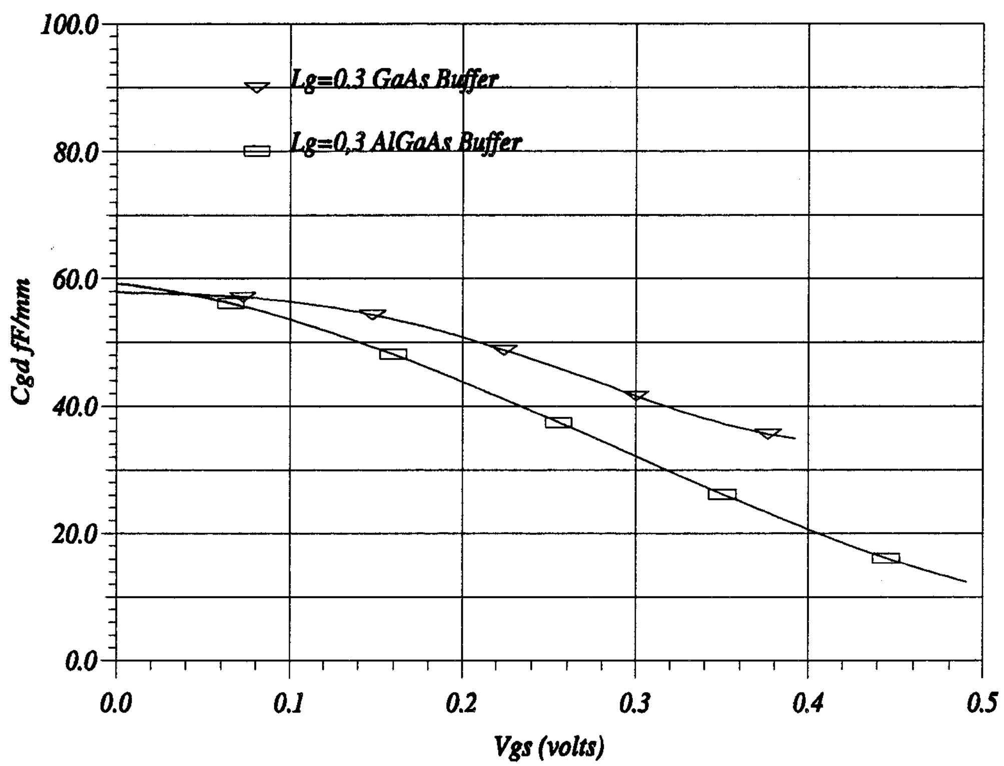


Figure III.24: Feedback capacitance versus gate voltage: Comparison of buffer type in a pseudomorphic AlGaAs/InGaAs HEMT with $0.3\mu m$ *gate length and planar doped at* $3 * 10^{12} cm^{-2}$.

maximum frequency of oscillation f_{max} at a drain voltage of 2.5V and at the gate voltage corresponding to the maximum f_c, for a $100\mu m$ device with $0.2\Omega.mm$ access resistance, we find 133 and 164 respectively for the device with GaAs buffer and 141 and 175 GHz for the device with AlGaAs buffer. Thus, an increase in F_{MAG} and f_{max} of about 6% seems to be the only advantage in using AlGaAs buffer layers.

III.7. Recapitulation

The pseudomorphic AlGaAs/InGaAs HEMT's show better performance than conventional lattice matched devices. This was confirmed by the simulations done using a simplified hydrodynamic energy model. Although the absolute values of the circuit parameters are somehow overestimated when compared to experimental measurements, the relative evolution of the different parameters is in the right sense and in the same ratio of improvement reported in experimental work.

The pulse doping of the barrier layer offers several advantages as compared to uniform doping, the first of which higher current gain cut off frequency. This is besides the major advantage in increasing the break down potential.

The use of AlGaAs buffer layers seems to be of a minor effect as compared to the GaAs buffers. An increase in the f_{MAG} and f_{max} of about 6% is apparantly the only advantage.

III.8. On The Transient Instabilities in HEMT's

III.8.1. Existence of Gunn oscillations

It is agreed upon that the III.V compound semiconductors devices could present a transient time instability manifested by the appearance of current pulses due to a travelling high field domain. This phenomenon has been seen in devices based on GaAs or InP[59].

Two terminal devices were conceived to produce such oscillatory currents and used later as oscillators in the centimetre and the millimetre wave range. At the same time, extensive research was carried on to comprehend and to master this phenomenon in the two terminal devices[60].

In the three terminal devices, current oscillations due to travelling Gunn domains were reported in several cases. The GaAs MESFET's have shown this type of instability and were characterised both through simulations[61] and experimentally [62]. Detailed investigations were carried on the nature and the conditions of formation of such domains and concluded to specific conditions of the domain formation as related to the channel thickness. The InP MESFET's have also shown Gunn instabilities specially for planar structures[63] where the oscillations existence and frequency were shown to vary with physical device parameters.

More recently, Gunn oscillations were reported in Heterojunction Bipolar Transistors (HBT) for very large base bias[64], and MESFET's with low temperature grown insulator layer which have shown Gunn instabilities over a wide range of gate bias and layer structures[65].

As for HEMT's, first reports of Gunn oscillation were reported by the author for conventional HEMT's at 77K[66]. Nanosecond instabilities were reported[67]. Moreover, Monte-Carlo simulations have demonstrated the existence of such domains in MESFET's[68], and more recently in our laboratory metamorphic HEMT's have exhibited such oscillations[69].

In all these devices, the Gunn instabilities represents an undesirable effect that sets a limit on their use specially in power applications.

III.8.2. Oscillations in HEMT's

In order to investigate the phenomenon of Gunn oscillations in HEMT's, we simulated a pulse doped HEMT wit $0.3\mu m$ gate, $100\AA$ undoped barrier layer, $3*10^{12}cm^{-2}$ doped plan and $140\AA$ of InGaAs with 20% In.

Fig.(III.25), gives the output current when the gate voltage is 0.5 V and the drain is 3V. We find that both the drain and source currents give temporal variations in a repeated pattern i.e. oscillations. The period of each cycle is 5psec as we could deduce from the plot. This means that we have a frequency of 200 GHz.

These oscillations are due to a travelling Gunn domain in the gate drain region. We could understand the mechanism of its motion when examining figures.(III.26,III.27) which give the position of the high energy domain and fig.(III.28,III.29) which give the associated equipotential contours, over an oscilation cycle (5psec) with a (1psec) step. We could see that the domain nucleates at the gate's exit, starts to grow up there and then starts travelling towards the drain. Upon reaching the drain the domain starts to discharge giving rise to a current pulse whereas a new domain starts to nucleate at the gate once more to restart the cycle. Let us now calculate the frequency of the domain cycle. The carriers in the domain move with their saturation velocity, this could be easily deduced since the energy in the domain is very high (0.6 V on the average), the saturation velocity is ($10^5 m/sec.$). The distance travelled by the high energy domain is the gate drain spacing, thus $0.5\mu m$, the thing which gives a frequency of 200GHz. We have systematically verified, like other authors did, the physical nature of these oscillations and we excluded the possibility of numerical instabilities.

But what happens for different gate bias? To answer this question, we systematically simulated the device while keeping the drain bias fixed, at different gate potentials. We found that for negative gate bias, the oscillations cease to exist. The output

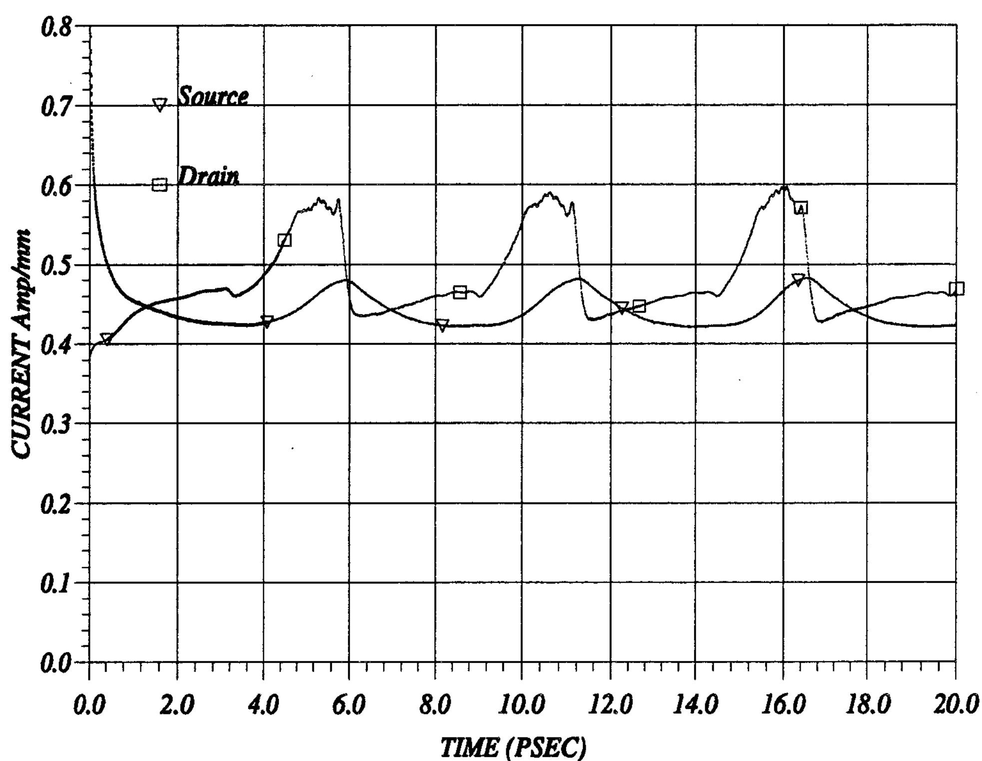


Figure III.25: Terminal currents in a $0.3\mu m$ gate pulse doped pseudomorphic HEMT with $3 * 10^{12} cm^{-2}$ doping level and 140Å InGaAs channel with 20% In. The gate bias is 0.5 V, and the drain is 3 V.

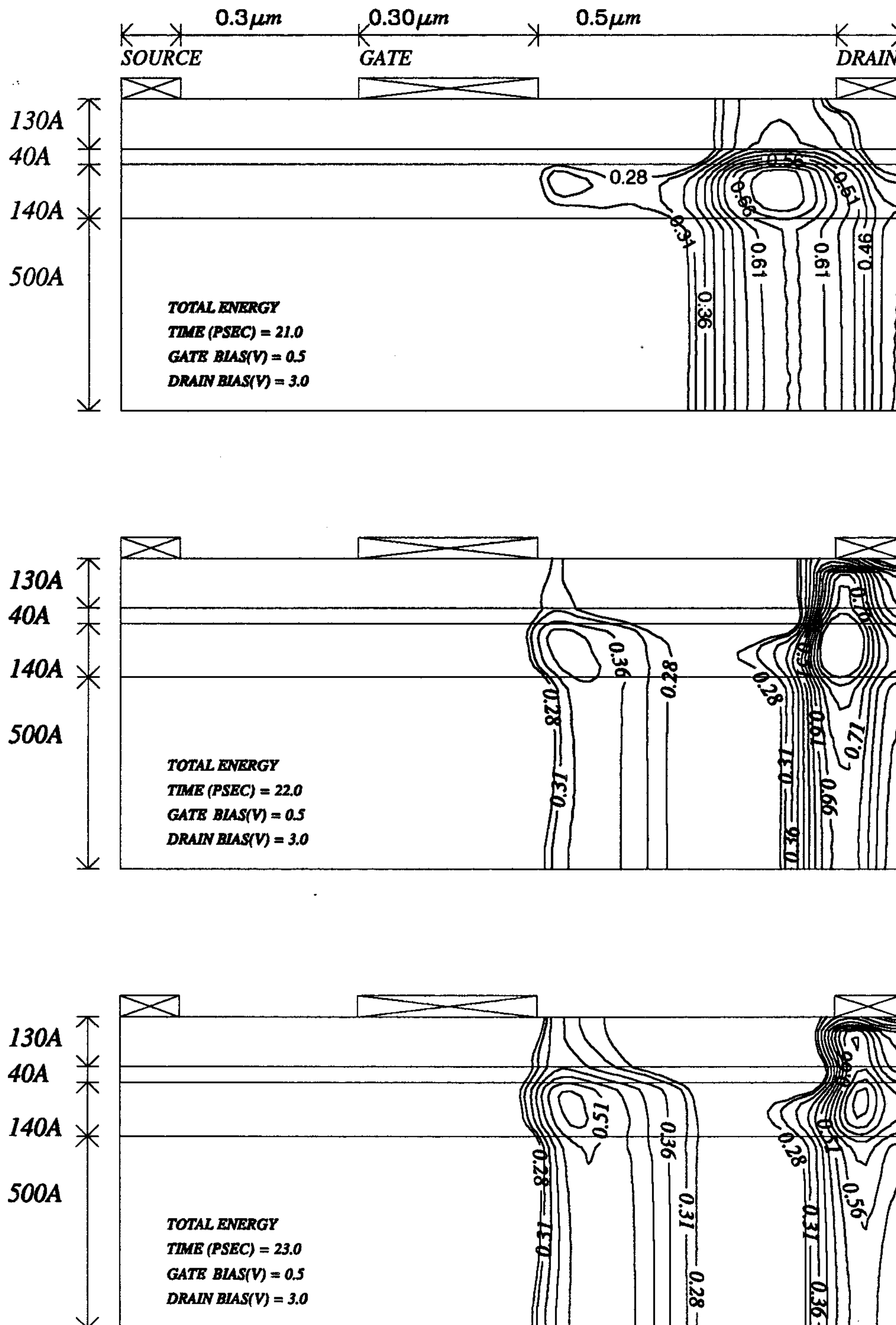


Figure III.26: Travelling Gunn Domain in the gate drain region of a pulse doped HEMT, V_{gs}=0.5V , V_{ds}=3 V, time instants at 20,21 and 22 psec

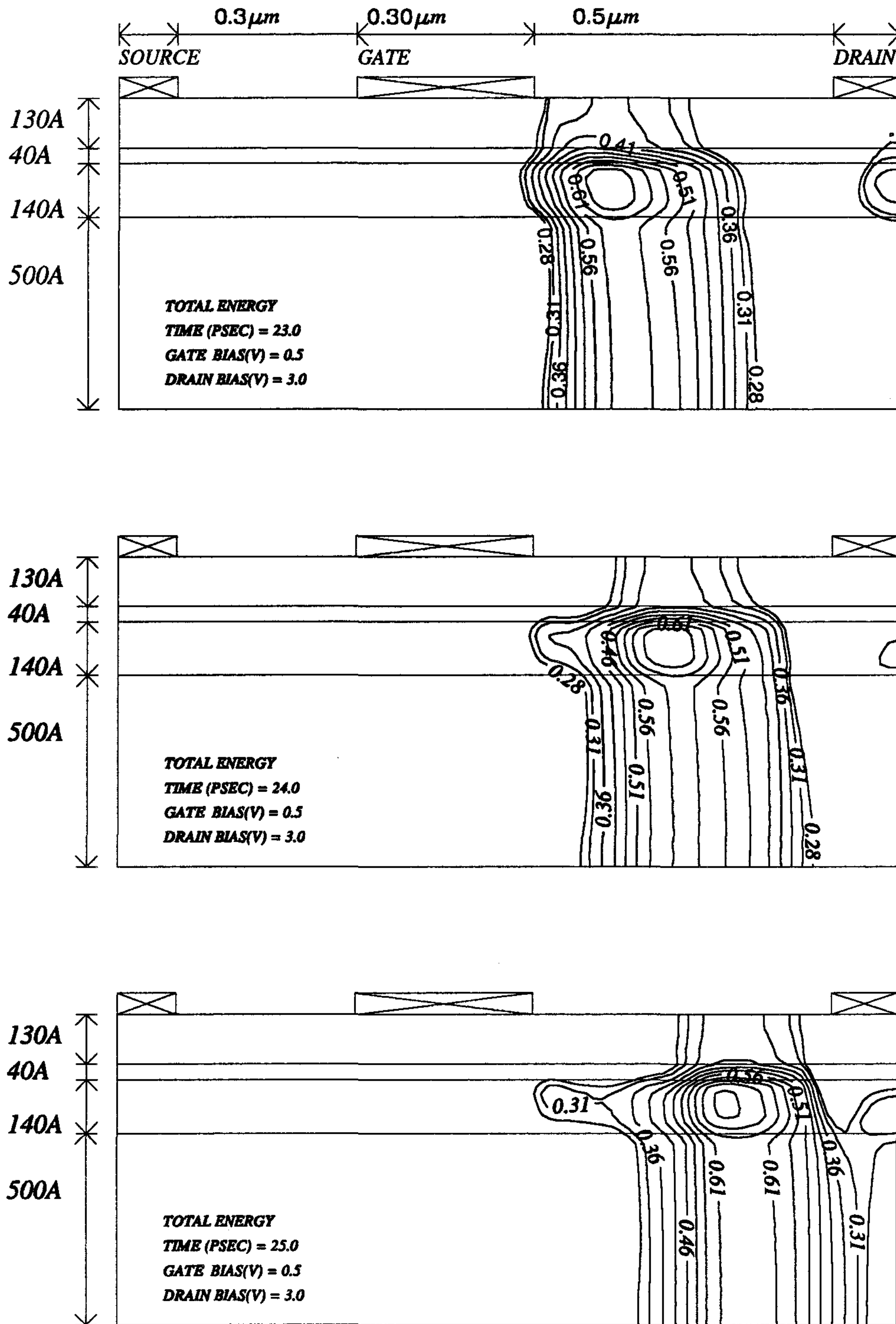


Figure III.27: Travelling Gunn Domain in the gate drain region of a pulse doped HEMT, V_{gs}=0.5V , V_{ds}=3 V, time instants 23,24 and 25 psec: the end of the cycle

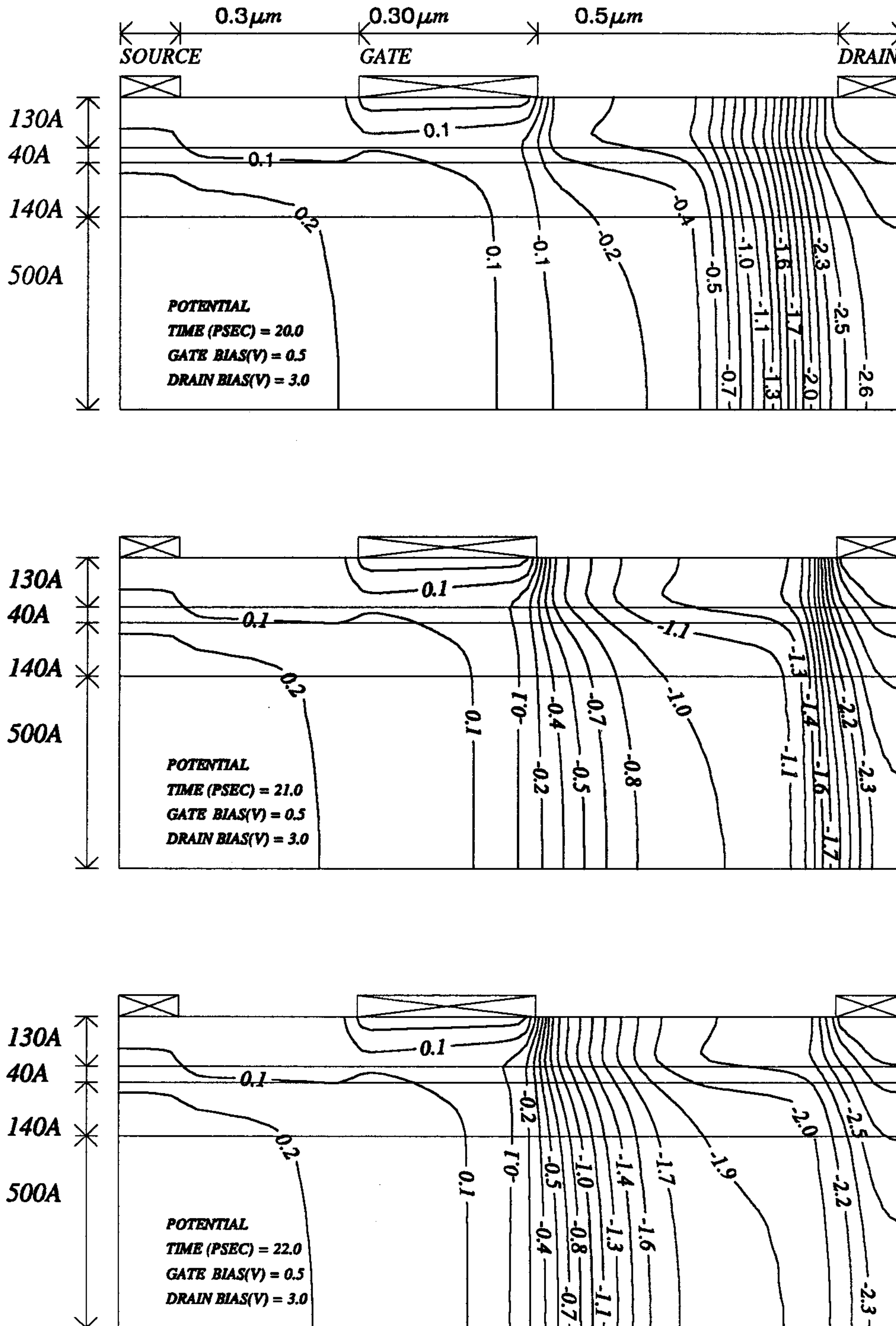


Figure III.28: Associated equipotential contours of the high energy-gigh field domain in a pulse doped 0.3μm HEMT, V_{gs}=0.5V , V_{ds}=3 V at the first three time instants 20, 21 and 22 psec.

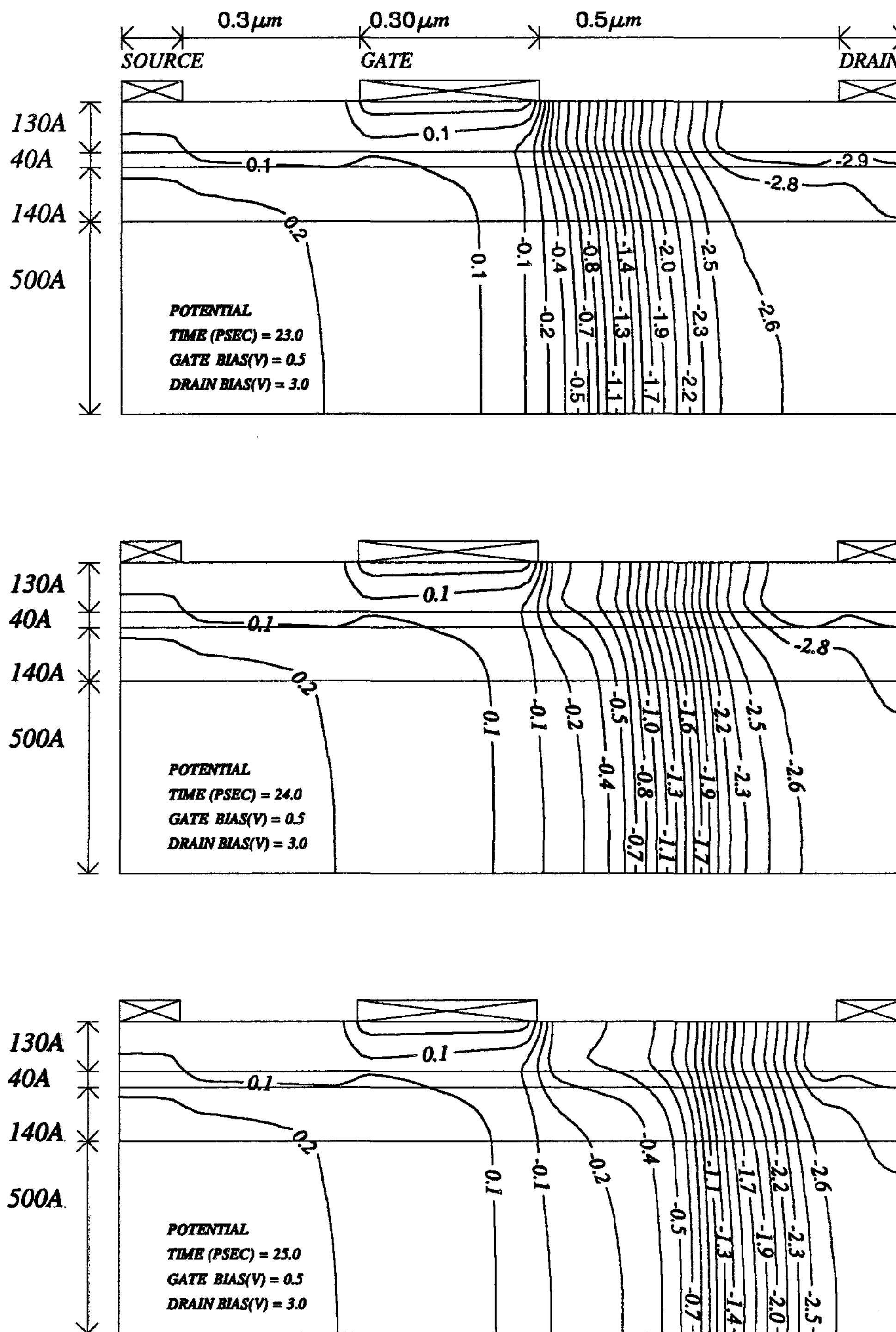


Figure III.29: Associated equipotential contours of the high energy-high field domain in a pulse doped 0.3μm *HEMT,* V_{gs}=0.5V *,* V_{ds}=3 V *at time instants 23,24 and 25 psec.*

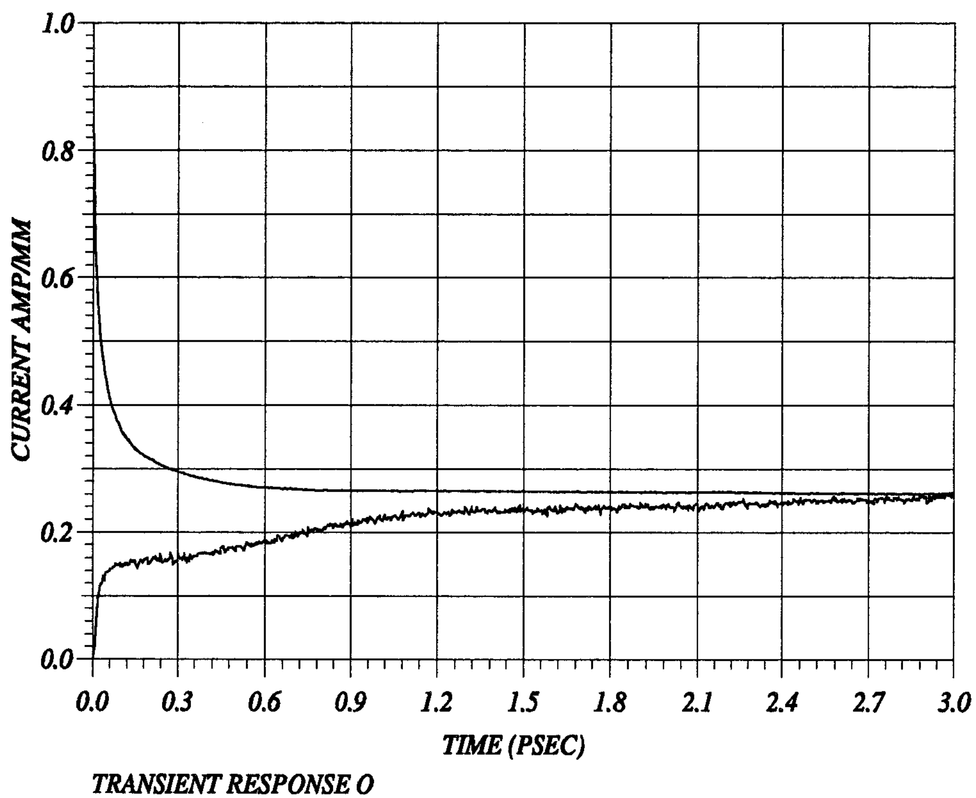


Figure III.30: Currents variation in a 0.3μm *gate pulse doped pseudomorphic HEMT with* $3 * 10^{12} cm^{-2}$ *doping level and 140Å InGaAs channel with 20% In. The drain is 3 V, Gate bias=0.3 V.*

currents at V_{gs}=.3V are shown in fig. III.30 where a stable response is found. The oscillations phenomena is then dependant on a physical quantity: the gate bias. The device remains stable for values of the gate bias as high as 0.45V, fig. III.31.

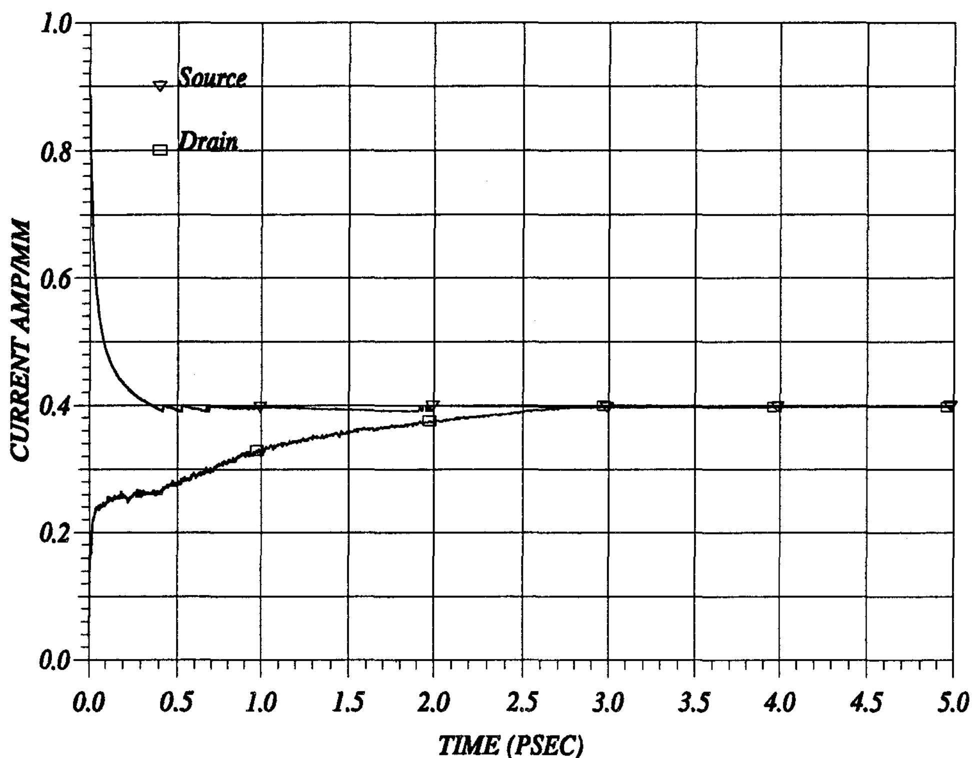


Figure III.31: Currents variation in a $0.3\mu m$ *gate pulse doped pseudomorphic HEMT with* $3*10^{12}cm^{-2}$ *doping level and 140Å InGaAs channel with 20% In. The drain is 3 V, Gate bias=0.45 V.*

The oscillations in the output current start to appear at V_{gs}=0.5 V and at drain voltage of 2.5 V. This could be very annoying for power operation as the device could be easily pushed to operate under these conditions, and in the same time it is less annoying for low noise operation where the output current is kept low.

In order to understand more what happens just before the oscillations start and why

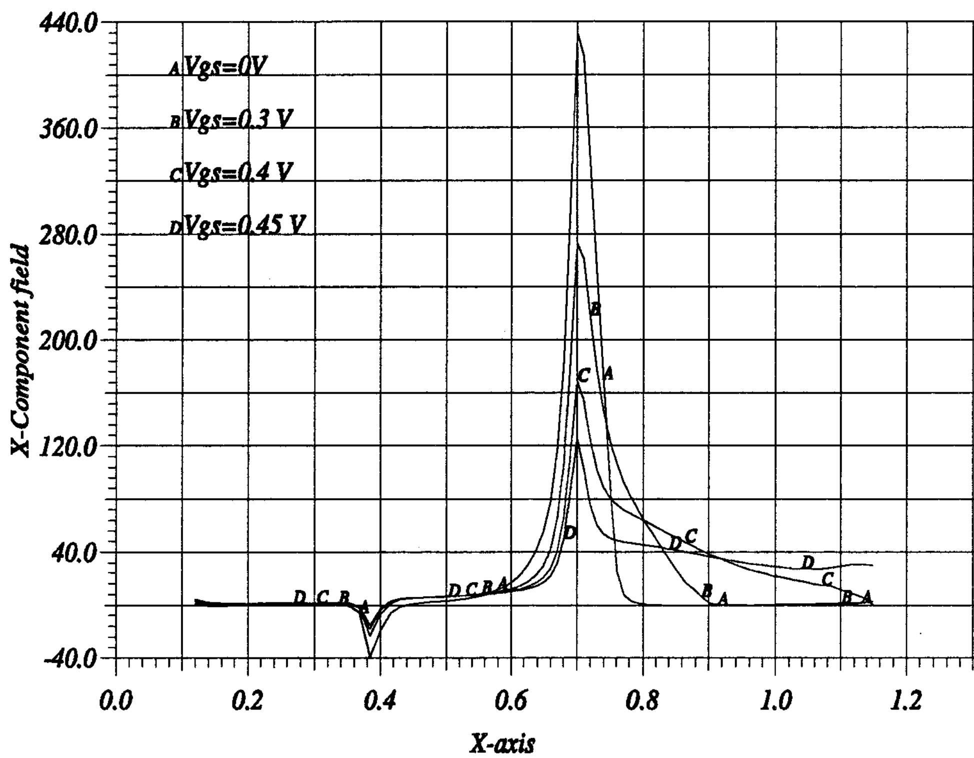


Figure III.32: The longitudinal electric field in the channel, variation in the form outside the domain when V_{gs} is increased in a 0.3 μm gate pulse doped pseudomorphic HEMT with $3 * 10^{12} cm^{-2}$ doping density.

a domain which is otherwise stable starts to move we had to thoroughly examine the physical quantities in the device. The most important quantity to observe is the electric field in the channel. We have plotted in fig. III.32 the longitudinal electric field in the channel while varying V_{gs} from values that gave stable operation (V_{gs}=0 V) till the last stable gate bias at V_{gs}=0.45 V (at V_{gs}=0.5V the device oscillates). When we observe the variation of the electric field, we find that the high field domain is located at the gate exit. The peak value of the field is smaller when the gate bias is more positive. The width of the base of the domain gets wider as the channel is more opened.

For all the cases, the field drops outside the domain to a much smaller value. This value is strictly zero except for the case when V_{gs}=0.45 V. There we notice that the field outside the domain remains uniform of a value higher than zero, exactly 39 Kv/cm. This value being higher than the sustain field for the high field domains InGaAs (or GaAs since the transport dynamics are the same cf. Chapter III) will permit the domain to start moving towards the drain once the gate voltage is slightly more positive.

From this we could conclude that the high field domain formed at the exit of the gate will remain stationary as long as the field in front of it is less than the sustain field. When the field passes this value, the domain will start travelling towards the drain creating the current pulses noticed.

III.8.3. Probability of oscillations in HEMT's

We extended our study to different HEMT structures in an attempt to know if it is possible for any HEMT to oscillate. These devices are shown in fig. III.33. We simulated two uniformly doped ($1 * 10^{18} cm^{-3}$) conventional HEMT's with different barrier layer thickness 360 and 600$\AA$, a uniformly doped ($1 * 10^{18} cm^{-3}$) pseudomorphic AlGaAs/InGaAs HEMT, a pulse doped ($3 * 10^{12} cm^{-2}$) conventional HEMT, and a pulse doped ($3 * 10^{12} cm^{-2}$) pseudomorphic AlGaAs/InGaAs HEMT with 330$\AA$ undoped barrier layer. All the devices had the same gate length: 0.3μm. The devices were set at a drain bias of 2.5V, then the gate bias was changed to obtain oscillating

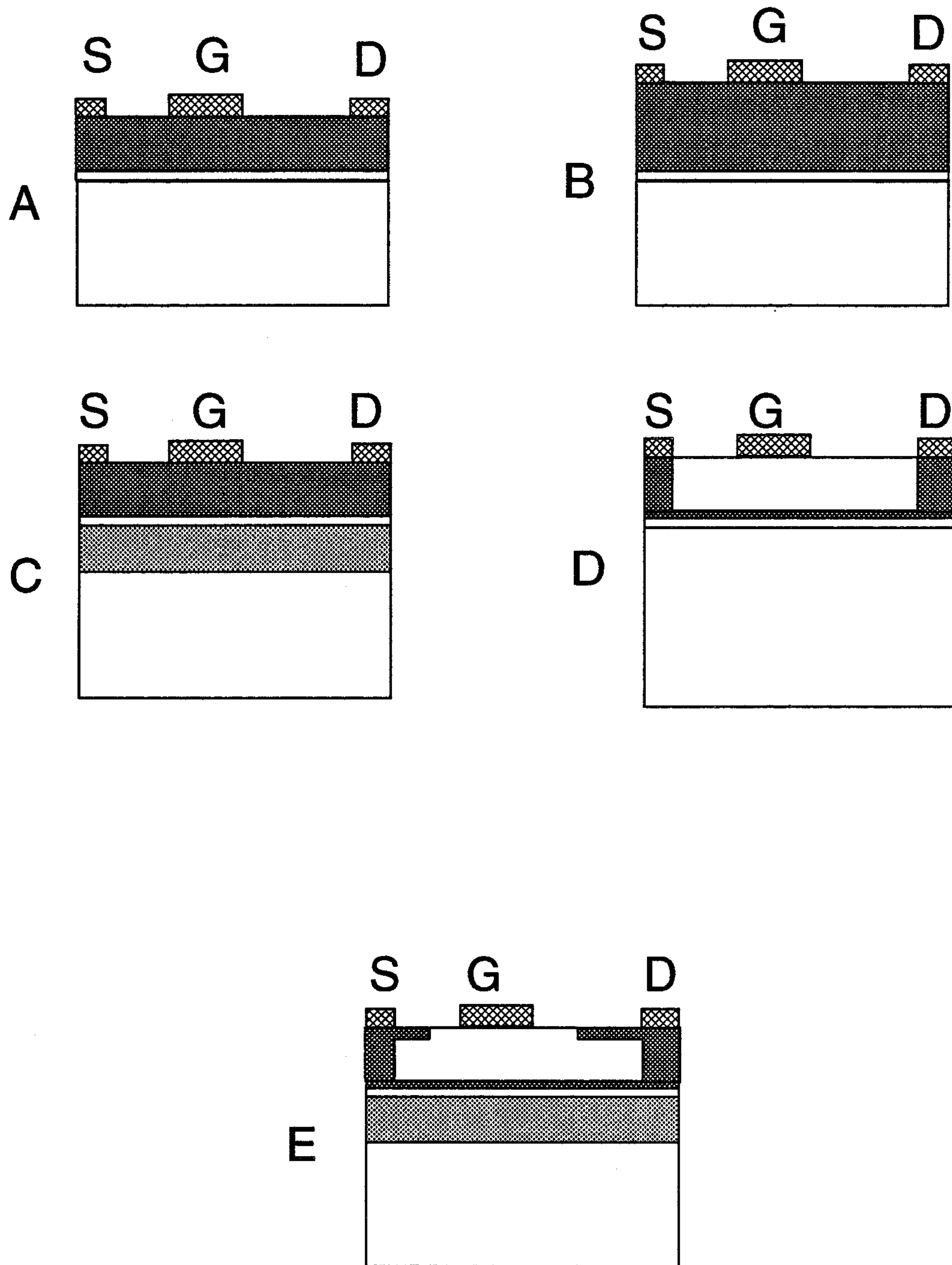


Figure III.33: Charge control of different HEMT structures, the simulated devices had $0.3\mu m$ *gate. a) is a* $360\AA$ *barrier layer, uniformly doped* ($1*10^{18}cm^{-3}$ *conventional HEMT, b)the same as (a) but with a 600Å barrier layer, c)a uniformly doped* ($1*10^{18}cm^{-3}$) *pseudomorphic AlGaAs/InGaAs HEMT, d)a pulse doped conventional HEMT* ($3*10^{12}cm^{-2}$) *and e)a pulse doped (same as d) pseudomorphic AlGaAs/InGaAs HEMT with 330Å barrier layer*

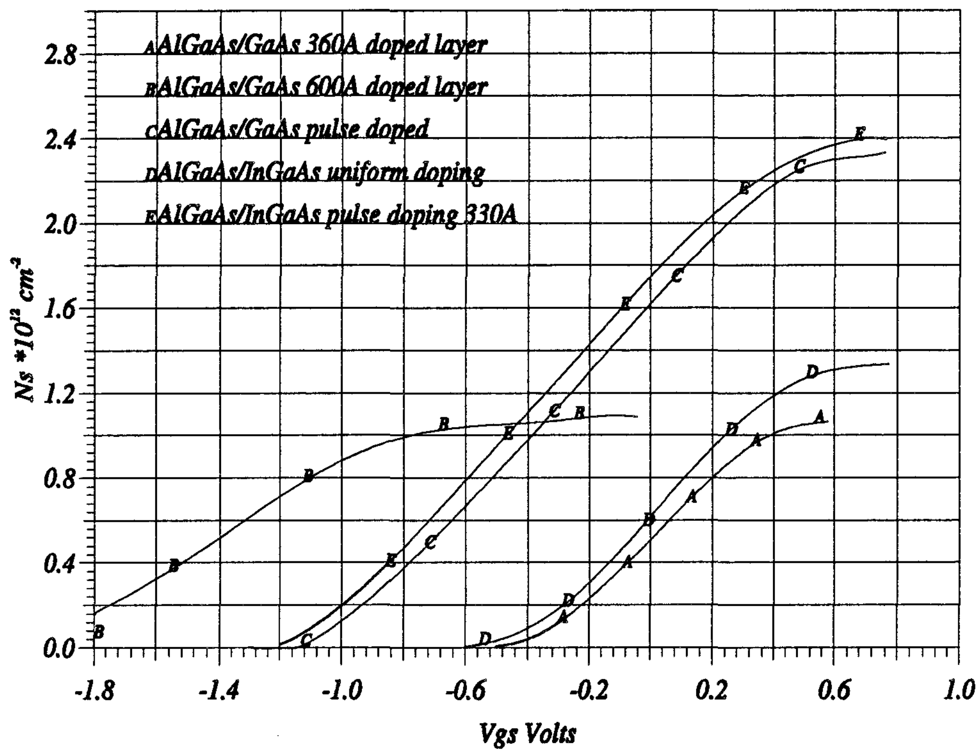


Figure III.34: Charge control of the devices described in the previous figure, drain bias is zero. Gate bias giving n_s more than $1.2 * 10^{12} cm^{-2}$ is more likely to give oscillating drain current when the drain is biased

drain currents on one hand and to determine the first stable operating point on the other hand. We were able to conclude that when concerning the charge control of the different devices, fig. III.34, *the gate bias giving an n_s greater than* $1.2*10^{12} cm^{-2}$ *is more likely to give oscillating drain current when the drain is biased.*

The maximum stable gate bias is then a function of the device structure. According to our criteria, this gives a maximum V_{gs} of 0.4 V for the conventional uniformly doped HEMT with 360Å barrier layer, V_{gs}=-0.8 V for the same type when the barrier layer thickness is increased to 600Å, V_{gs}=0.3 V for the pseudomorphic uniformly doped AlGaAs/InGaAs HEMT, and about -0.5 V for the two pulse doped structures.

If we consider the case of the pulse doped pseudomorphic HEMT, the device of interest in this work, we see that varying the thickness of the barrier layer would have a positive effect in shifting the operation voltage range. When we compare the charge control of two devices, fig. III.35, the first with 330Å undoped barrier layer and the second is with 100Å only, we see that the second device permits the rise in V_{gs} till almost 0.45 V, which was seen previously. A more detailed study of the effect of the undoped layer thickness is to follow in the next chapter.

III.8.4. Effect of loading circuit

The fact that the drain current presents time variant values means that the voltage drop across the external circuit will vary too. It is thus of less credibility to assume a fixed drain or gate bias.

In practical applications we have considered two cases:

- The source parasitic elements presented by a series resistance and inductance, these will react in such a way to alter the gate bias.
- The load circuit generally coupled to the drain, this is more complicated containing capacitive elements and will have a reaction on the drain bias if the drain current present temporal variations

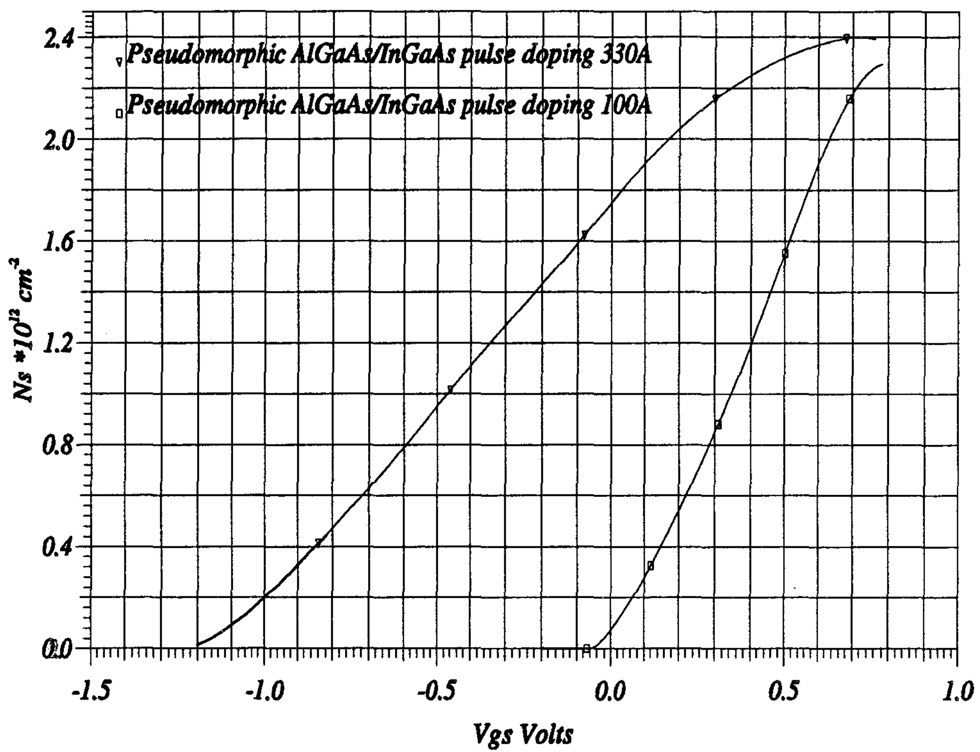


Figure III.35: Charge control of two pulse doped ($3 * 10^{12} cm^{-2}$) pseudomorphic AlGaAs/InGaAs HEMT with 330 and 100Å undoped barrier layer thickness.

The simulation of the coupled circuit using a hydrodynamic model is rather complicated due to the enormous amount of numerical noise that is produced, however we attempted to simulate the effect of the source parasitic elements through modifying the gate bias as:

$$V_g = V_{dc} - L_s \frac{\Delta I_s}{\Delta t} - I_s R_s \quad \text{(III.6)}$$

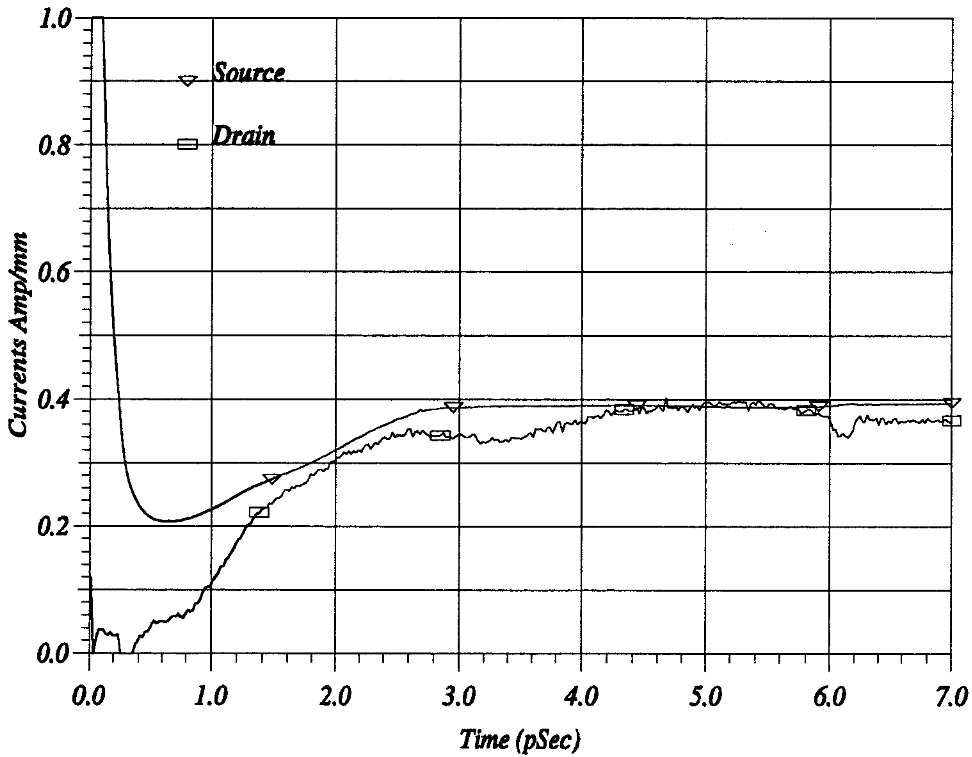


Figure III.36: Output drain and source currents in the pulse doped $0.3\mu m$ *device, the transistors is coupled to a 3Ω series resistance and 1pH series inductance, the drain bias is 3.5 V*

This formula was introduced into the simulator where the gate bias was corrected

accordingly every 0.1pSec, this small correction time step was unavoidable due to the domination of numerical instabilities. We changed systematically the values of the source resistance and inductance, small values of Rs(typically 0.5 to 2 Ohms) had no effect on the existence of the oscillations and on the other hand, high values of the inductance provoked numerical instabilities. For a high source resistance (3 Ohms) and a source inductance of 1pH.[2] and a drain bias of 3.5V, the output drain current oscillations were radically reduced, fig. III.36, although on the expense of a highly numerically perturbed drain current.

We attempted then to include the effect of a charge coupling impedance on the drain side, however this turned out to be an enormous task from the modelling point of view and is one which deserves a self standing work.

III.8.5. Discussion

We have presented the phenomena of Gunn oscillations in HEMT's where we have pointed out two criterion: the first concerning the form of the electric field in the channel and its role in causing the a priori stationary domain to travel towards the drain, and the second concerned the probability of instability associated with high n_s devices.

The presence of such oscillations is not abundant in practice, this could be attributed to several reasons:

- The fact that we simulate planar structures while practical transistors are mostly recessed gate with a highly doped cap layer. Mouatakif[37] have demonstrated that in the case of InP MISFET's, the existence of a highly doped cap layer could eliminate the oscillations. However, recent work in our laboratory where attempts to simulate fully recessed structures with cap layers confirmed the existence of travelling domains even in these structures[70]

[2] Finally we assumed these values based on a device width of 100 μm, this would be fine for the calculation of the resistance as we have then practical values. As for the inductance, values as low as 6pH could be obtained today thanks to the advanced technology, so the value we took is not a bad estimation after all.

- The commonly used techniques for characterising the transistors might not be sensible to the frequency of oscillations found, 200GHz, which could be blocked at an early stage as being a high frequency noise.

- In spite of the extremely difficult simulations of a coupled circuit, we have seen that oscillatory drain current could be stabilised if a higher source resistance is introduced. This is a very critical task as the external circuit could also act to accentuate the appearance of the oscillations.

To conclude, we managed to highlight the existence of travelling high energy domains in the gate drain region of most of HEMT structures under appropriate biasing conditions. These instabilities present an inconvenience in the use of transistors for power amplifications. Special care is to be taken in the design of transistors to avoid the zone of instabilities, this work is most suitably done through our simulations.

This page is intentionally left blank

This page is intentionally left blank

This page is intentionally left blank

Chapter IV

Optimisation of pulse doped pseudomorphic HEMT's

Abstract

In this chapter, the structural optimisation of pulse doped HEMT's will be discussed. Two major criteria will be considered. Scaling the device dimensions for better microwave performance in which the effect of the undoped barrier layer, the gate length and the recess offset are going to be discussed. Two approaches for increasing the current carrying capabilities are investigated: increasing the doping level in the fore plan and the introduction of a pulse doped buffer. The summary will present the main outlines for such an optimisation procedure.

IV.1. Optimisation of HEMT's

Generally speaking, the optimisation procedure passes by what is called scaling of the device. It is well known that the performance of submicronic HEMT's improves with reduced dimensions, specially the gate length, as we approach the ballistic transport. For low noise applications, the approach will be of entirely different nature than for the power applications, the applied voltages are moderate and the noise is the parameter that counts most. On the contrary, for power applications, the noise is of less importance and the endurance of the device in high potentials is important. Thus the approaches applied previously for the scaling of MESFET's [71], and of HEMT's [36] might not be of direct use in our case.

For power application optimal performance we need to maximise(c.f. chapter one):

- The breakdown potential.
- The output current, which together with the maximisation of the potential will result in increased output power.
- The current-gain cut-off frequency, while keeping an acceptable linearity.

The parameters that would be considered are quite numerous and we had to rely on qualitative reasoning to simplify the problem. From this aspect, we were based upon the experimental founding giving advantage in using pulse doped layers for higher breakdown potential [9]. In the following chapter we will present the primary results concerning the estimation of the breakdown potential of HEMT's.

To increase the output current, the 2DEG charge density has to be increased, for this several possibilities are available:

- Increasing the donor concentration in the doped plan of the barrier layer; this would definitely result in a threshold voltage shift which might necessitate, in the case of a real advantage in increased doping level, the modification of the layer thickness [1].

[1] As a matter of fact this what is known as the constant $N_d a$ scaling scheme

- Increasing the conduction band discontinuity height (ΔE_c), this was shown not to be of infinite interest as a saturation starts to occur in the 2DEG charge density [72].
- The use of a supplementary carrier injector as for example a doped channel or a doped buffer, the first may be less favourable (cf. chapter one), while the second is subject to investigation.

The increase in the current gain cut off frequency could be achieved through the scaling down of the device i.e. the reduction of the gate length since the f_{ci} increases when the gate length is decreased. To investigate these different strategies we were based on a basic structure and then systematically the considered parameters were changed.

IV.2. The Device Structure

The basic device structure is presented in fig. IV.1. As we pointed out before, we consider only the active part of the device. This consists of an undoped GaAs buffer layer: 500$\AA$ are sufficient since little carrier injection takes place in the buffer[36]. Above this layer, we have a 140$\AA$ of undoped InGaAs with 20% In conforming with Matthew's stability criterion. Following this layer, we have a 40$\AA$ undoped GaAs spacer with 20% Al. This gives a conduction band discontinuity of 0.31 ev. The doped plan is found after the spacer layer where basically the doping is set to $3.10^{12} cm^{-2}$ over 30$\AA$ thick layer. The superficial AlGaAs layer, non intentionally doped, follows. The Al content is always kept at 20%.

IV.3. Particularities about the Simulation

Always, aiming as realistic as possible simulations, some fine tunings have been done while simulating the pulse doped HEMT.

The relatively high volumetric doping in the doped plan ($1.10^{19} cm^{-3}$), specially when assuming complete ionisation, means that a higher concentration of positive

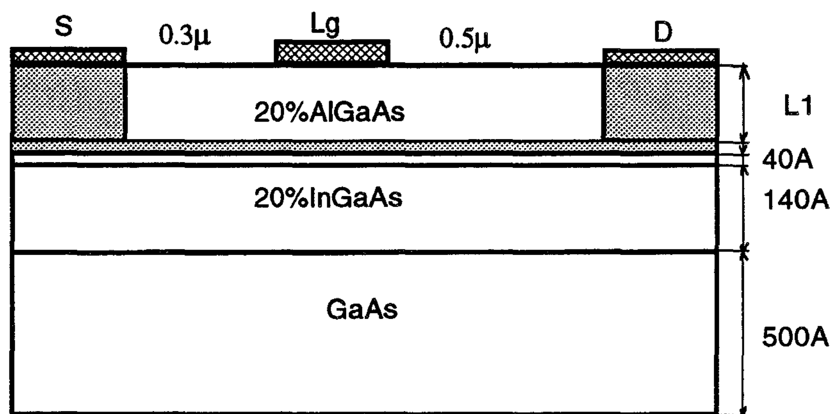


Figure IV.1: The basic structure of the pseudomorphic AlGaAs/InGaAs/GaAs pulse doped HEMT

ions is found just across the spacer layer. For electrons, essentially hot, traversing this layer, an important quantity of Coulomb interactions will occur with the ionised donors. This means that their mobility will not be exactly what is found in an undoped bulk AlGaAs, and at the same time we could not claim it is of the doped bulk. Accordingly, we had chosen to apply in this region an interpolated transport dynamics to account for this phenomenon. So, instead of the n.i.d. mobility in AlGaAs with 20% Al of 5500 $cm^2/V.sec$, we used a low field mobility of 3600 $cm^2/V.sec$.

The second aspect, concerning the doped plan, dealt with the possibility of having carriers diffusion into the undoped neighbourhood. We assumed that at a distance equal to the Debye length, given by:

$$L_D = \sqrt{\frac{\epsilon K_B T_0}{q^2 N_d}} \qquad \text{(IV.1)}$$

the charge concentration will return to the n.i.d. status, i.e. at a distance of $12\mathring{A}$. Accordingly, at a distance $L_D/2$, intermediate transport dynamics are to be applied. Since the minimum mesh step in the y direction, in this zone, is $10\mathring{A}$, we applied the interpolated dynamics at the first mesh step rather than $L_D/2$ A above the doped plan. This resulted in having a low field mobility of 4000 $cm^2/V.sec$ at a distance of $10\mathring{A}$.

In the same framework, as in [73], for the simulation of a recess structure we used the technique presented in fig. IV.2. To approach as much as possible to realistic devices, we introduced a highly doped GaAs layer in the zone to the left of the supposed edge of recess to the electrodes as well as under the electrodes. The surface potential of 0.5 V was introduced by an equivalent depletion field in the recess zone.

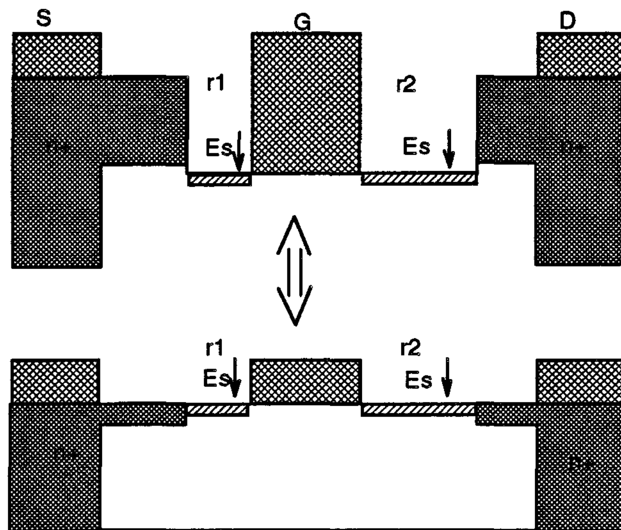


Figure IV.2: Simulation of recessed structures, a more realistic approach is applied

IV.4. Optimisation of Aspect Ratio

IV.4.1. Why should we optimise the aspect ratio?

The aspect ratio is an important factor in the optimisation of microwave FET's in general. It gives a link between the gate length Lg and the gate channel separation a, such that

$$L_g/a > A_{min} \tag{IV.2}$$

Where A_{min} is the minimum aspect ratio. It has been shown in previous work [74][75] that for conventional HEMT, the aspect ratio had to be between 5-7 otherwise, degradation in performance due to the apparition of 2D effects was noticed. Even in the case when there is no real drop in performance, the improvement was not as expected. Classically the investigation is directed first towards the gate channel separation a, however, in the case of pulse doped HEMT's this layer is no longer a single component. The different components are the undoped layer thickness, the doped plan thickness and the spacer layer. The influence of the spacer layer was extensively studied [76] and optimum value seems to be $40\AA$. At The Same Time we do not have much choice in changing the doped plane thickness since this thickness $30\AA$ indicates the use of $10^{19}cm^{-3}$ doping level which is quite high. This leaves the undoped layer thickness L_1 as the parameter that must be studied.

IV.4.2. Effect of Undoped AlGaAs Layer Thickness

IV.4.2.1. Effect on g_m

For $0.3\mu m$ gate length, we simulated four different devices with undoped barrier layers of thickness 100, 150, 180 and $330\AA$. The intrinsic transconductance, is presented in fig. IV.3 and traced against the normalised quantity (Vg-Voff) where Voff is the pinch off potential. This figure shows that there is a slight variation with the layer thickness when this is increased from $100\AA$ to $150\AA$. Further increase of

the layer thickness to 180Å results in a drop in the maximum g_m by about 20%, this drop increases to about 45% in the maximum value when the layer thickness becomes 330Å. There exists, thus, an increase in the maximum g_m with reduced

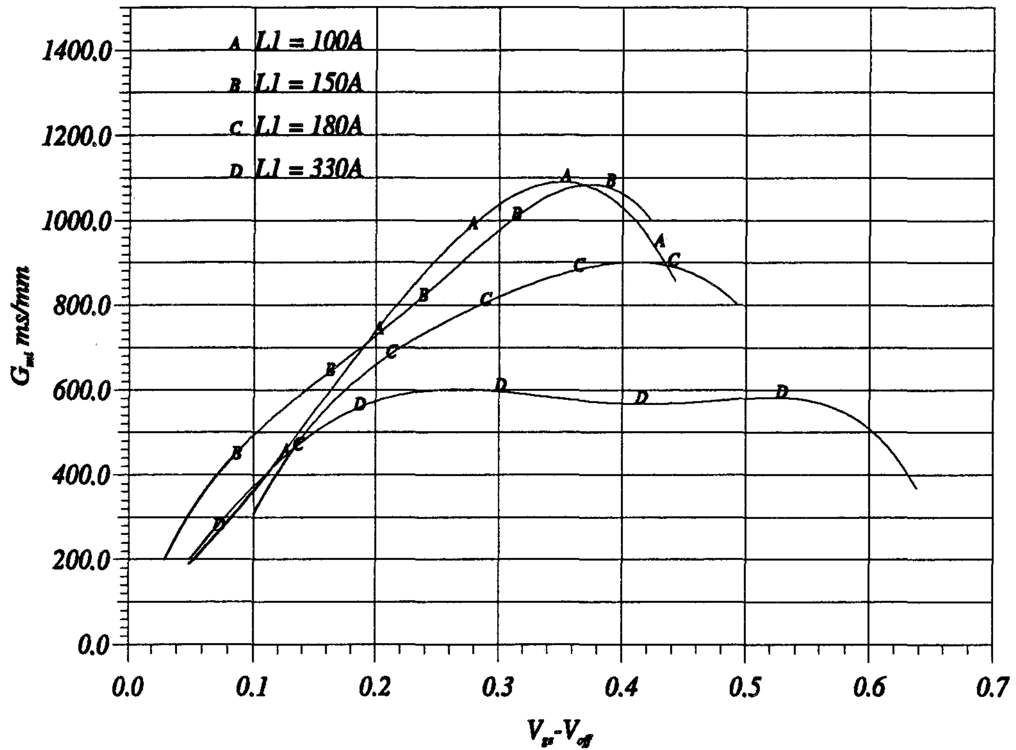


Figure IV.3: Variation of transconductance with gate bias for different thickness of barrier layer. Gate length is $0.3\mu m$*, doping level is* $3 * 10^{12} vm^{-2}$ *(*$V_{ds} = 2.5V$*)*

L_1 although on a first look this increase does not correspond linearly to $1/a$ as theoretically predicted by the formula:

$$g_m = c\frac{\epsilon < v >}{a} \tag{IV.3}$$

where c=1[77].

In the case of a pulse doped HEMT, a is not merely the separation of the gate to

the doped plan, but rather it is sum of the different layers

$$a = L_1 + L_2 + L_3 + 0.7 * L_C \tag{IV.4}$$

the term $0.7 * L_C$ gives the approximate position of the peek of the electron wave function in the quantum well ($\approx 100\AA$)[78], and

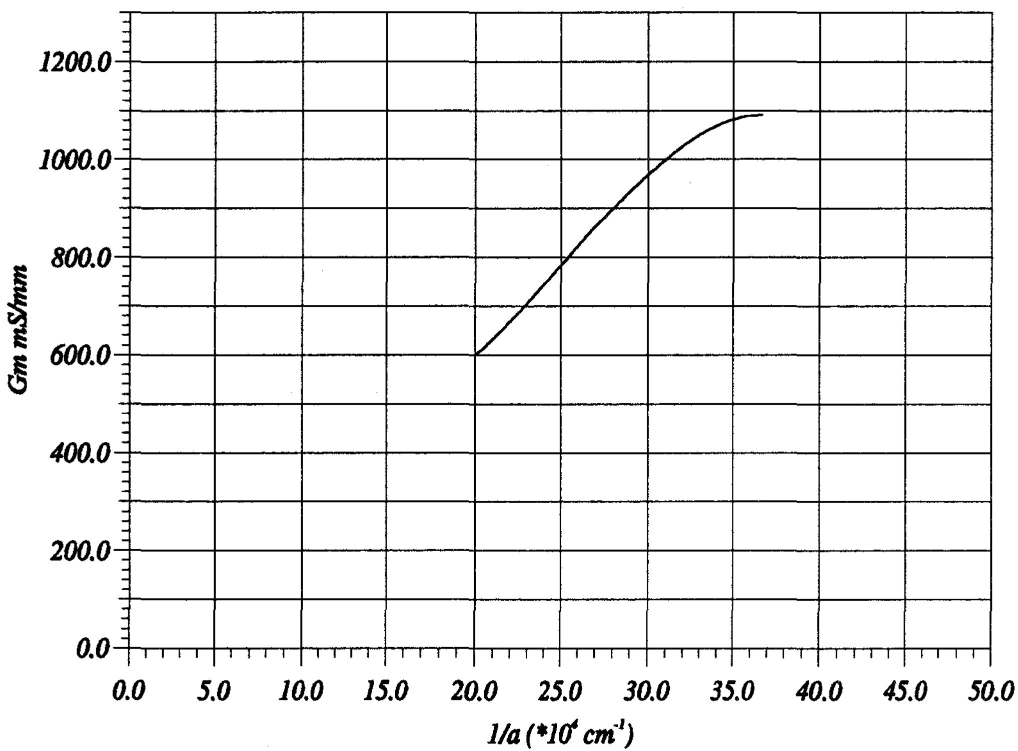


Figure IV.4: Variation of maximum Transconductance with the inverse of a in a $0.3\mu m$ *gate pulse doped pseudomorphic HEMT with doping level* $3 * 10^{12} cm^{-2}$ *(*$V_{ds} = 2.5V$*)*

IV.4.2.2. Effect on Gate capacitance

Fig. IV.5 gives the variation in the gate source capacitance when the undoped layer thickness L_1 is varied. The first observation that the maximum value increases when

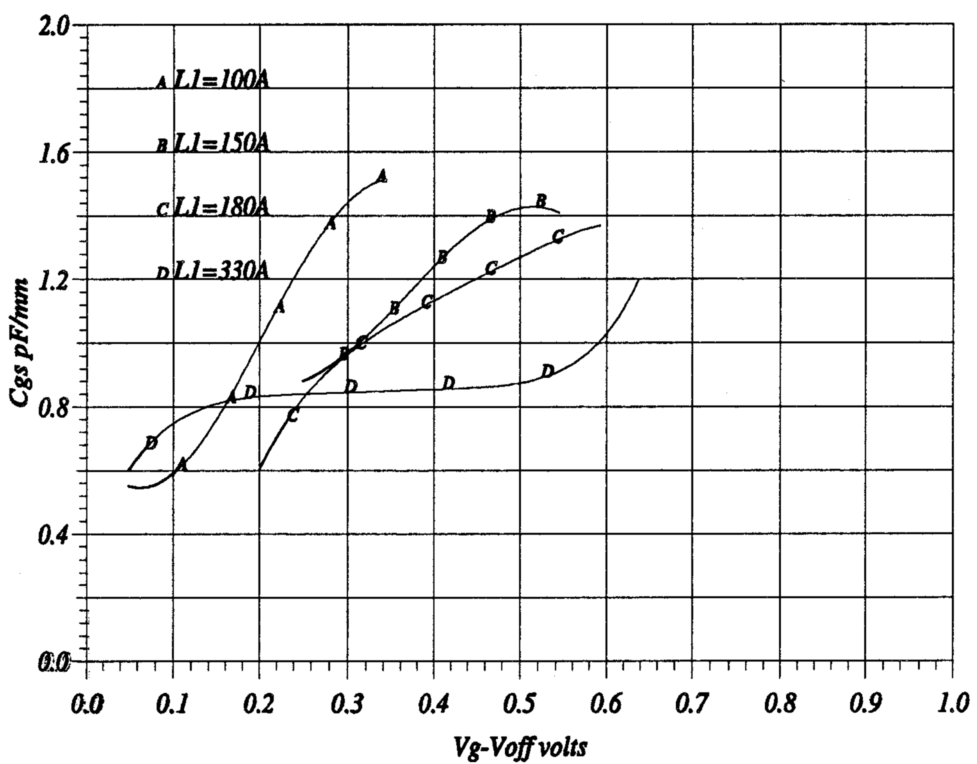


Figure IV.5: Variation of gate-source capacitance with gate bias for different thickness of barrier layer in a pseudomorphic AlGaAs/InGaAs HEMT with a $0.3\mu m$ *gate and doping level* $3 * 10^{12} cm^{-2}$. $(V_{ds} = 2.5V)$

L_1 decreases. The form of the capacitance with gate voltage is of a continuos rising for L_1 equals to 100 and 150Å, this however tends to give a less sharper rise when the thickness L_1 is increased to 180Å and then becomes more flat at L_1 equals 330Å. This could be explained in the same context of the relative position of the 2DEG, in the same way as we did in §III.4.2.

IV.4.2.3. Effect on f_{ci}

In fig.(IV.6) we see variation of the microwave figure of merit, the intrinsic current gain cut-off frequency (f_{ci}). Slight difference exists between the layers of thickness 100 and 150Å. An admirable linearity is obtained at L_1 = 150Å. We also notice that for L_1 equals to 180Å the f_{ci} starts to suffer a drop specially for open channel operation, this drop is more pronounced when L_1 is increased to 330Å.

IV.4.2.4. Recapitulation

This study was done for a pulse doped HEMT of gate length $0.3\mu m$ and a doped plan at $3 * 10^{12} cm^{-2}$. We deduced that for L_1 larger than 150Å, a slight drop in the device performance starts to appear, whereas thinner layer thickness (100Å) is of minor effect upon improving the performance. However, in the scope of using smaller gate lengths, it is favourable to use thinner L_1 layer thickness; this will also be limited by the compatibility with breakdown voltage performance. Studies in our laboratory[79] have given the minimum layer thickness to be again of about 100Å. All these factors converge to set an optimum value for the aspect ratio to be 10.

IV.4.3. Effect of Gate Length

Gate length reduction results in a double fold effect upon the device performance. First, it reduces the transit time under the gate resulting in reducing the propagation delay which is of particular interest for digital applications. Second, a reduced gate length gives rise to velocity overshoots effects due to the difference in the energy and the momentum relaxation times[80]. These effects give rise to an increase in

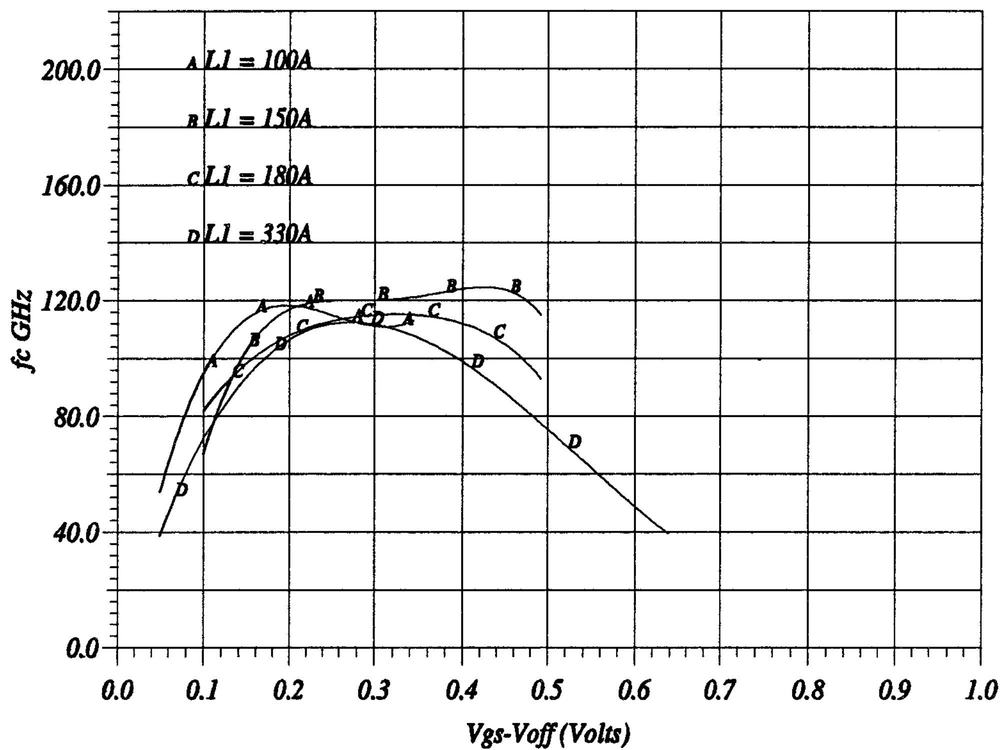


Figure IV.6: Variation of intrinsic current gain cut off frequency with gate bias for different thickness of barrier layer in a pseudomorphic AlGaAs/InGaAs HEMT with $0.3\mu m$ *gate and doping level* $3 * 10^{12} cm^{-2}$. *(*$V_{ds} = 2.5V$*)*

the average carrier velocity which in turn improves the maximum f_c that could be obtained. To understand the effect of gate length reduction, a group of four devices with different gate lengths were simulated. The devices had otherwise the same layer structure: $100\mathring{A}$ of undoped AlGaAs with 0.2 Al mole fraction, $30\mathring{A}$ doped plan at $3 * 10^{12} cm^{-2}$, $40\mathring{A}$ AlGaAs spacer, $40\mathring{A}$ of InGaAs with 0.2 In mole fraction and finally a $500\mathring{A}$ of undoped GaAs buffer. The gate lengths simulated were 0.1, 0.15, 0.2 and $0.3\mu m$.

IV.4.3.1. Effect on g_m

Fig.(IV.7) gives the variation of the intrinsic transconductance G_{mi} with gate bias when the gate length is scaled down. We observe that the threshold potential suffers a substantial shift. This is due to the reduction in the gate length when we keep the gate channel separation fixed. We notice also that there exists an increase of about 20% in the transconductance for gate length varying from $0.3\mu m$ down to $0.15\mu m$ (i.e. a 50% reduction). The transconductance drops for the $0.1\mu m$ gate length by about 25%. The increase which is found at the first reduction, as in all FET structures, is attributed to the increase in the velocity overshoot effect for shorter gates. The drop upon further reduction of Lg, on the other hand is due to the 2D effects (short channel effects), also a characteristic of FET's.

IV.4.3.2. Effect on Gate capacitance

When we examine the gate source capacitance variation with gate length reduction we notice that the maximum value presents a continuous drop when the gate length is decreased. The second observation is that the variation of the gate capacitance with the gate bias is of a sharper rise when L_g increases.

The increase in the capacitance with gate length is normal since the C_{gs} is proportional to L_g. As for the variation of the capacitance with gate voltage, it seems to occur more slowly when the gate length is reduced. This is associated with the reduction of the aspect ratio as we have seen that for small aspect ratio we tend to

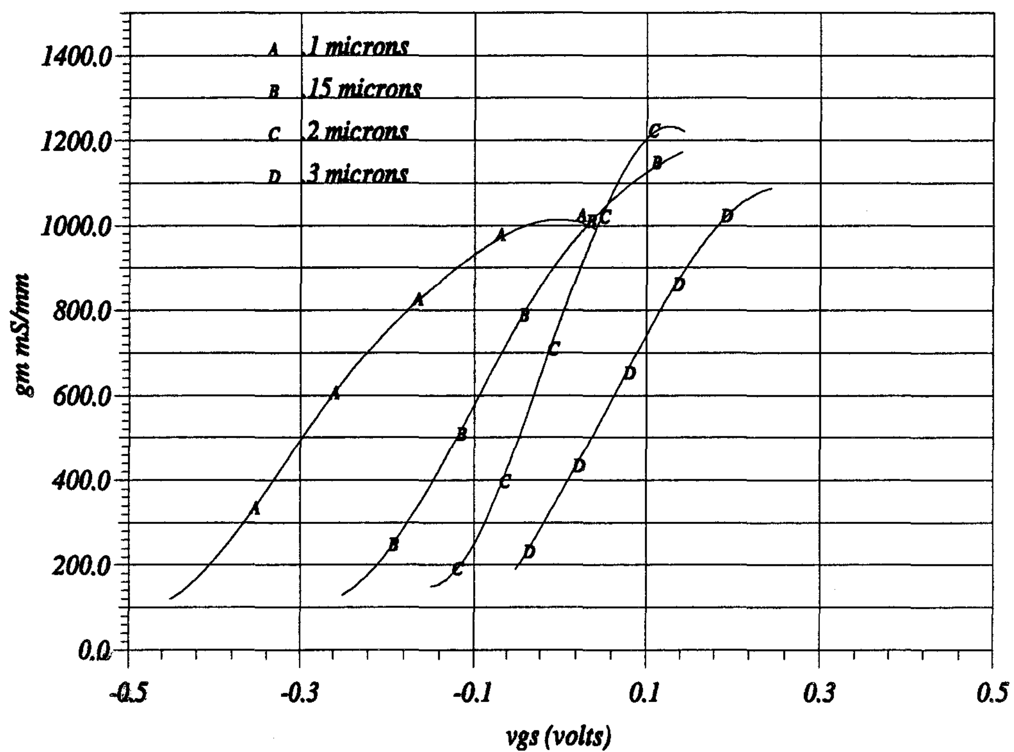


Figure IV.7: Effect of gate length reduction (scaling) upon intrinsic transconductance in a pseudomorphic pulse doped HEMT, the undoped barrier layer is kept at $100\AA$ *and the doping level is* $3 * 10^{12} cm^{-2}$

have more flatter response of the Cgs.

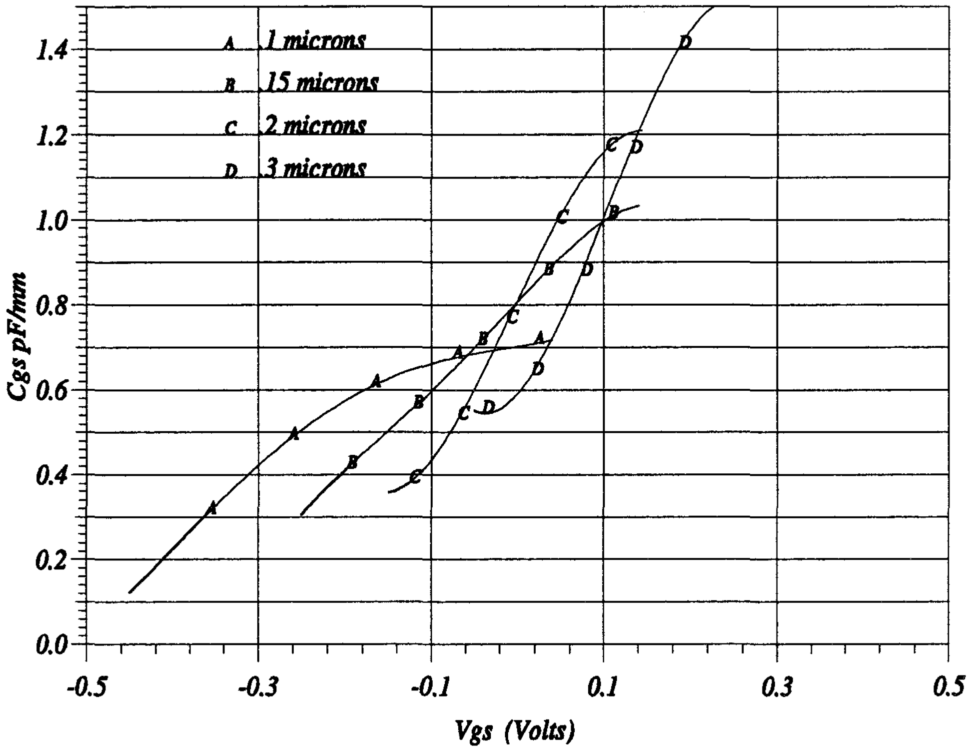


Figure IV.8: Effect of gate length reduction (scaling) upon the intrinsic gate source capacitance in a pseudomorphic pulse doped HEMT, the undoped barrier layer is kept at $100\AA$ *and the doping level is* $3 * 10^{12} cm^{-2}$

The feedback capacitance presented in fig. IV.9 presents apparently a drop for the shorter gate length, however, when the shift in the pinch off potential is eliminated by tracing the variation as a function of (V_{gs}-Voff), we find that the variation is only but small compared to the reduction in the gate length. This means that the feedback capacitance is almost independent of the gate length. The consequence of such a result is the degradation of the ratio f_{MAG}/f_c and f_t/f_c.

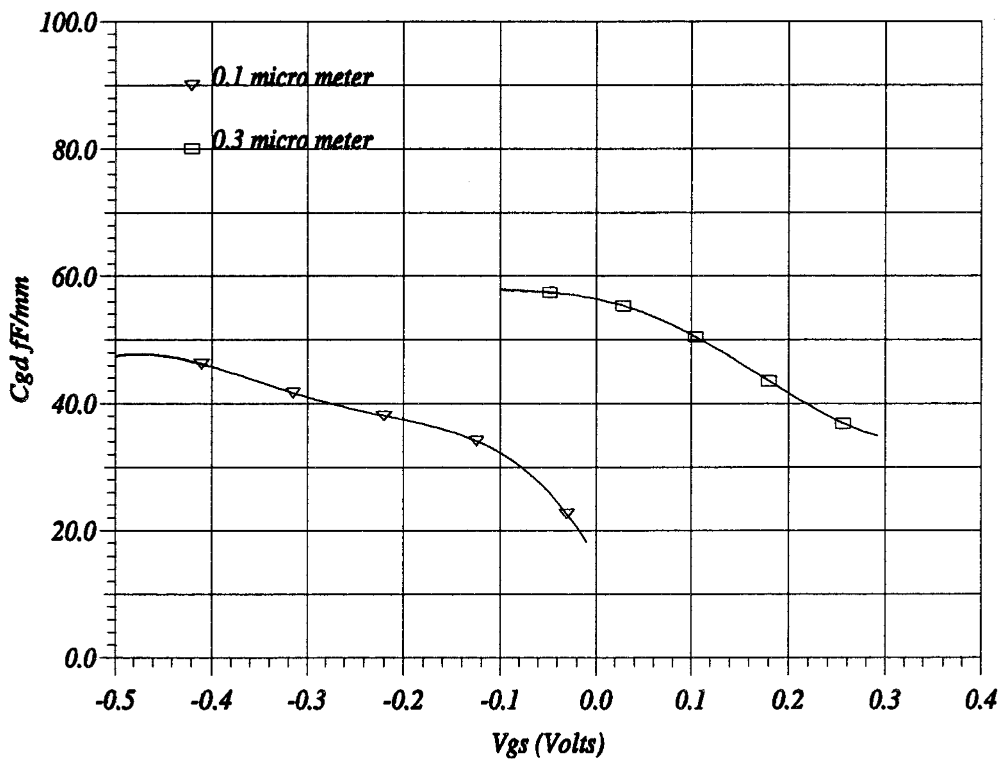


Figure IV.9: The feedback capacitance variation with V_{gs} corresponding to 0.1 and 0.3μm gate lengths, the devices are planar doped with 100Å undoped barrier layer thickness.

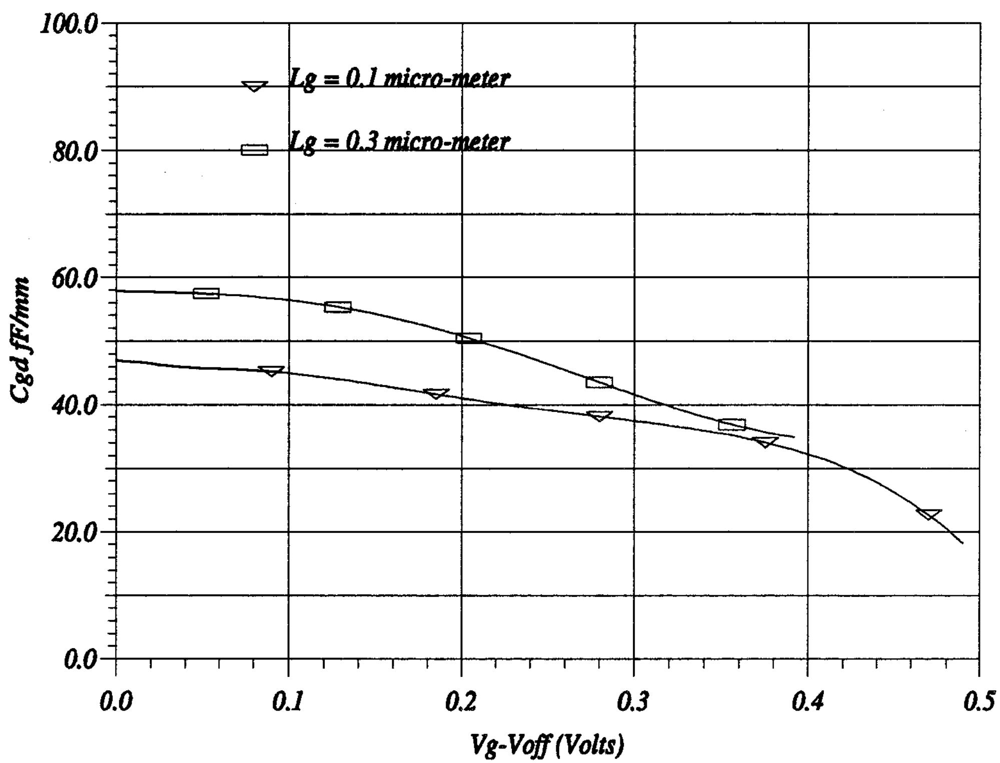


Figure IV.10: The feedback capacitance variation with V_{gs}-Voff corresponding to 0.1 and 0.3μm gate lengths, the devices are planar doped with 100Å undoped barrier layer thickness.

IV.4.3.3. Effect on f_{ci}

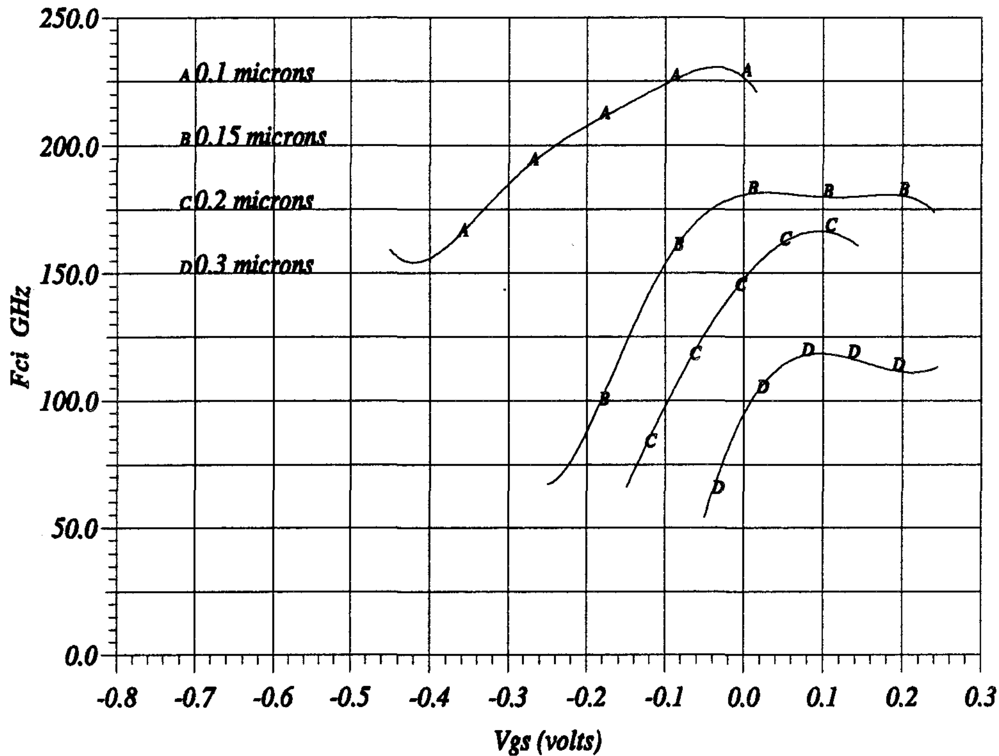


Figure IV.11: Effect of gate length reduction (scaling) upon the intrinsic current gain cut off frequency in a pseudomorphic pulse doped HEMT, the undoped barrier layer is kept at 100Å and the doping level is $3 * 10^{12} cm^{-2}$

When comparing the different current gain cut-off frequencies for different gate lengths, we found it a chance to compare our model with experimental results published. For the $0.15\mu m$ gate device, comparison with results obtained in the laboratory [81] for HEMT's fabricated at THOMSON CSF, give a satisfactory agreement of the maximum value obtained. For the other gate length, we find perfect agreement with the results found in [82], the comparison with experimental results is summarised in table IV.1.

Lg μm	exp. f_{ci} GHz	Sim. f_{ci} GHz	Ref.
0.1	250	235	[82]
0.15	173	180	[81]
0.2	150	160	[82]
0.3	120	120	[82]

Table IV.1: Comparison of maximum obtained f_{ci} with published experimental results of similar devices, all the devices are pulse doped pseudomorphic AlGaAs/InGaAs/GaAs HEMT's

As for the behaviour of f_c upon the reduction of the gate length, we have two observations. The first observation concerns the shift in the pinch off potential in the same way as for the g_m. This shift is more pronounced upon the use of $0.1\mu m$ gate obviously as a result of increased short channel effects. We observe second that the maximum value of f_c increases upon reducing the gate length. We have at $L_g = 0.3\mu m$ a frequency of 120GHz, this value rises to 160GHz when the gate length is reduced to $0.2\mu m$ with about 30% increase in f_c for a 30% reduction in L_g. Upon reducing the gate length to $0.15\mu m$, the maximum f_c increase to 180 GHz thus a 12.5% increase in the frequency for a 25% decrease in the gate length. When the gate length is further reduced to $0.1\mu m$, the maximum f_c rises to 230GHz, thus we have about 26% increase for a 30%reduction in the gate length. Here a question should be posed about the evolution of the maximum f_c when the gate length is reduced.

Elementary analytical expressions give the current gain cut off frequency as:

$$f_{ci} \approx \frac{< v >}{2\pi L_g} \tag{IV.5}$$

where $< v >$ is the average velocity of the carriers. which means that we ought to find a straight line when we trace the f_c with the inverse gate length. That is what we have done in fig.(IV.12), but, we did not find the relation to be a straight line. The plot of fig.(IV.12) implies then that the average velocity decreases when the gate length is reduced which is exactly the inverse of the overshoot effect found in the submicronic gate devices. The increase in the g_m maximum value with the reduction of the gate length is but a proof of the existence of velocity overshoot. We conclude then that the average velocity increases when the gate length is reduced.

From where comes the difference then? As a matter of fact, the simple formula

assumes that the gate length is exactly the mask length, whereas it has been shown by Monte-Carlo simulation that the two dimensional effects are more important in the case of short gates which results in having an effective gate length longer than the supposed mask length. In addition, this formula ignores the injection effect related to the shorter gate length and which result in having a different path for the electrons with a different effective velocity: the short channel effects.

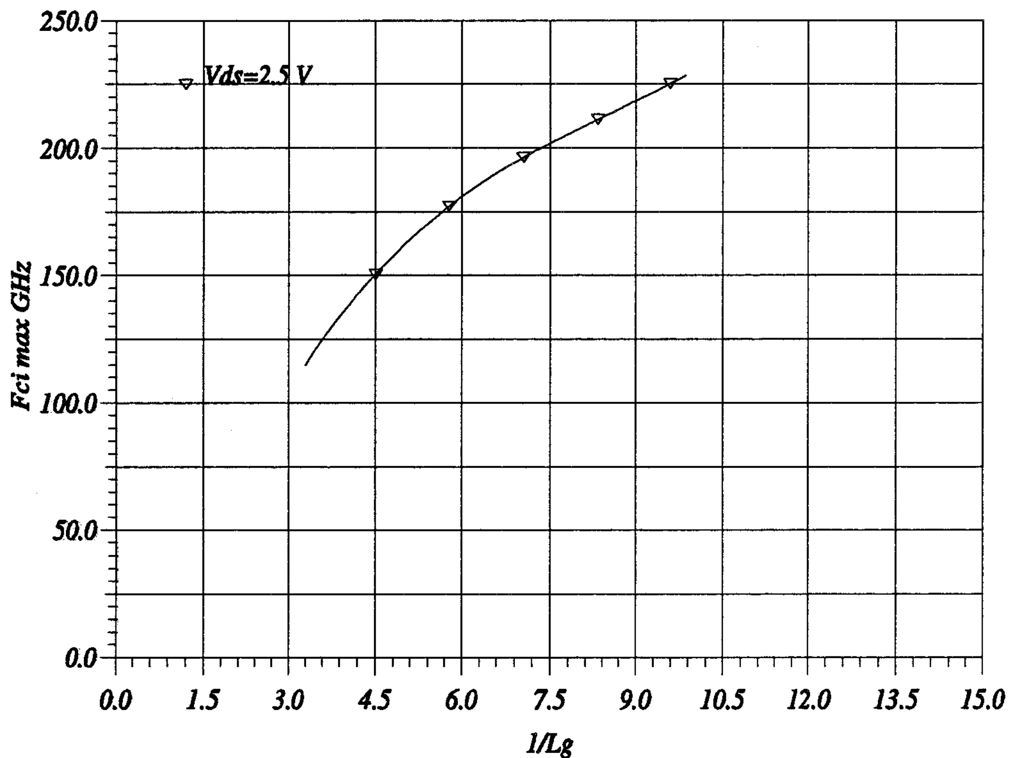


Figure IV.12: Variation of the maximum current gain cut off frequency with the inverse of the gate length in a pseudomorphic pulse doped HEMT, the undoped barrier layer is kept at $100\AA$ *and the doping level is* $3 * 10^{12} cm^{-2}$

IV.4.3.4. Effect on g_d

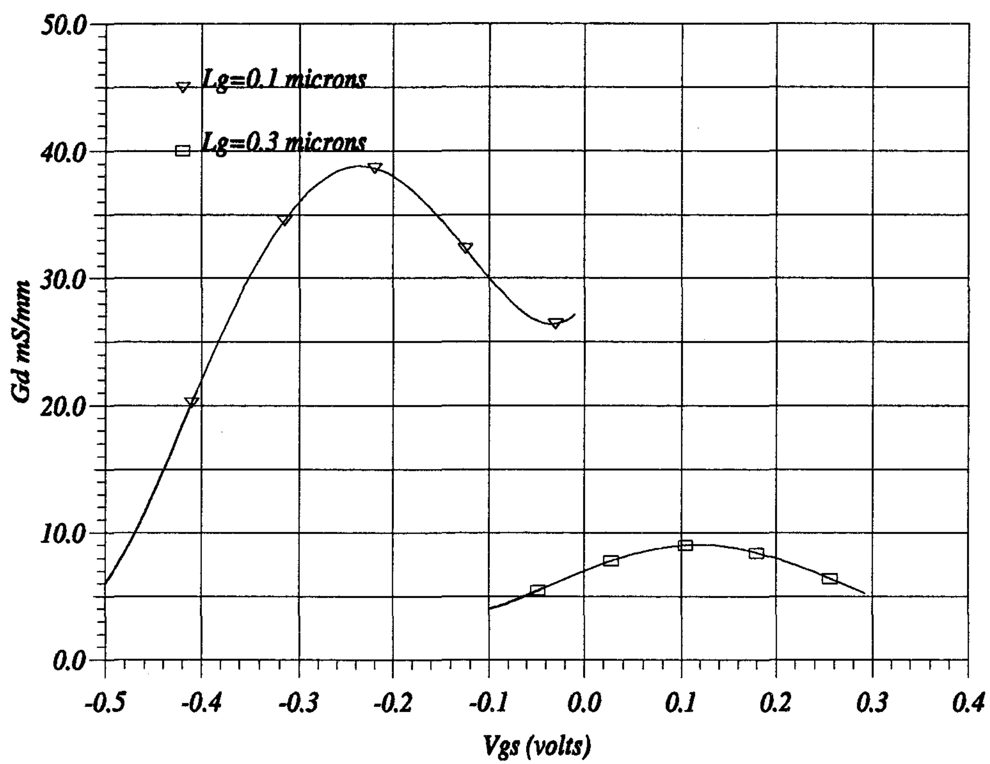
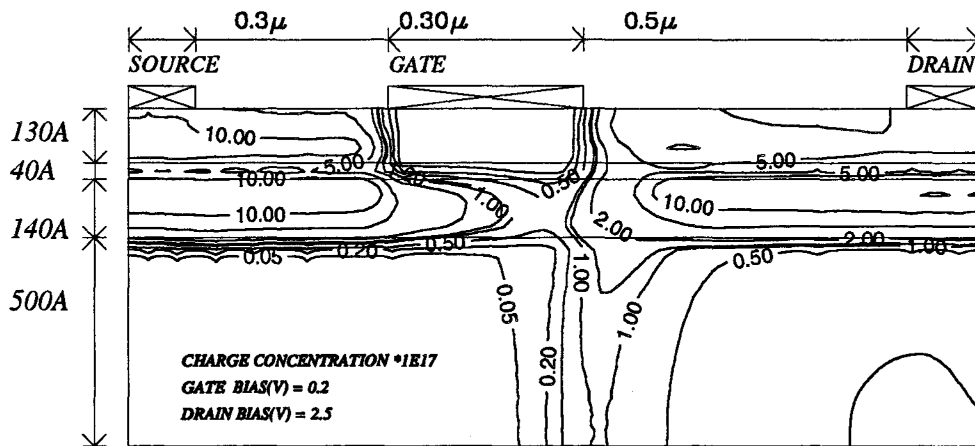
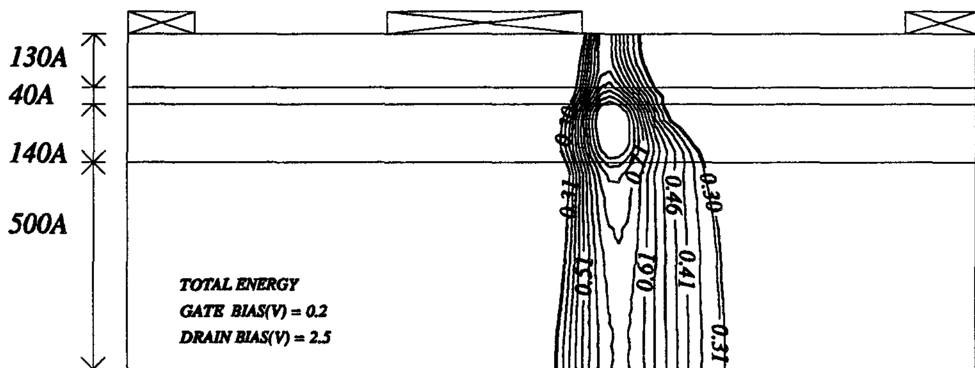


Figure IV.13: The intrinsic drain conductance of 0.3μm *and* 0.1μm *gate devices with planar doping of* $3.10^{12} cm^{-2}$ *and the same gate channel separation*





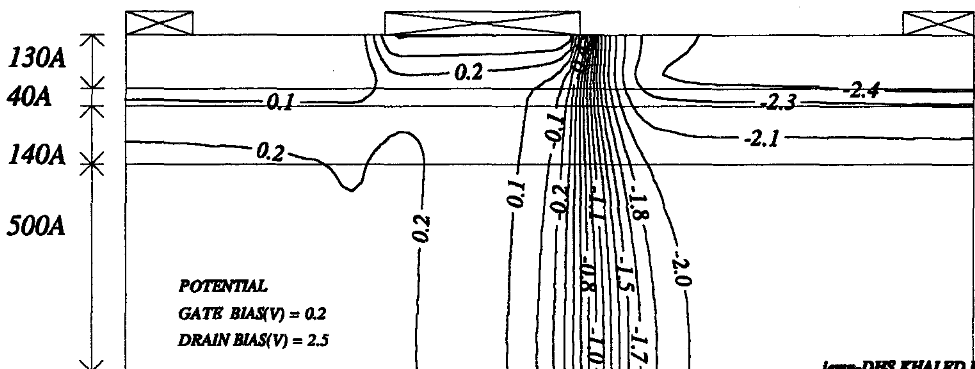


Figure IV.14: Plot of the distribution of electrons charge density, average total electron energy and electrostatic potential for a $0.3\mu m$ *gate pulse doped pseudomorphic HEMT, the fore plan doping is* $3.10^{12} cm^{-2}$

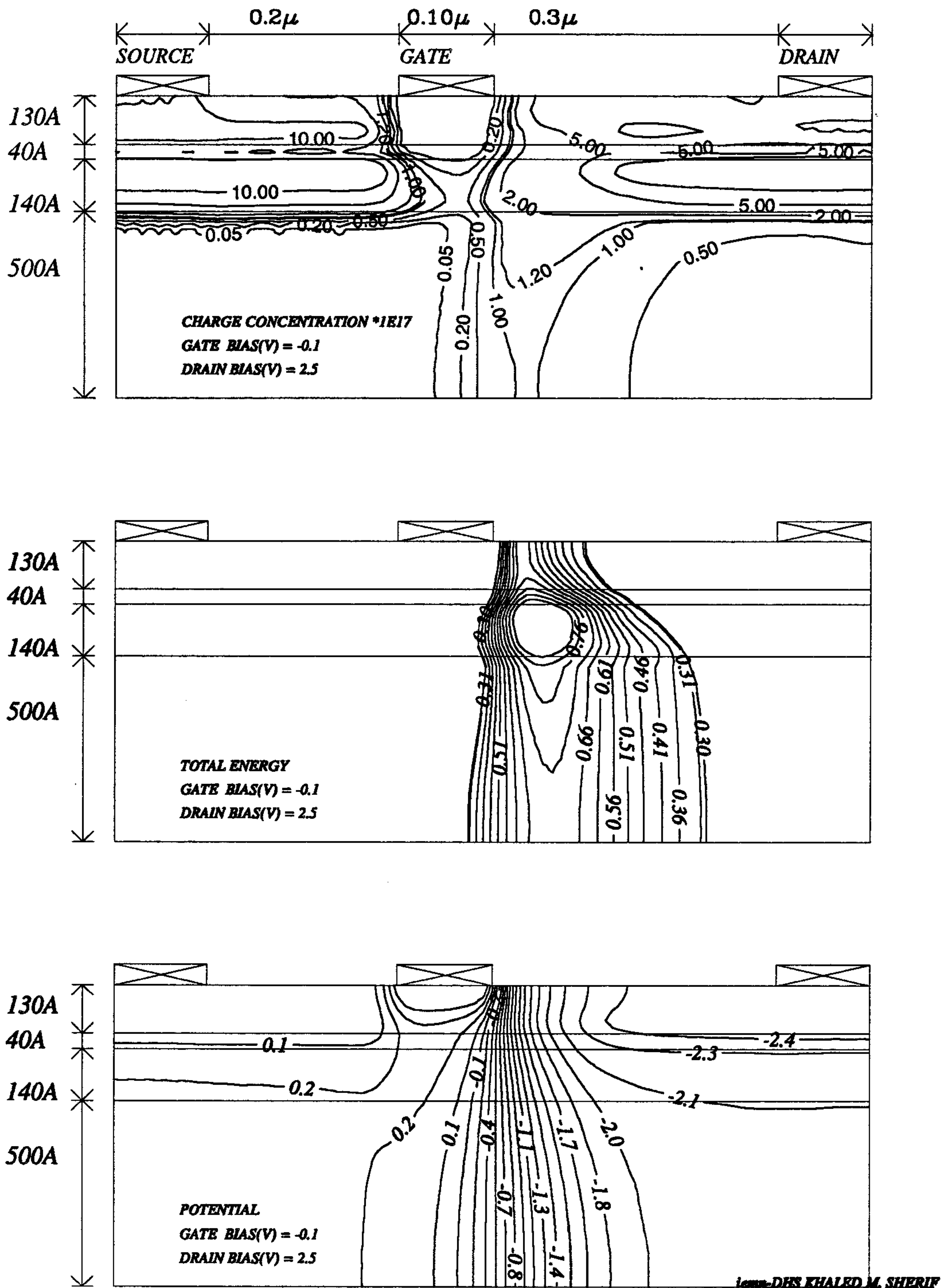


Figure IV.15: Plot of the distribution of electrons charge density, average total electron energy and electrostatic potential for a $0.1\mu m$ *gate pulse doped pseudomorphic HEMT, the fore plan doping is* $3.10^{12}cm^{-2}$

The drain conductance g_d, in addition to the gate capacitance (C_{gs}, C_{gd}), is a deterministic parameter for the calculation of the maximum frequency of oscillation f_{max} and the maximum available gain frequency f_{mag}. We have compared the variation of g_d for two gate length presenting the extremes that we have, at $0.1\mu m$ and $0.3\mu m$, fig. IV.13. A dependence of the g_d on the inverse gate length is evident from the rise in g_d when the gate length is decreased, however this dependance is by no means linear. To understand the reason for this increase in g_d, we plotted the distribution of physical quantities in the two devices at the same (V_{gs}-Voff) values, fig.(IV.15, IV.14). In fact, we could notice two things. An increased injection of carriers could be seen at the gate exit in the $0.1\mu m$ gate device. The second thing is that the high energy domain is wider for the shorter gate length and at the same time deeper (normal because of increased injection). Since the drain conductance is proportional the injection of carriers, this could be a reason for increased g_d. On the other hand, g_d is inversely proportional to the square width of the domain (the width of the space charge region), this should reduce g_d. We have here two contradictory effects where one seems to prevail. In short, the injection of carriers into the buffer could be taken as the major reason behind the increased drain conductance.

This represents the cost of gate scaling; the gain in the transconductance and the gate capacitance is unfortunately paid for by an increase in the feedback conductance.

IV.4.3.5. Recapitulation

The use of extremely short gates in pulse doped devices does not produce analytically predicted performance due to the more important short channel effects. A minimum recommended gate length would be a $0.15\mu m$ over an undoped barrier layer of $100\AA$ and a doping level of $3.10^{12}cm^{-2}$. The use of smaller gate length gives rise to undesirable short channel effects resulting in a slower rise in the f_c. We have found that for the $0.1\mu m$ gate, the drain conductance increases sharply, almost doubles when the gate length is reduced to the third and presenting even worse, an increase at low noise conditions. This will mean a drop in the M.A.G. and

degraded performances of the open circuit voltage (g_m/g_d). This suggests another approach in order to improve the drain conductance which is to optimise the gate recess dimensions.

IV.4.4. Effect of Recess shape

IV.4.4.1. Interest in recess offset

One of the most critical parameters in power FET's would be the recess offset, which is the separation between the edge of the gate and the edge of the recess, fig.(IV.16). First of all, this distance determines to great extent the access resistance R_s, increasing the recess offset on the source side results in increased resistance. On the other hand, increasing the recess offset on the drain side of the gate could result in better breakdown behaviour (we will discuss this in the next chapter) but at the same time it might hamper the device current gain cut off frequency. Thus, a close examination of the effect of this parameter is important. The influence of the gate recess offset on device performance has been studied extensively for MESFETS's[83], and for conventional HEMT's[36].

We will examine the behaviour of the current gain cut off frequency in a device where the increase in the recess offset is favoured in order to allow a high breakdown potential. Such a device will be able to support, most probably, drain bias higher than 10 V which we will take accordingly to be the limit of a voltage swing. Then we will see how another vital parameter, the f_c would behave in such a voltage swing. A $0.3\mu m$ gate device and $150\AA$ undoped AlGaAs layer, is used as a test vehicle. The channel was a pseudomorphic InGaAs with 20% In and a thickness of $140\AA$, a doped plane of $3 * 10^{12} cm^{-2}$ was introduced after the barrier layer and before a $40\AA$ spacer. A symmetric recess offset distance was introduced at each side of the gate. The distance, gate edge of recess separation, was varied between 0, $0.1\mu m$ and $0.2\mu m$, and the drain bias was varied from 2 to 10 V over 2.5 V step.

IV.4.4.2. Effect on f_{ci}

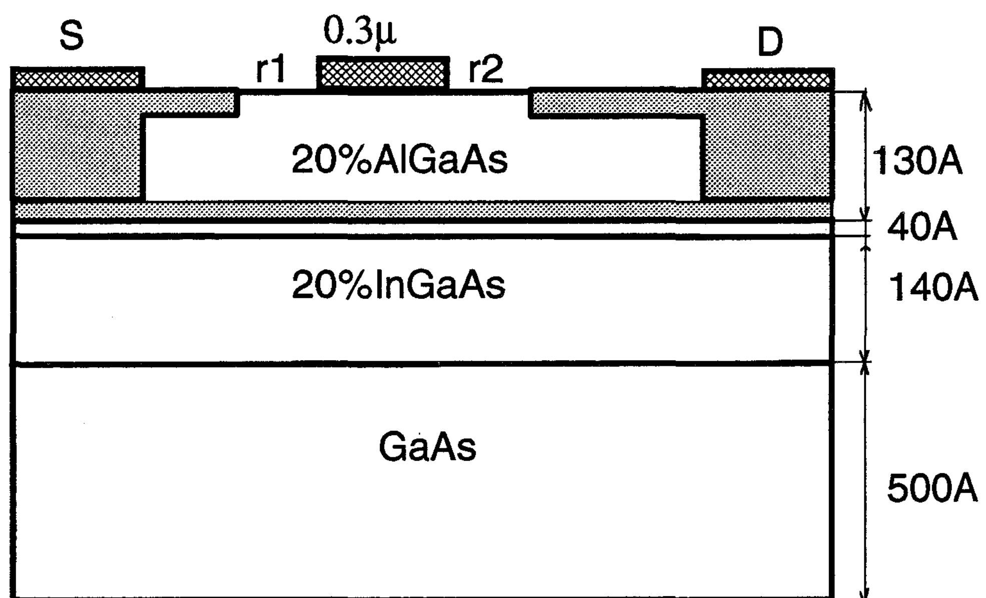


Figure IV.16: An HEMT with a recessed gate structure.

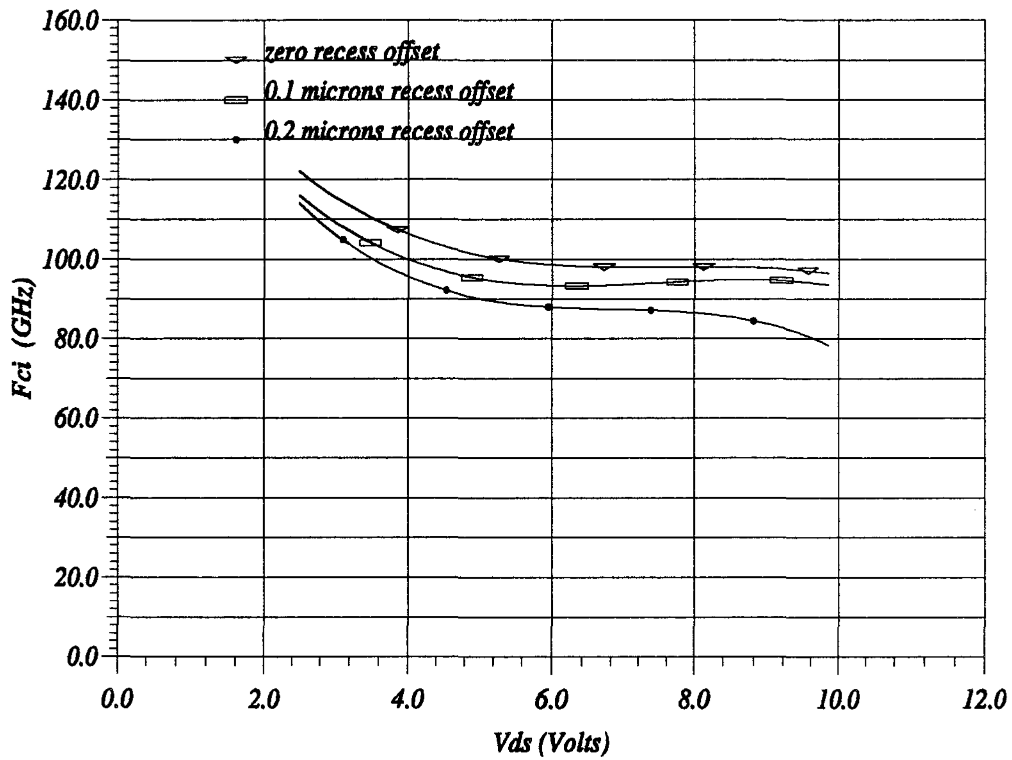


Figure IV.17: Variation of the maximum current gain cut off frequency with recess offset in a voltage swing (2-10 V) in a pseudomorphic pulse doped HEMT with $0.3\mu m$ *gate, the undoped barrier layer is kept at* $150\mathring{A}$ *and the doping level is* $3 * 10^{12} cm^{-2}$

The maximum f_c function of V_{ds} for the three devices is plotted on fig. IV.17. From this figure, two conclusions could be drawn. First, for the same V_{ds}, the f_c decreases when r increases. Second, a recess offset of $0.1\mu m$ has hardly a small effect on the f_c over the whole voltage range: at $0.3\mu m$ gate length, an offset of 30% of the gate length is of little effect upon the transport dynamics and thus the maximum f_c. At a recess offset of $0.2\mu m$, we find small changes at low $V_{ds}(< 5$ V). For higher values of V_{ds}, a drop in the f_{ci} is noticed. Although not drastic, upon the use of larger offsets ($r > 60\%$ of gate length) or for higher drain bias (the two go in the same sense), a loss in the f_{ci} is probable.

IV.4.4.3. Effect on g_d

As we previously hinted, the drain conductance would be a sensitive parameter when discussing the effect of recess offset distance. In fig. IV.18 we have the output conductance variation with gate potential for three recess offsets. Fortunately enough, the output conductance presents a drop when the recess offset is increased. This drop is enormous (almost 50%) just upon the introduction of this offset, additional drop of about 25% is obtained when the offset is increased from 0.1 to 0.2 μm. The increased recess offset results in increasing the effective gate length which reduces the injection effects at the gate exit.

IV.4.4.4. Impact on device performance

It could be understood from the previous discussion that it is of interest to use a gate recess offset of at least $0.1\mu m$ because of the improvement found in the g_d. In addition, the f_c represents practically very small variation due to this recess offset distance. The increase of the recess offset distance till $0.2\mu m$ is of great interest if the slight drop in f_c at high V_{ds} could be tolerated. The major advantage will be in the maximum frequency of oscillation which we calculated using the information presented by the simulations that we have done. Table IV.2 summarises the main characteristics of the simulated devices. We could clearly see the impact of increasing the recess offset. The maximum available gain cut off frequency f_{mag} is increased by

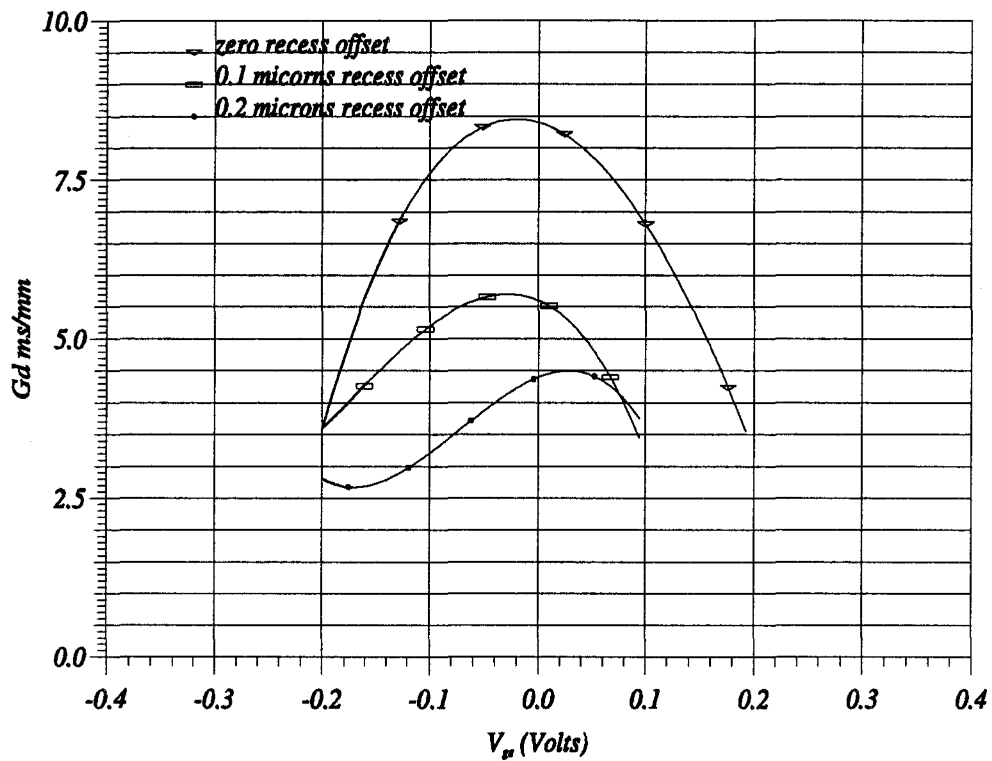


Figure IV.18: Variation of the output conductance with gate bias for different recess offsets in a planar doped HEMT of $0.3\mu m$ *gate and doping level of* $3*10^{12}cm^{-2}$*, the barrier layer thickness is* $150\AA$ *and the drain voltage is 2.5V*

Parameter	Device I	Device II	Device III
Lg (μm)	0.3	0.3	0.3
L1 ($\AA$)	150	150	150
R1 (μm)	0	0.1	0.2
R2 (μm)	0	0.1	0.2
Z (μm)	100	100	100
g_m(mS/mm)†	870	780	760
Cgs(pF/mm)†	1.403	1.302	1.305
g_d(mS/mm)†	7.6	5.2	3.2
Cgd (fF/mm)†	60	54	42
f_c (GHz)	99	95	92
f_t (GHz)	94	91	89
f_{mag} (GHz)	115	120	136
f_{max} (GHz)	144	154	176

Table IV.2: The performance of different devices with different recess offsets, the three devices are pseudomorphic AlGaAs/InGaAs pulse doped at $3 * 10^{12} cm^{-2}$, Rs=Rd=0.2Ω.mm, Ri=1.3Ω.mm, Ls=2pH.mm. (†parameters calculated at the maximum current gain cut off frequency

about 23%, and the maximum frequency of oscillation is increased by about 27%.

IV.5. Optimisation of n_s

The optimisation of the 2DEG charge concentration n_s has for goal the increase in the current level through pushing up the charge density. However this is a critical process since there would be undesirable effects arising from the increase in n_s, for example the Gunn type instabilities associated with high n_s devices (c.f. Chapter III). For this study we considered a $0.3\mu m$ gate device with a barrier layer of $100\AA$. We used two approaches, the first was increasing the fore plan doping level, and the second was introducing a rear plan in the buffer layer.

IV.5.1. Fore Plan Doping approach

We simulated the device described above using two different fore plan doping levels. The first $3 * 10^{12} cm^{-2}$ and the second $5 * 10^{12} cm^{-2}$, the maximum n_s obtained has increased from $2.3 * 10^{12} cm^{-2}$ to $2.9 * 10^{12} cm^{-2}$. The current variation with the normalised gate potential is shown in fig. IV.19. The maximum available current is almost doubled upon the increase of the planar doping level by 60%. However,

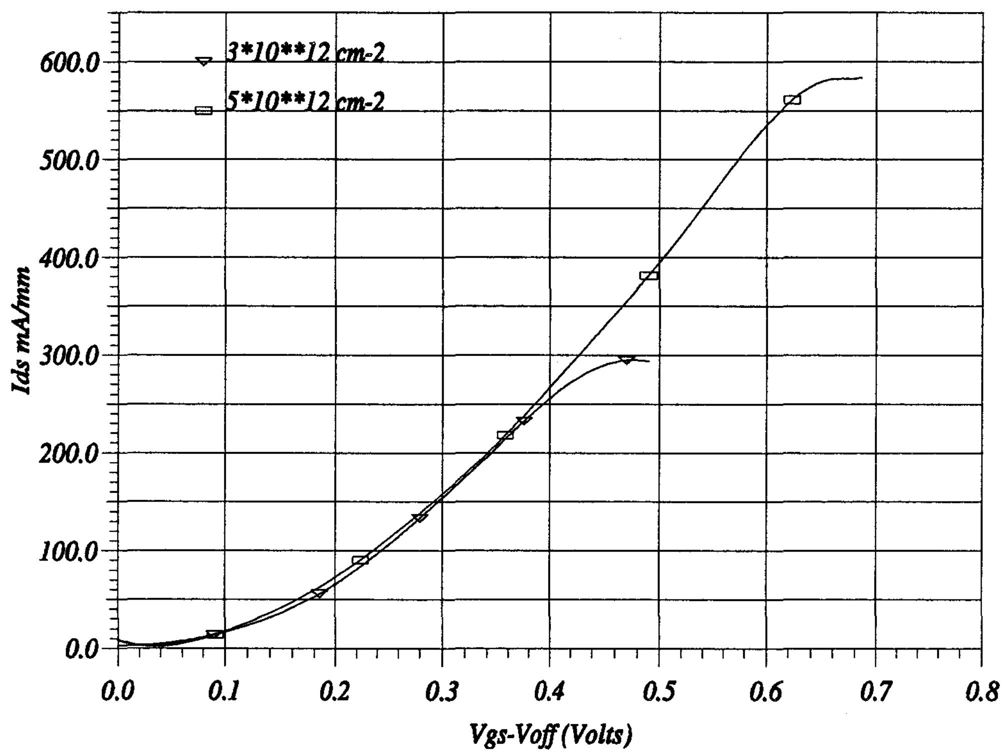


Figure IV.19: Variation of the drain current in a $0.3\mu m$ *gate pulse doped HEMT with two different doping levels. The drain bias is 2.5 V*

the behaviour of the current with gate voltage is almost unchanged. The maximum value of the current gain cut-off frequency, fig.(IV.20), is almost unchanged, yet the variation of fc with I_{ds} suggests that the higher doping level would provide a flatter response of the device. This is a very important advantage for power applications.

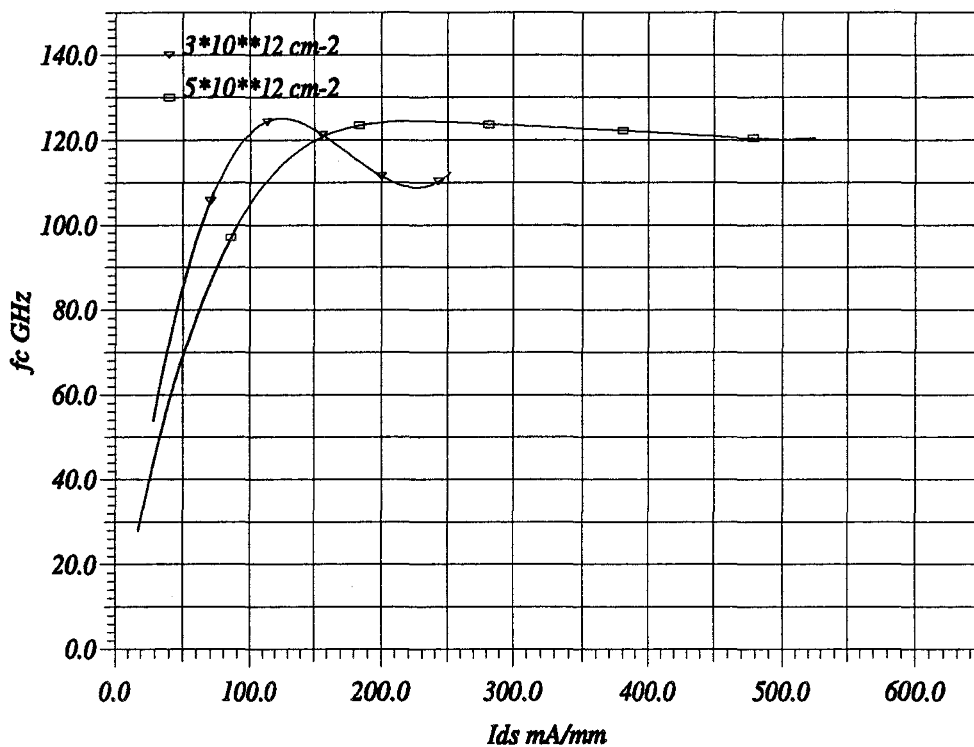


Figure IV.20: Variation of f_c in a $0.3\mu m$ gate pulse doped HEMT with two different doping levels, the barrier layer is kept at $100\AA$. The drain bias is 2.5 V

IV.5.2. Rear Plan approach

Another possibility to increase the n_s is to use a pulse doped buffer. We simulated two devices of fixed fore plan doping of $3.10^{12}cm^{-2}$ and two different doping levels

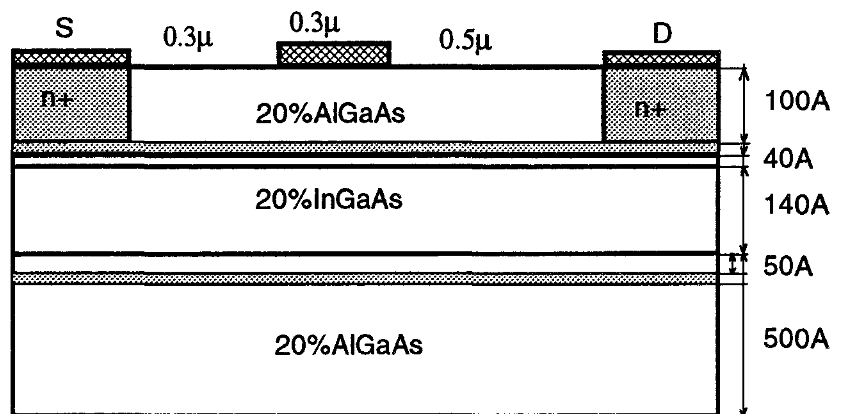


Figure IV.21: A planar doped HEMT with a rear plan doping; gate length $0.3\mu m$*, barrier layer thickness* $100\AA$

in the buffer $1.5.10^{12}cm^{-2}$ and $2.5.10^{12}cm^{-2}$, fig.(IV.21). The buffer material, again following practical guide lines, was changed to n.i.d. AlGaAs with 20% aluminium content, the doping plane in the buffer was placed after a spacer of $50\AA$ below the hetero-interface and its thickness is taken to be $30\AA$. The rest of the buffer was kept unchanged. The maximum n_s increased from $2.3 * 10^{12}cm^{-2}$ for the device with a single doped plane, to $4.2 * 10^{12}cm^{-2}$ upon introducing a rear plane of doping $1.5 * 10^{12}cm^{-2}$, and then increased to $4.6 * 10^{12}cm^{-2}$ when the rear plan doping is increased to $2.5 * 10^{12}cm^{-2}$.

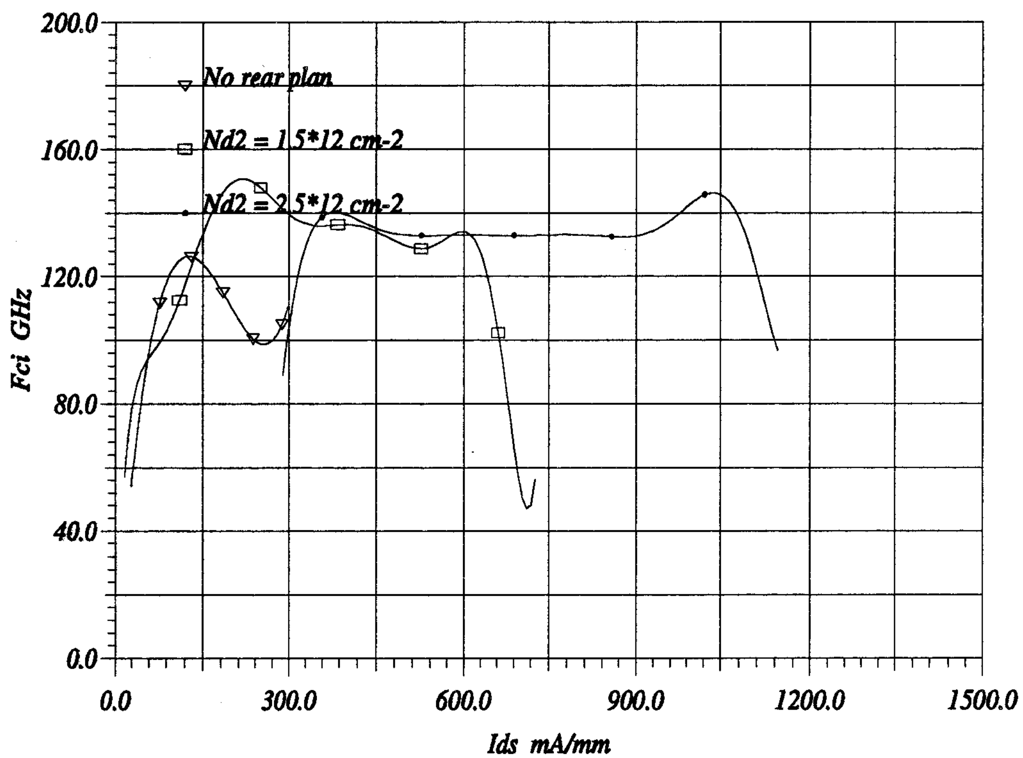


Figure IV.22: Variation of the current gain cut off frequency with the drain steady state current for pulse doped devices: effect of buffer doping while having a fore plan doping fixed at $3 * 10^{12}cm^{-2}$ and a $0.3\mu m$ gate

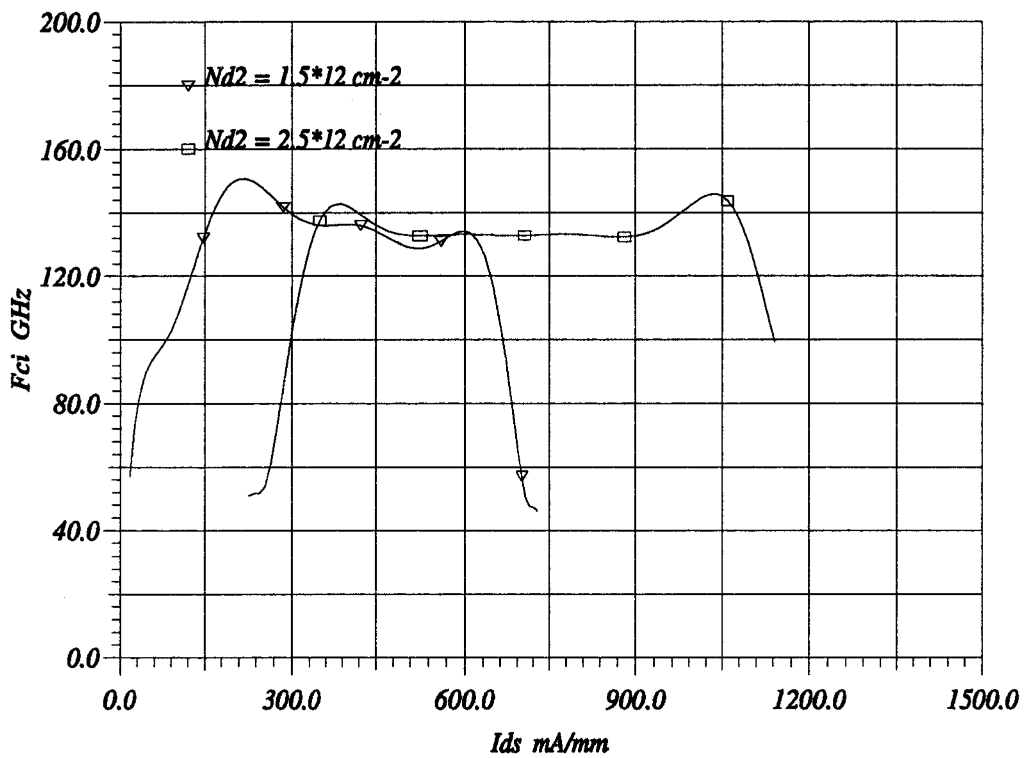


Figure IV.23: Variation of the current gain cut off frequency with the drain steady state current: effect of buffer doping level (replot of the previous figure)

Fig. IV.22 gives the current gain cut off frequency versus the steady state drain current for three devices, the two with the rear plan and the one with only a single plan doping. This last device highlights the importance of adding a rear plan when we notice the difference in the shape of $f_{ci}(Iss)$. For the devices with rear plan, a wider plateau of f_c indicates better linearity. The maximum current level possible increased from about 300mA/mm for the device with no rear plan, to 740mA/mm for the rear plan doping of $1.5 * 10^{12} cm^{-2}$, and then to 1160mA/mm for the device with $2.5 * 10^{12} cm^{-2}$ rear plan doping. We have thus more than 100% increase upon the introduction of the rear plan while 75% increase in its doping gave about 50% increase in the maximum current level.

Between the two devices with rear plan (replotted in fig. IV.23 for clarity), we find that a higher rear plan doping level does not add much to the maximum f_c. On the other hand, we notice a sharp drop in the frequency for the low current region of the characteristics. For the first device, we have a maximum f_c over a range of 150-600 mA/mm of the output current. The other device shows a dramatic drop at currents as high as 300 mA/mm. This means that in microwave operation; a drop in efficiency will be encountered at higher rear plan doping.

To further understand the reason of this drop, we examine the distribution of charge density in the two devices, fig. IV.24 for the device with low rear plan doping and fig. IV.25 for the other device with high rear plan doping. For higher rear plan doping, a large portion of the carriers are in the AlGaAs buffer at negative gate bias. This means that carriers with less favourable transport dynamics are contributing to the current transport which will directly result in lower f_{ci} at a given current.

IV.5.2.1. Recapitulation

In table. IV.3 we summarise the different devices considered in the n_s optimisation study.

In conclusion, the introduction of a doped buffer layer results in both improving the maximum value of the current, and increasing the linearity of the device by having

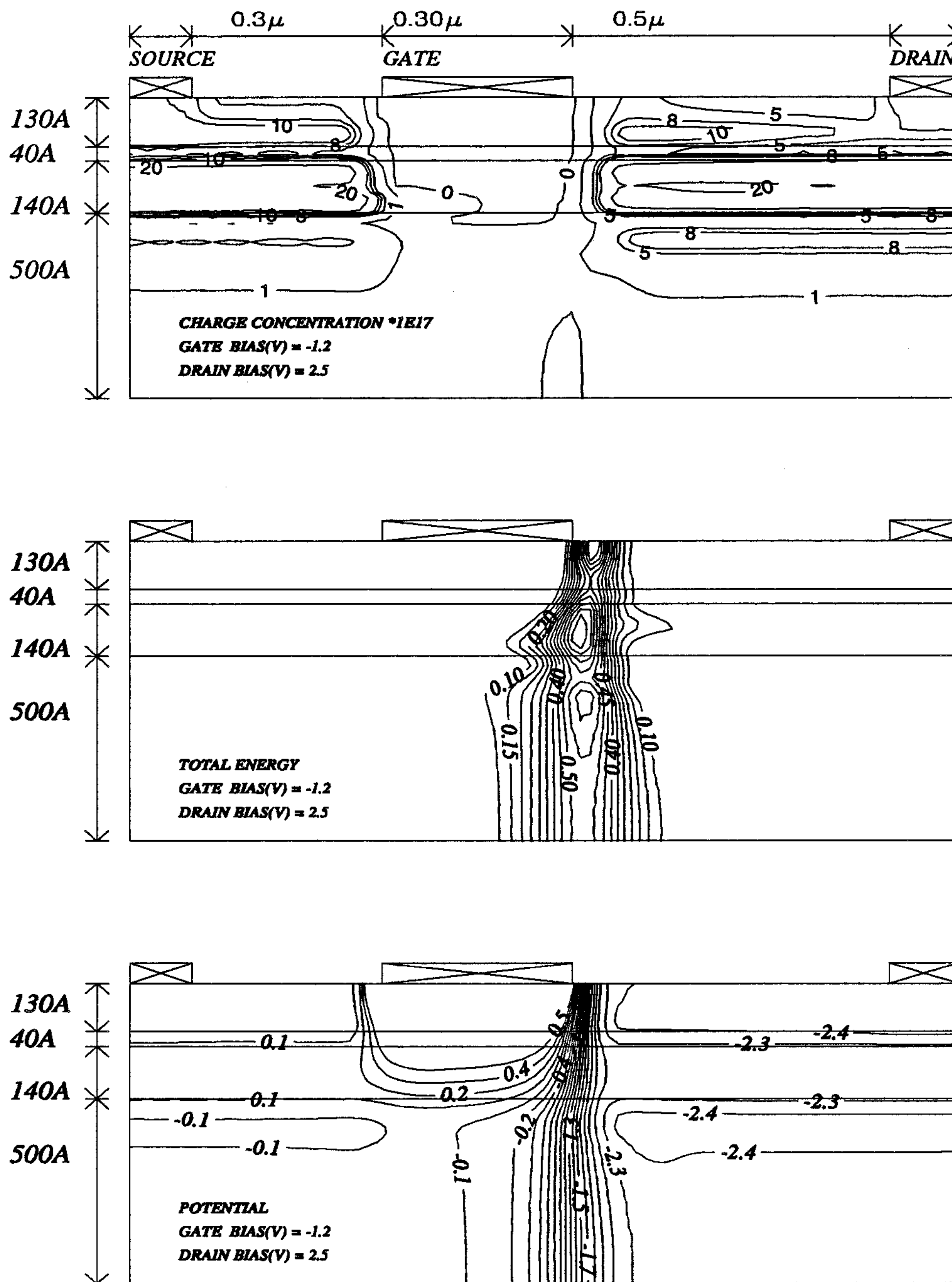


Figure IV.24: Plot of the distribution of electrons charge density, average total electron energy and electrostatic potential for a buffer doped pseudomorphic HEMT, the fore plan doping is $3.10^{12}cm^{-2}$ *and the rear plan doping is* $1.5.10^{12}cm^{-2}$

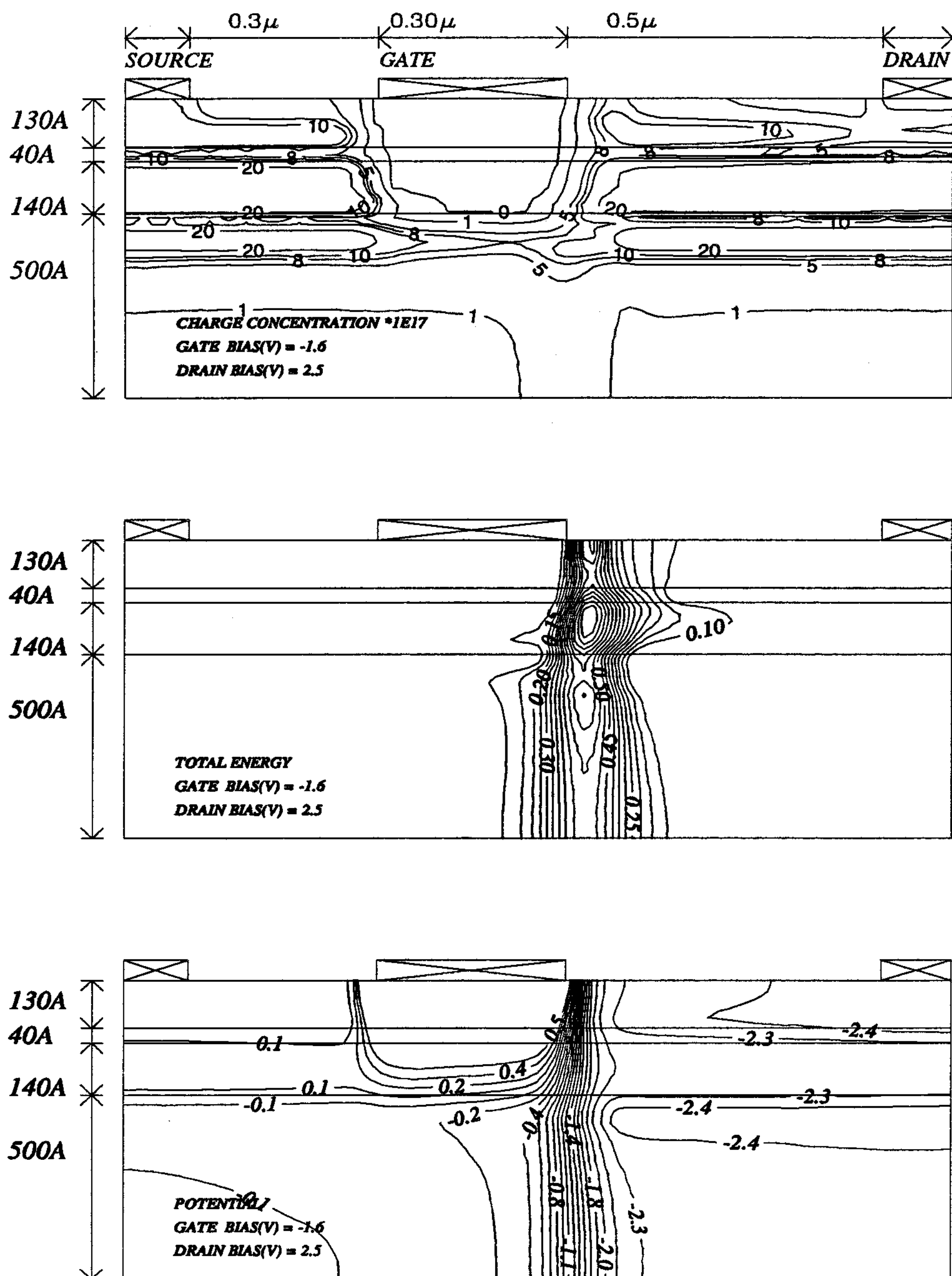


Figure IV.25: Plot of the distribution of electrons charge density, average total electron energy and electrostatic potential for a buffer doped pseudomorphic HEMT, the fore plan doping is $3.10^{12}cm^{-2}$ and the rear plan doping is $2.5.10^{12}cm^{-2}$

Fore Plan $*10^{12}cm^{-2}$	Rear plan $*10^{12}cm^{-2}$	$n_s$max. $*10^{12}cm^{-2}$	I max mA/mm	f_{ci}max. GHz	Range of f_{ci}max mA/mm
3	—	2.3	300	120	100-150
5	—	2.9	580	120	150-500
3	1.5	4.2	740	135	150-600
3	2.5	4.6	1160	135	350-1000

Table IV.3: Characteristics of the different devices with different planar doping levels. The gate length is $0.3\mu m$*, the undoped barrier layer is* $100\mathring{A}$*, the current and the current gain cut off frequency were obtained for* V_{ds}*=2.5 V*

a flatter response of the current gain cut off frequency. The increase in the value of buffer doping might not necessarily mean improved performance due to the presence of carriers in the donor layers resulting in mediocre efficiency. We managed to point out a quasi optimum value of the rear plan doping to be about $2 * 10^{12}cm^{-2}$.

IV.6. Summary

In this chapter we discussed some structural optimisation approaches for obtaining better microwave power performance of pseudomorphic HEMT's. The major outlines for device scaling indicate that short channel effects begin to have considerable effects for an aspect ratio (L_g/a) less than 10 for a $0.3\mu m$ gate length. In order to avoid such effects we have to keep the undoped barrier layer thickness L_1 small but not extensively small as no improvement is found, this is to be done within the compatibility limits of the breakdown potential. The value that we found to be suitable is $100\mathring{A}$ which gives a gate channel separation of $180\mathring{A}$.

If we want to get the best performance, the gate length is to be kept at $0.15\mu m$, smaller values could result in higher f_c but will cause strong degradation in the performance specially in the factor (g_m/g_d), due to short channel effects. The use of recess offset distance is then a must to improve the maximum available gain cut off frequency.

The recess offset distance is of absolute importance for improving the g_d, however a maximum recess offset should be not more than $0.2\mu m$, beyond this value considerable degradation in f_c is probable.

For power applications, a rear plan doping is of interest both for the current level supplied and the linearity in the response. The optimum range of this doping level is to be kept between 1.5 and $2 * 10^{12} cm^{-2}$.

This page is intentionally left blank

This page is intentionally left blank

Chapter V

Simulation of HEMT's in breakdown

Abstract

In this chapter, we will present the results obtained through the two dimensional transient time energy modelling for HEMT's under power operation when the drain bias is pushed to the breakdown limit. We use our model in studying the effect of recess shape upon the breakdown potential in a uniformly doped HEMT.

V.1. Introduction

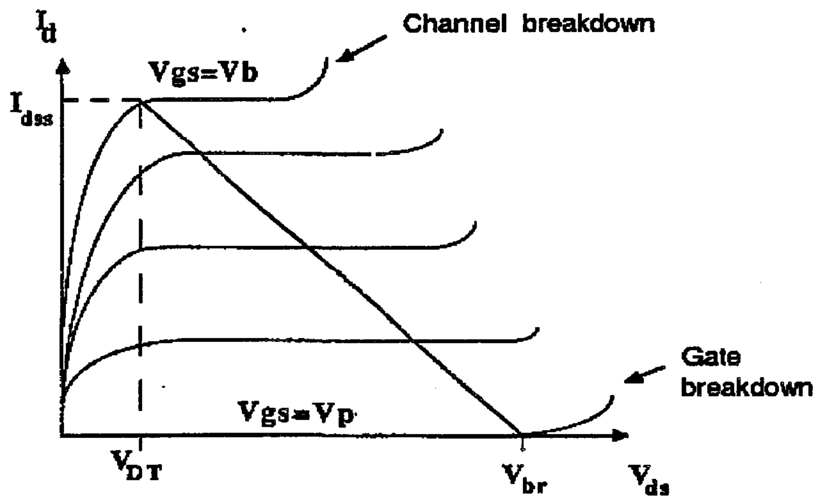


Figure V.1: Schematic presentation of I-V characteristics in a FET. The breakdown at pinch-off is the limiting value on V_{ds}.

In power operation, the transistor is submitted to a microwave signal at the gate which swings the gate voltage from open to even beyond pinch off potentials. On the drain side, a high drain-source voltage is found which might in certain cases cause the device to breakdown. As schematically shown in fig. V.1, the breakdown potential sets a limit on the usable range of the device specially at pinchoff where it could easily be found on the load line of the device. A through analysis of the device breakdown behaviour is essential to guarantee efficient and safe operation of the device.

The simulation models proposed so far do not present a reliable tool for studying breakdown behaviour. The models that use a steady state solution of the transport equation [84] might not be reliable for sub-half micron gate devices when the non stationary electron carrier transport becomes important. Another category of models solves Poisson equation with the continuity equations for electrons and holes while assuming zero currents [85], these models are of limited application. Quasi-two dimensional models recently introduced into the simulation of HEMT's in breakdown are to be validated before being certified as reliable simulation tools, this of course when the various related problems are to be solved, mainly the inclusion of minority carriers [86]. There are other types of simulation techniques that could

account for the avalanche multiplication but they are too complicated, for example the Monte-Carlo methods simulation models[87].

The experimental studies, although difficult, exhaustive and often result in device damage, have been able to demonstrate a satisfactory guide lines concerning the behaviour of HEMT's in breakdown, [88]. These studies have put in evidence the importance of carriers ionisation in the channel as the factor governing breakdown specially in open channel operation and also the role of gate tunnelling currents and surface fields in the breakdown near and beyond pinchoff. The major parameters that govern the breakdown potential are thus shown to be dependant on the operation regime; open channel or pinch off, device dimensions; the gate drain separation, thickness of the barrier layer, device type; planar doped, uniformly doped.

In the search for a more complete analysis, we attempted to use the two dimensional hydrodynamic transient time energy modelling. We were first interested in the impact ionisation as the cause of breakdown, no tunnelling currents were considered and care was taken so that operation regime do not get close to pinchoff for this reason.

In the simulation process, the effect of energy relaxation was taken into account (c.f. II.8.3) justified by the fact that an electron that has just passed an ionising collision, although present in a high field zone, needs a certain time for its energy to acquire the ionisation energy. Thus the electrons ionisation coefficient α_n is taken function of the average total energy ε to account for electrons non stationary and non isothermal transport. The holes ionisation coefficient β_p is taken a function of the local electric field E assuming that, due to their heavy effective mass, their energy is an instantaneous function of the electric field. This is a first approximation based upon the infinitesimal contribution of holes to the total current.

The simulations done in the context of studying the effect of impact ionisation contained two parts:

- The first part was about studying the breakdown characteristics of a conventional AlGaAs/GaAs HEMT with a simplified quasi planar structure.

- The second part was studying the effect of recess offset distance upon the breakdown characteristics of an HEMT.

We present the major results obtained in the following sections.

V.2. Breakdown in quasi planar HEMT

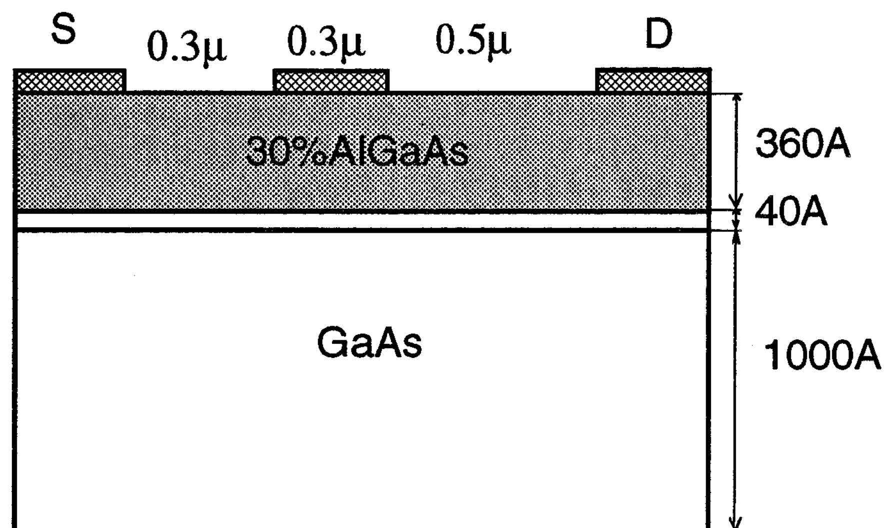


Figure V.2: The conventional HEMT considered for the study of breakdown characteristics. The gate length is $0.3\mu m$, the doping level is $1 * 10^{18} cm^{-3}$ uniformly over a layer $360\mathring{A}$ thick. The supply layer is AlGaAs with 30% Al, and the channel is n.i.d. GaAs

In order to understand what happens upon the breakdown of HEMT's we simulated a quasi planar structure where no surface effects were considered[1]. The HEMT was taken to be a conventional AlGaAs/GaAs structure, with a $0.3\mu m$ gate and a doping density of $1.10^{18}cm^{-2}$ over a layer of $360\AA$ thickness. The other structural parameters are indicated on fig. V.2.

The breakdown condition is set to be when the drain current rises to 10 times the current found at a smaller drain bias. Classically the breakdown condition is determined when the ionisation integral:

$$\int_0^{L_a} \alpha_n(r) e^{\left[-\int_0^r (\alpha_n(r')-\beta_p(r'))dr'\right]} dr = 1 \tag{V.1}$$

where r is the position along the electric field line and L_a is the total path length. We used (V.1) as a verifying condition by performing systematically at the end of each simulation point, when the breakdown is detected by our criteria, the integration of the ionisation coefficients in the channel. For all the cases, a sum of one was found along the path passing in the channel which confirmed our prediction of breakdown.

The I-V characteristics of the device are shown in fig. V.3. The behaviour of the device confirms with the experimental observations. For the gate bias varying from near pinchoff ($V_{gs} = -0.3V$ where $Vp \approx 0.5V$) to open channel operation ($V_{gs} = 0.3V$ where $V_{bi} = 0.8V$) we find that the breakdown potential decreases with increased gate potential. The values for the breakdown potential varying from 8 V near pinchoff to (5 V) at open channel are in accordance with the experimental results presented in fig. V.4[88] for a device of similar structure as the simulated one. The major difference in the precise values for the breakdown potential could be explained in view of the theory [89][88] which argues that the breakdown results from the combined effect of the impact ionisation in the gate drain region and tunnelling currents across the Schottky gate, this later is not accounted for in our simulations which might result in an over estimated breakdown potential in some

[1] A planar structure would have surface states in the inter electrode spacing corresponding to the doping density of the surface layer

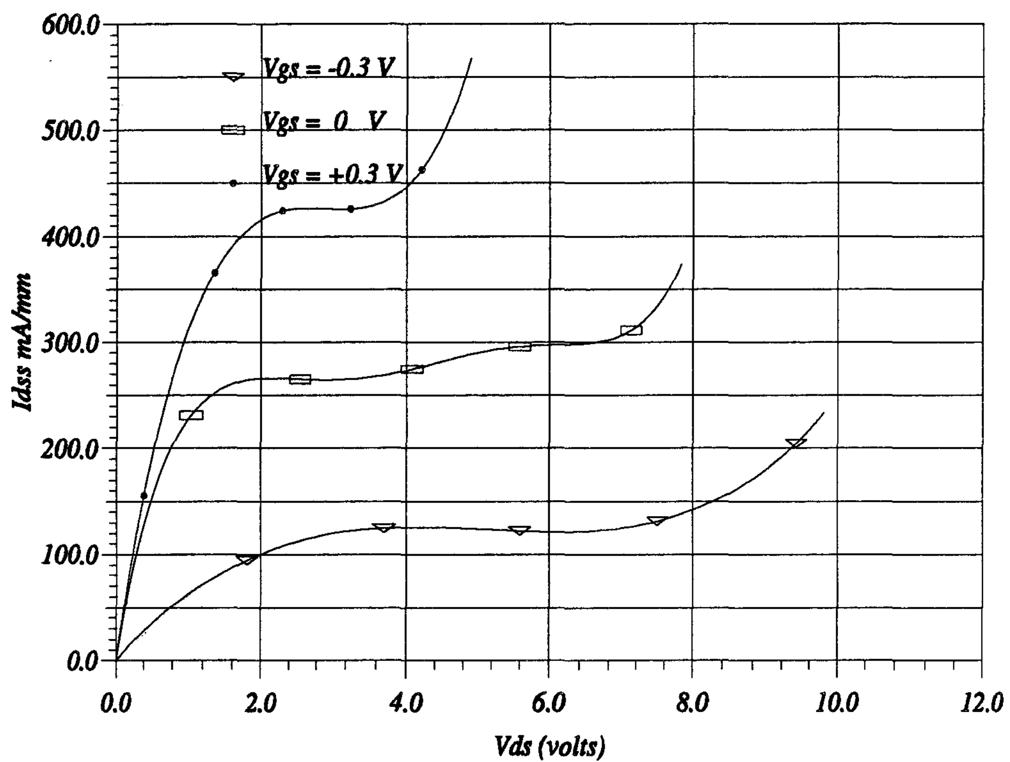


Figure V.3: Drain current versus Drain potential characteristics of a quasi planar lattice matched HEMT with uniform doping $10^{18}cm^{-2}$ at different gate potentials

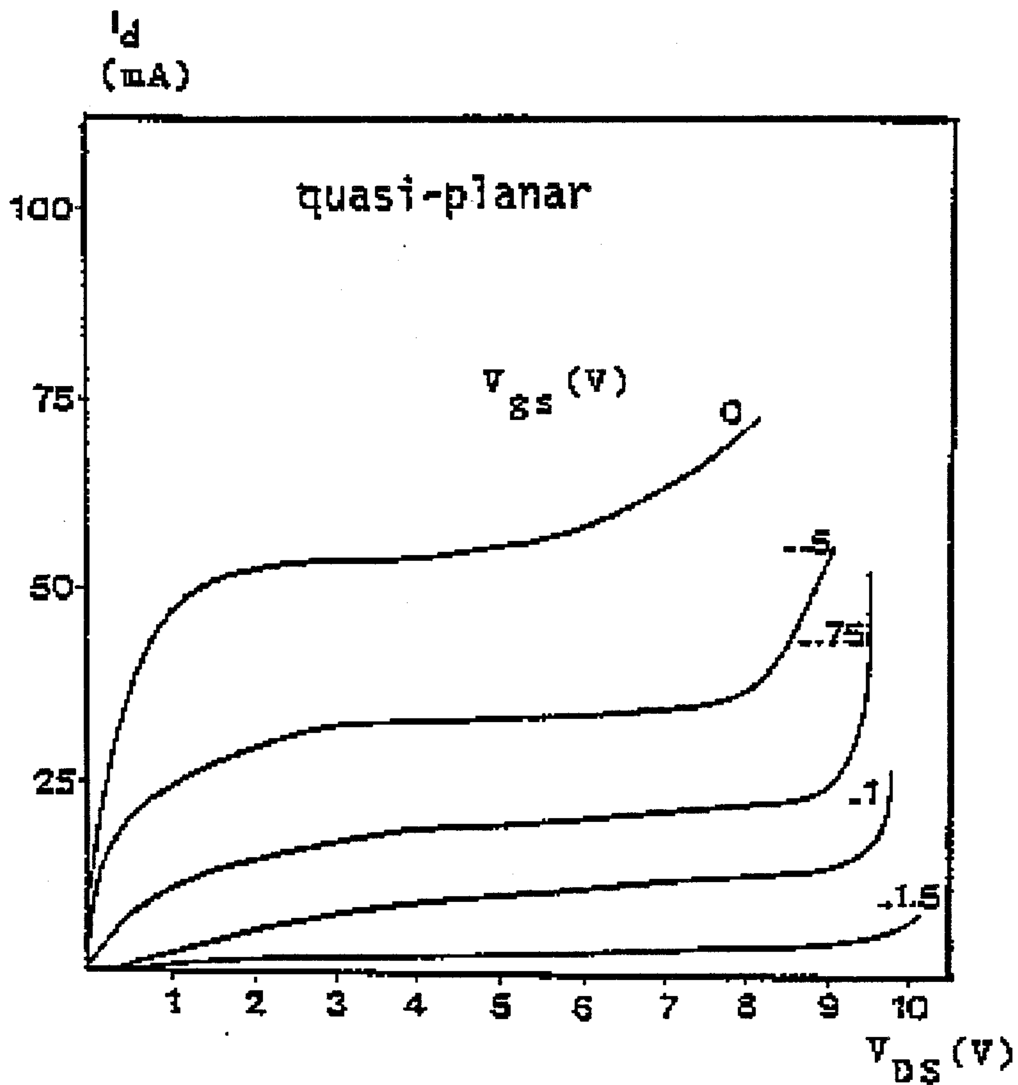


Figure V.4: Experimental drain current versus Drain potential characteristics of a quasi planar lattice matched HEMT with uniform doping $10^{18}cm^{-2}$ *at different gate potentials, gate length* $0.3\mu m$, *28%Al, barrier layer 380Å*

cases specially towards pinch off where the tunnelling is the phenomenon that mostly governs breakdown.

V.2.1. Analysis of breakdown

We have plotted the distribution of different physical quantities in the device under breakdown drain bias for the extreme cases of gate bias that we simulated. Figures (V.5, V.6) give the distribution of the electrons charge density, electrons generation, average electrons energy, electrostatic potential for gate biases -0.3, and 0.3 V respectively. From these figures we can observe the following:

- The major electrons generation zone corresponds to the high-energy-high-field domain situated at the gate exit. For negative gate bias the maximum generation rate is about $5*10^{28}sec^{-1}$, this rate goes down to $1*10^{22}sec^{-1}$ when the gate potential is set at 0.3 V. However, this could not be taken as an indication that the generation rate drops for higher gate bis, in fact these plots were done when the drain current just passed 1Amp/mm; in the case of an open channel, this would happen earlier than in the near pinch-off case. The difference in the time response of the device in the two cases makes it difficult to observe the behaviour strictly at the breakdown.
- The high energy domain is diffused all over the whole gate drain region, when the channel is opened, with an average energy of about 0.55 ev. In the case of near pinch off bias, this domain occupies about two fifth of the gate drain region but with a relatively higher average energy of about 0.7 ev.
- The equipotential contours reveal that the potential lines are perpendicular to the channel, and that there exists a high field region associated with the high energy domain.

Now let us see what happens to the high energy domain at a given gate bias when varying the drain bias. Fig. V.7 presents the energy in the quantum well (at a distance of $420\mathring{A}$ from the gate) in the device at different drain bias when the gate

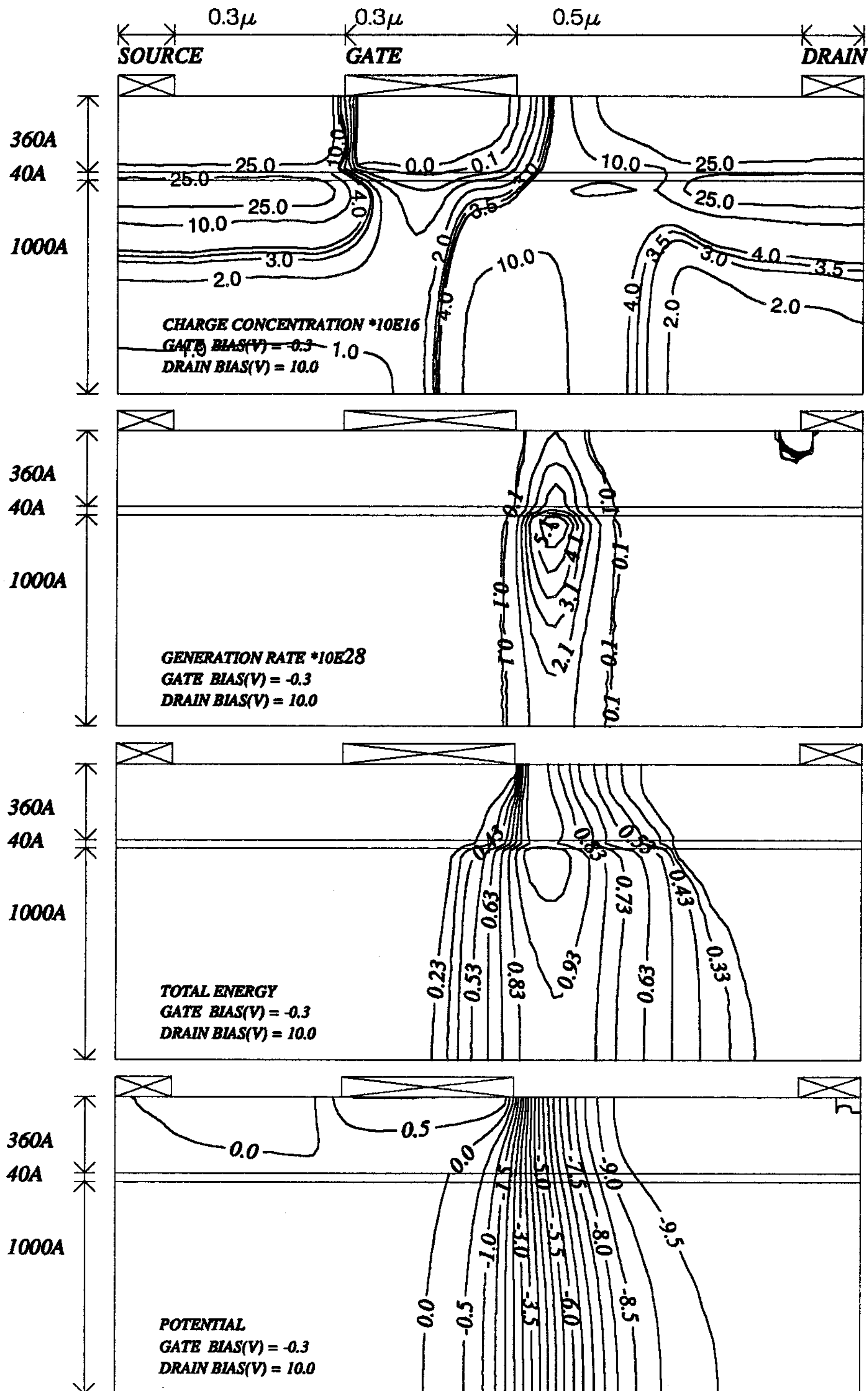


Figure V.5: Distribution of physical quantities inside the HEMT upon breakdown where the gate potential is near pinchoff.

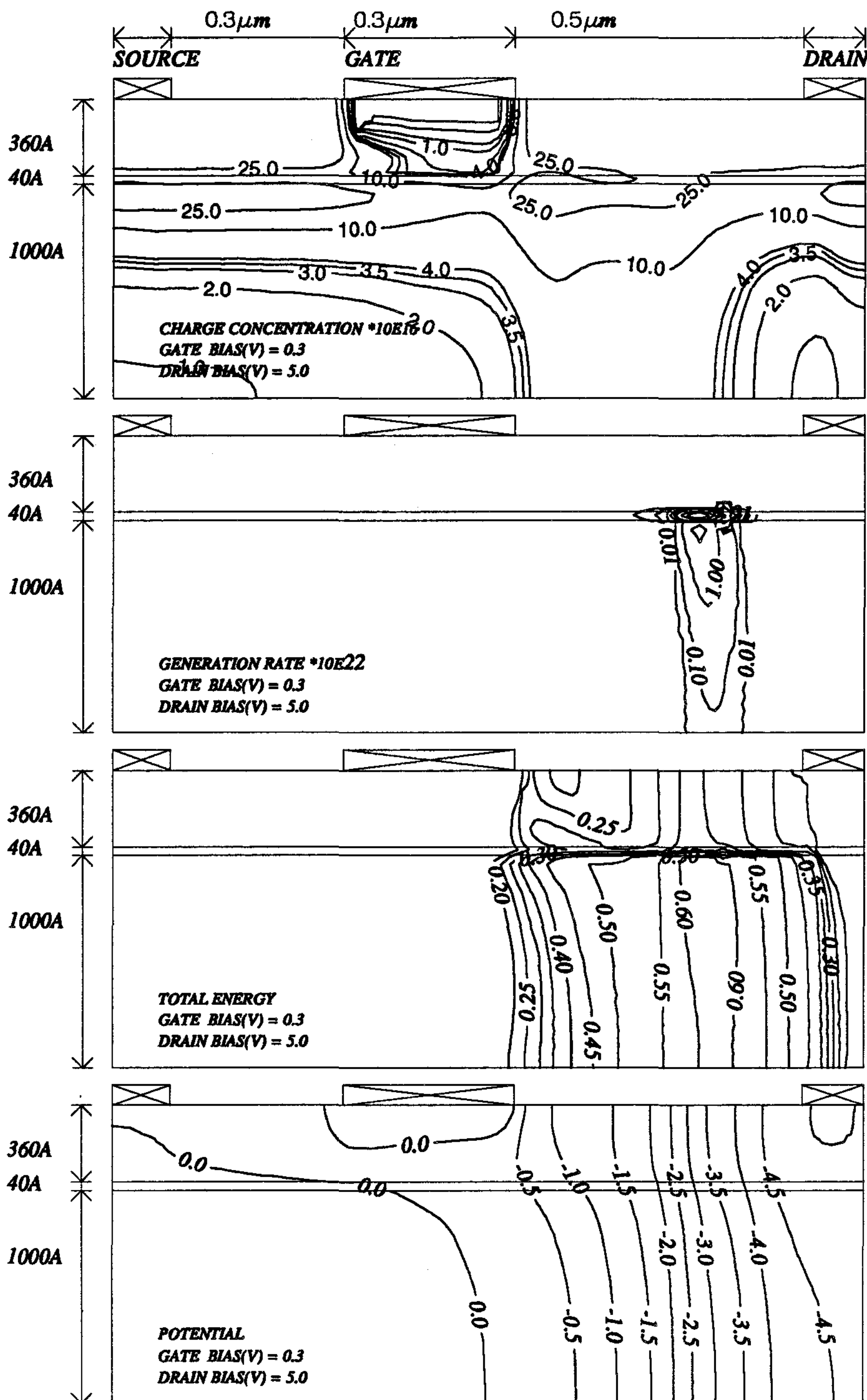


Figure V.6: Distribution of physical quantities upon breakdown with an open channel in a conventional AlGaAs/GaAs HEMT, uniformly doped at $3 * 10^{12} cm^{-3}$. The gate length is $0.3\mu m$, V_{gs}=0.3 V, V_{ds}=5.0V .

bias is kept at 0.3 V. At drain bias 1 V, the maximum energy is about 0.5 ev. When the drain bias is increased the maximum energy inside the domain starts to increase,

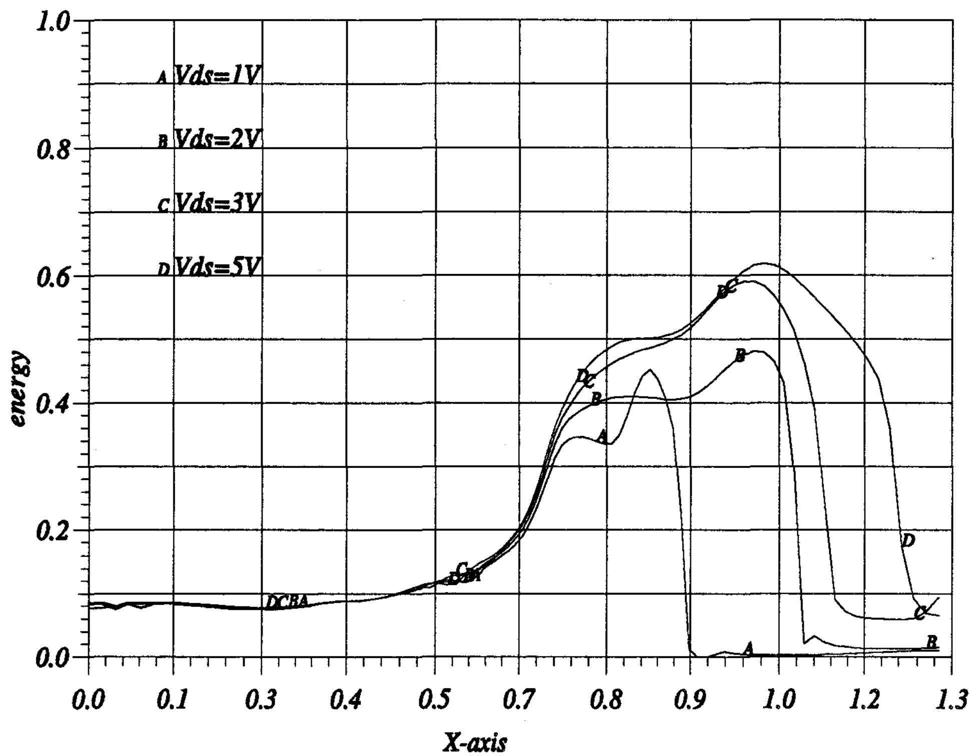


Figure V.7: The high energy domain variation with drain bias at an open channel operation

this increase is small when we change V_{ds} from 1 to 2 V, sharper upon the change to 3 V to be at a value of about 0.6 ev, and then slightly changing for V_{ds} 5 V.

The shape of the domain is affected by the increase in the drain bias. At 1 volt the domain is of a distinct shape starting at the gate exit and its width is about $0.2\mu m$. This width jumps by about 50% when the drain bias is increased to 2 V. At drain bias of 3 V an increase in the domain width puts it at the drain edge (starting at $1.2\mu m$) and then at 5 V the domain extends in the whole gate drain region.

The next study we will make is observing the behaviour of the high energy domain at the same V_{ds} when the gate bias is changed. Fig. V.8 gives such a plot where the drain bias is kept at 5 V. We notice that for near pinchoff operation, the peak energy is about 0.86 ev, this energy is reduced when the channel is more open to be

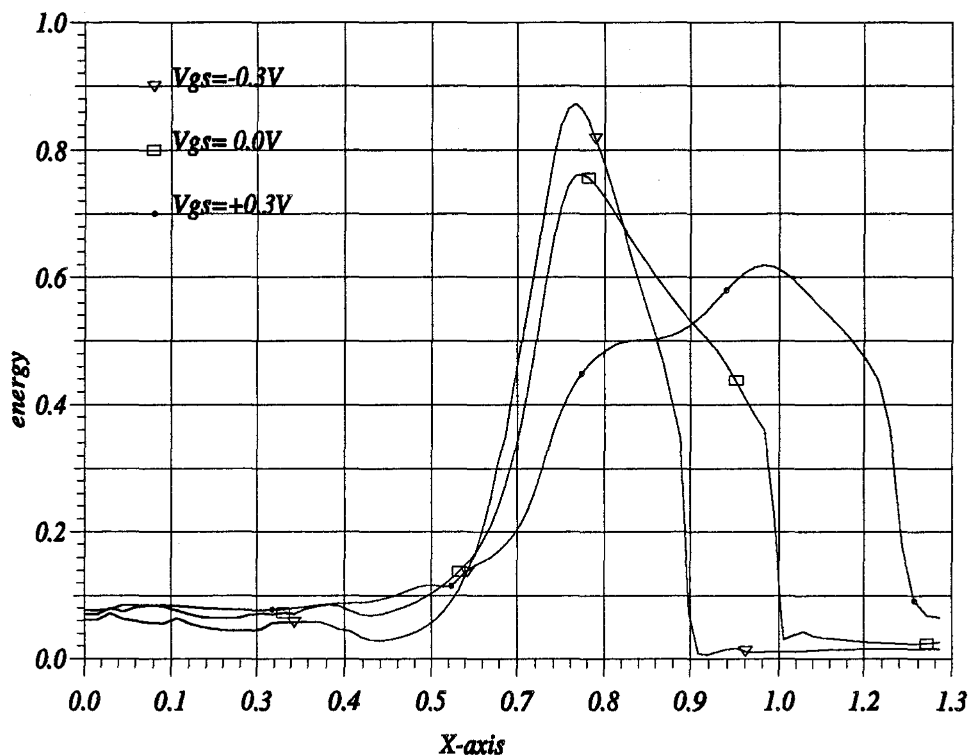


Figure V.8: The high energy domain variation at the same drain bias and for the three operation regimes of the gate bias.

0.75 ev at zero gate bias and 0.6 at 0.3 V of gate bias. The domain shape changes in its turn from a sharp distinct domain at the gate exit of average width of $0.2\mu m$ to a more diffused one of average width extending over the whole gate drain region.

In fig. V.9 we compare the high energy domain at breakdown at the different gate biases. We observe that the value of the maximum energy found decreases with

increased gate bias confirming with the previous observation. As the channel is more opened, the shape of the domain changes form a more or less a domain with

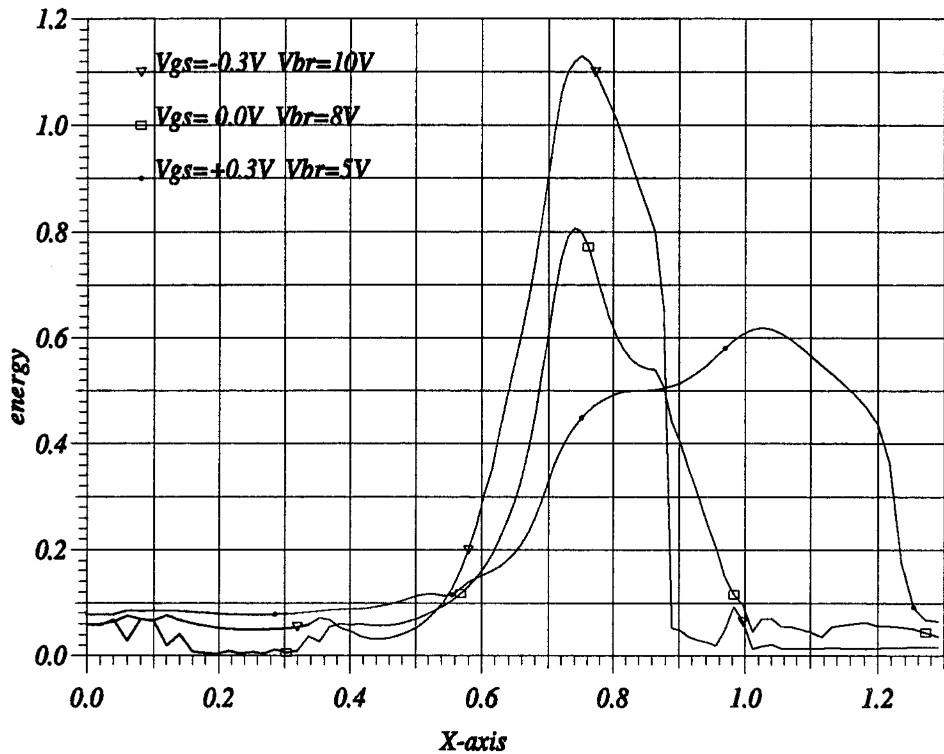


Figure V.9: Position of the high energy domain at breakdown for different gate bias

a distinguished peak to the one with more diffused shape

These curves permit to explain the difference in the breakdown potential since we take the ionisation coefficient as a function of energy $\alpha_n(\xi)$. This functional dependance (c.f. fig. II.5) gives a quasi saturation in the ionisation coefficient for energy values higher than 3 ev, thus it is the length of the high energy region that will determine the breakdown potential at such or higher energies. The more the domain is wide, the easier the avalanche multiplication will occur and the lower the

breakdown potential will be.

V.3. Study of the effect of recess offset

V.3.1. Motivation of the study

Ideally, the introduction of a recess offset at the source side would increase the access resistance and reduce the g_m and the f_c. It is thus of interest to have a small offset at the source side. On the drain side, a recess offset reduces the f_c and increases f_{max} (c.f. Chapter IV); the breakdown voltage would also increase. Accordingly, it is of interest to have a large recess offset on the drain side of the gate. Yet, the realisation of such an assymetric recess is cumbersome and often a symetric recess is used. The situation is different then because we will be having the combined effect of both offset distances. For longer recess offset, the drop in f_c might start to be dramatic while the breakdown voltage could keep improving. It is clear then that a compromise has to be made. This could only be possible when we understand the effect of different recess offsets on the breakdown potential.

V.3.2. Breakdown in recessed gate HEMT's

For this study, we simulated a set of three devices, these devices had the same gate length, the same doping density but differed in the recess offset distance. The first device had a zero offset distance with the n^+ zone extended till the gate, the second and the third device had recess offsets of $0.1\mu m$ and $0.2\mu m$ respectively with a surface depletion potential of 0.5V applied on the recess zone, fig. V.10. Figure (V.11) give the current voltage characteristics of the three devices. A clear increase in the breakdown potential is obtained upon the introduction of a recess offset as could be seen when comparing the first two figures. The variation of the breakdown potential with the recess offset is summarised in fig. V.12 where we could notice that the value of the breakdown voltage for different recess offsets, depends strongly upon the operation regime. In the open channel regime, ($V_{gs} = 0.3V$), a continuous increase in the breakdown voltage is found to occur upon the increase in

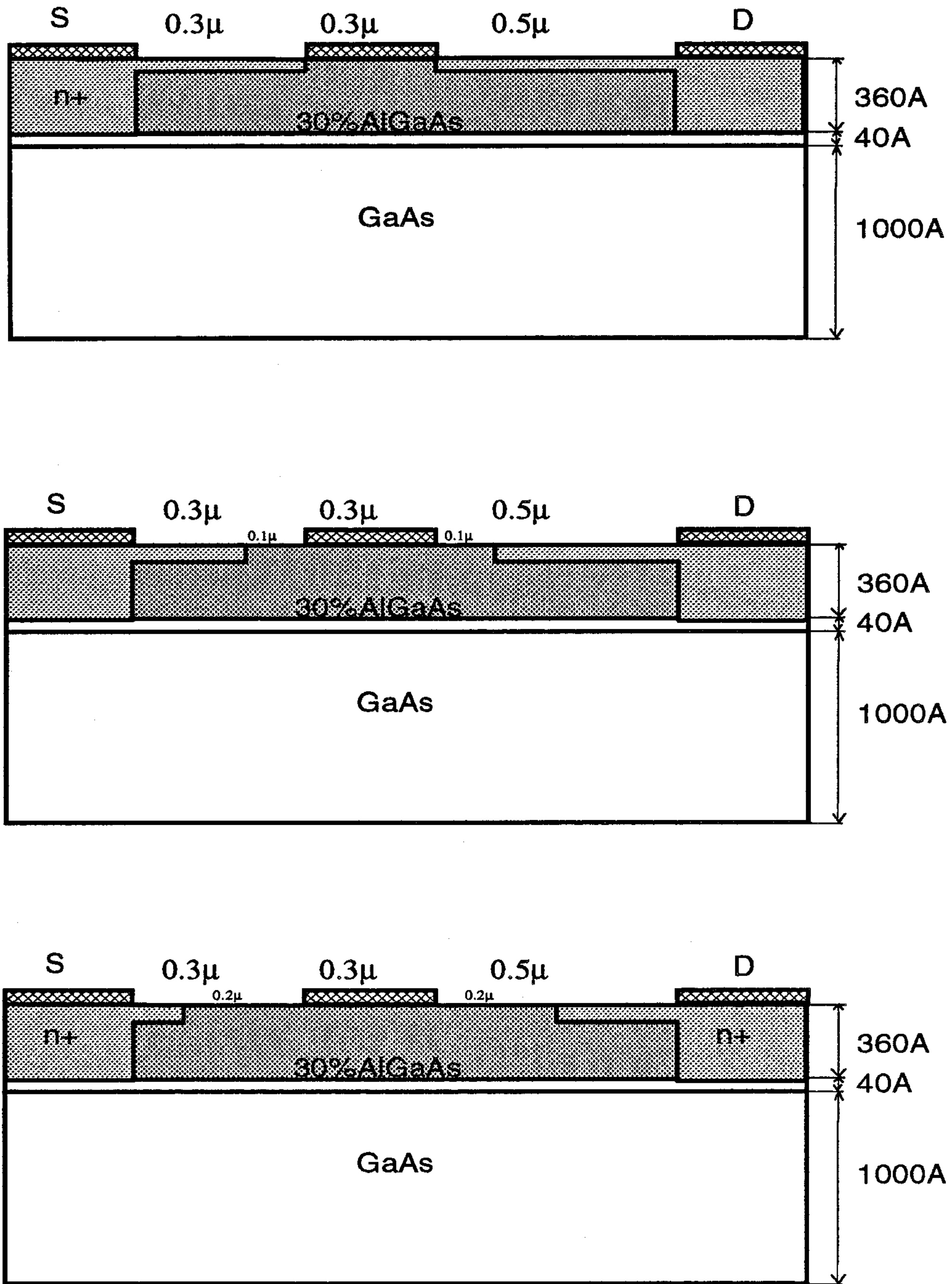
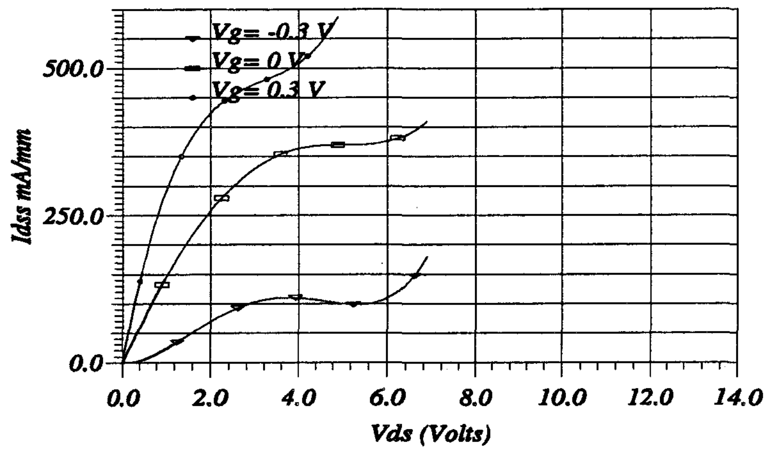
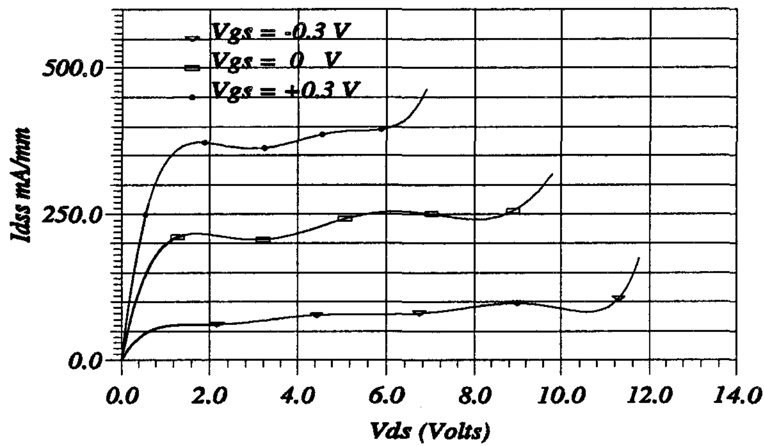


Figure V.10: Schematic presentation of the three devices considered for the study of the effect of recess offset. The first device is with a zero recess offset, the second device is with $0.1\mu m$ *offset and the third device is with* $0.2\mu m$ *recess offset distance. The gate length is* $0.3\mu m$ *and the doping level is* $1 * 10^{18} cm^{-3}$ *uniform over* $360\AA$





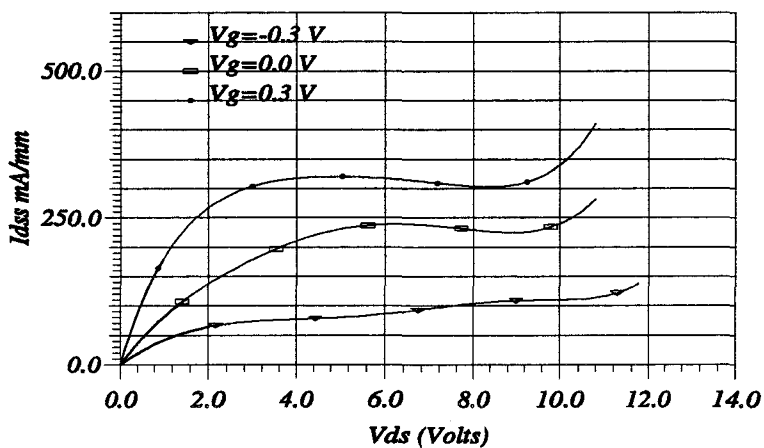


Figure V.11: I-V characteristics of a $0.3\mu m$ *conventional uniformly doped at* $1 * 10^{18} cm^{-3}$ *AlGaAs/GaAs HEMT with different recess offset distance. The top curves are for a device with a zero recess offset, the middle curve is for a device with* $0.1\mu m$ *offset and the bottom curve is for the device with* $0.2\mu m$ *recess offset distance*

the recess offset. For the other two values of gate operation regime (0 and -0.3 V), an improvement is found upon introducing an offset distance of $0.1\mu m$. However, this improvements tends to a saturation for increased recess offset.

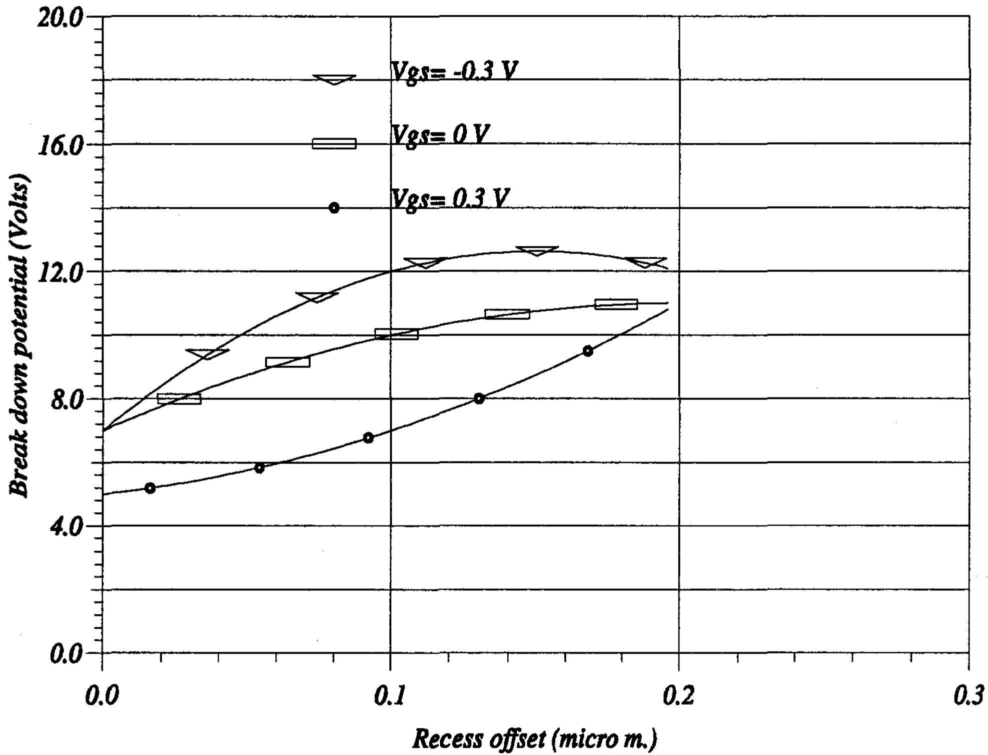


Figure V.12: Influence of recess offset upon breakdown potential for different gate bias

The improvement found upon the introduction of a gate recess offset distance of $0,1\mu m$ could be explained in view of the shape of the high energy domain once more. In fig. V.13 we have the distribution of the average electron energy along the channel at open channel condition $V_{gs} = 0.3V$ and at the breakdown drain bias. In the case of the lowest breakdown potential (zero recess offset) we recognise that the high energy domain is the one with wider shape than for the cases with recess.

Accordingly and in view of the same reasoning presented in the previous section, the wider domain shape will result in a smaller breakdown potential.

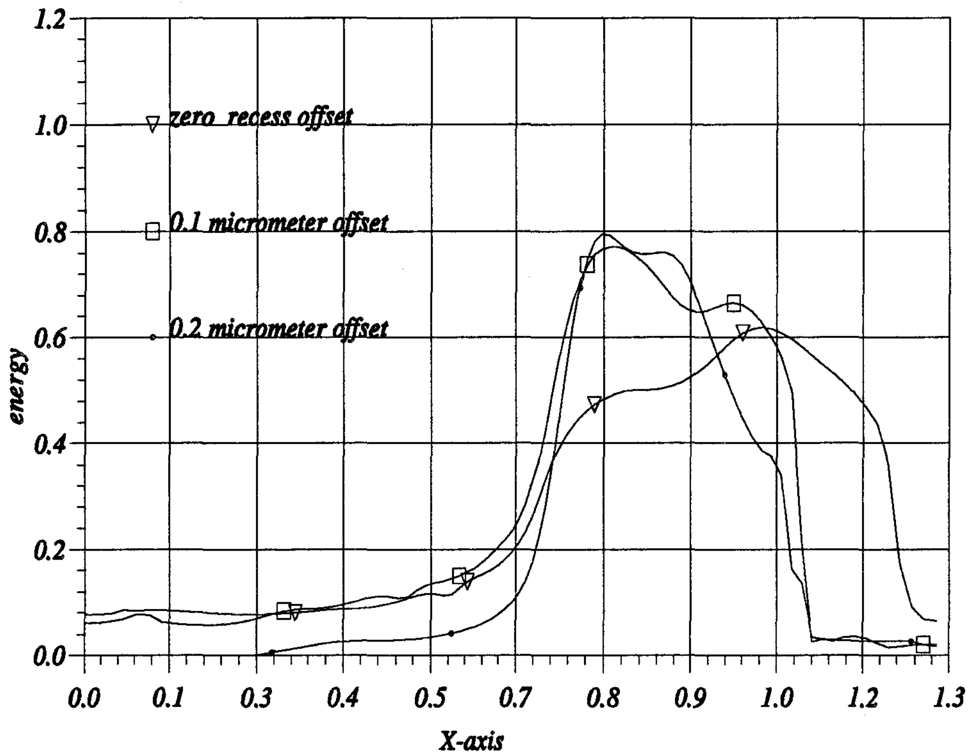


Figure V.13: Variation in the shape of the high energy domain in the quantum well for different recess offsets at breakdown in an open channel

For fixed gate and drain voltages, fig. V.14 gives the high energy domains in the three devices with different recess offsets. It is obvious once more that at this drain bias which results in breakdown for the zero offset device, the domain is of the widest form when compared to the others in the devices with superior offset distances.

For a near pinchoff operation, $V_{gs} = -0.3$, we do the same in plotting the high energy domain in the well for the three offset distances, fig. V.15. We find that at zero recess offset the domain is the wider with the highest absolute value (1.1 ev), for

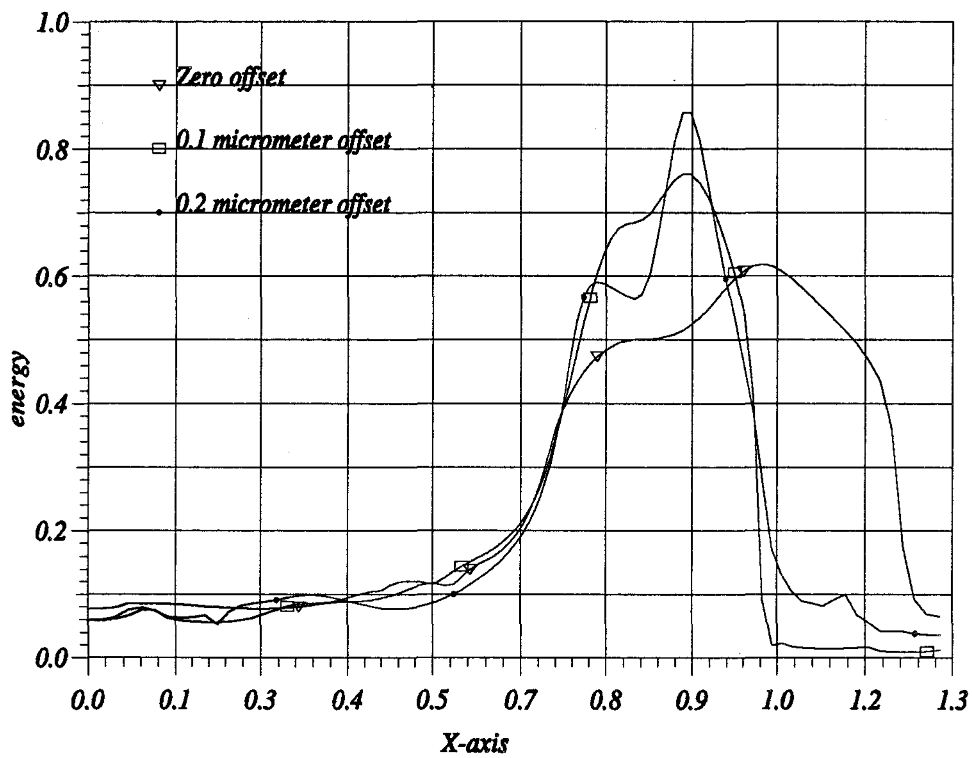


Figure V.14: High energy domains in the quantum well for different recess offsets at the same drain bias (5V) and the same gate bias (3V)

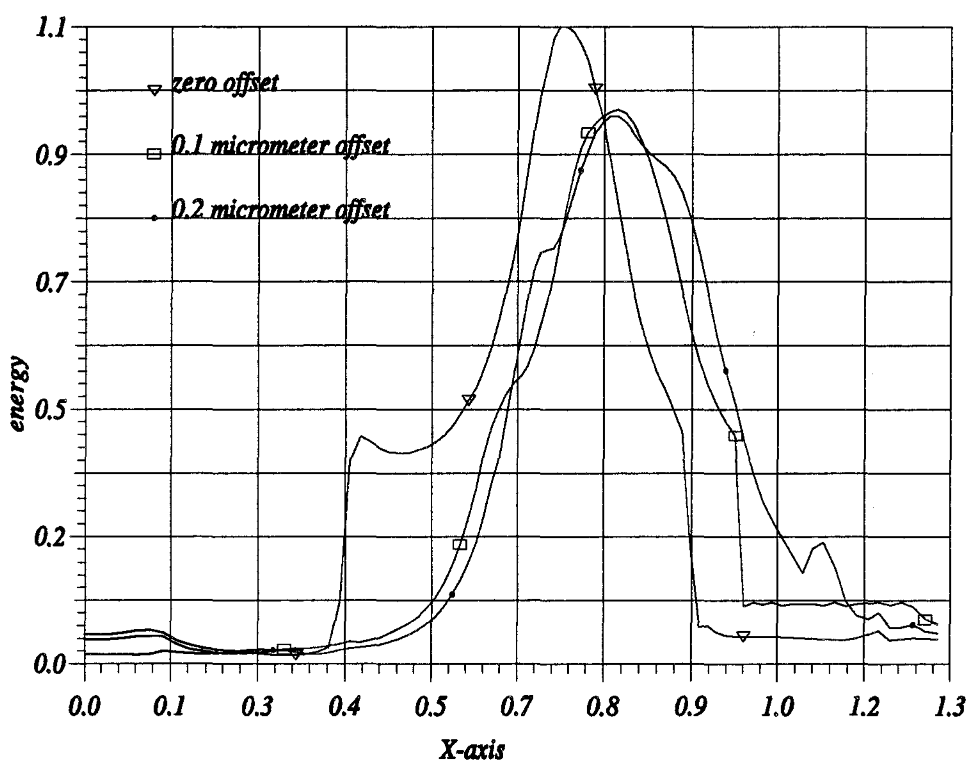


Figure V.15: Variation in the shape of the high energy domain in the quantum well for different recess offsets at breakdown for a near pinch off operation

the other two cases (0.1 and $0.2\mu m$) the high energy domains are almost identical. This is very interesting as it explains why we find a saturation in the breakdown potential at near pinchoff operation in spite of doubling the recess offset distance.

After this exhaustive study we can conclude that the introduction of a recess offset distance results in reducing the high energy domain width which in its turn is responsible for the increase in breakdown potential. At near pinchoff operation, the width of the high energy domain stays unchanged which gives a tendency of saturation in the breakdown potential for recess offset larger than $0.2\mu m$. This study holds in the frame of the imposed hypotheses first of which that the breakdown is due only to impact ionisation in the channel. Generalisation by introducing the tunnelling current is mandatory for the study of breakdown performance. We have recently verified the technique of simulating the gate recess by comparing the outcomes of a modified version of our model which accounts for a realistic recess shape, the agreement that we found with the approximated method assures its validity in this study[90].

V.4. Conclusion

We have managed to study the breakdown behaviour of HEMT's using our hydrodynamic energy model. The first findings confirmed with the experimental ones and the physical reasoning. When we used the simulator to study the effect of the recess offset distance, we found that the use of an offset distance improves the breakdown potential by a factor of $\approx 30\%$. For wider recess offsets, the breakdown potential tends to saturate at mid channel and near pinchoff, yet it continues improving for open channel.

This is quite interesting as in practical applications what counts more is the breakdown near pinchoff. The results we found imply that for a conventional uniformly doped HEMT with $0.3\mu m$ gate, a recess offset of $0.1\mu m$ is sufficient.

A detailed analysis of the distribution of average electrons energy in the quantum well (channel) revealed that the breakdown potential is strongly dependant upon

the shape of the high energy domain. The wider this domain is, the lower the breakdown potential will be. This study revealed also that the saturation in the breakdown potentials at near pinchoff operation for different recess offsets is due to identical distribution of the high energy domain in the channel.

This page is intentionally left blank

This page is intentionally left blank

Chapter VI

Summary and Conclusions

Abstract

In the closing of this work, we summarise up what was done to make the point. The objectives, the means used the major results obtained are concretised. The future extensions are presents and discussed.

VI.1. Motivation of this work

In the need for higher power and higher operating frequency the pseudomorphic AlGaAs/InGaAs HEMT's have shown to be the most promising devices, however they suffered from a drop in the output power density for frequencies higher than 60GHz. To improve their response, an optimisation process imposed the need for studying the influence of device scaling upon its potential performance.

The huge number of the parameters to be investigated separately and collectively posed a question mark upon the possibility of using uniquely the experimental approach and the use of simulation techniques is a must.

Among the different choices for simulation models, the two-dimensional hydrodynamic energy modelling presented a good compromise between accuracy and rapidity. We had chosen a model previously conceived for the simulation of conventional HEMT's. This model needed to be modified to suit the simulation of pseudomorphic devices through the introduction of a number of phenomenon not accounted for previously.

VI.2. The Simulator

VI.2.1. The physical model

The model is based upon the solution of three transport equations driven form the Boltzmann Transport equation for non homogenous semiconductors, coupled with Poisson's equation.

The following effects were **originally added** to the model:

- A quantification effect correction by assuming that the carriers are found in the first energy subband and not in the bottom of the conduction band. The energy step is calculated through the triangular well or the finite square well approximation according to the device.

- The introduction of screening effect through the use of empirical data relating the 2DEG low field mobility to its interface charge density, this phenomenon had a major effect upon the terminal characteristics. **When compared to experimental work, the model gave very acceptable results.**

- The simulation of minority carriers transport by introducing a drift diffusion type set of equations for the holes.

- The introduction of avalanche multiplication by impact ionisation in the conservation equations by taking an energy dependant ionisation coefficient rather than field dependant for the electrons. This permitted the simulation of devices under breakdown conditions.

VI.2.2. Mathematical and computational models

The set of six partial deferential equations was written into a finite difference mesh scheme under vectorial formulation. The spatial mesh used a highly non uniform step and the temporal discretisation was done using a dynamic time step. The mesh step and the time step are both governed by the material parameters and device structure.

The resulting sets of equation was solved using a parallel processing technique where the decoupling was based upon the relaxation time approximation.

VI.3. Main results obtained

- The systematic use of our model demonstrated the advantage in using planar doped AlGaAs/InGaAs HEMT's to uniformly doped ones.

- The choice of buffer type have shown to be of little effect upon the current gain cut off frequency and the feedback parameters.

- The choice of the undoped barrier layer thickness have shown not to be of major effect provided the aspect ratio greater than 10. The choice of its thickness will be governed by breakdown considerations and short channel effects.

BU LILLE I

- The scaling of the gate length indicates a limit of 0.15μm as shorter gate lengths if we want to get the best out of the device.
- The recess offset distance is of a great impact upon the maximum available gain cut off frequency. Its introduction is necessary to sustain this frequency at a high value when the gate length is reduced.
- Increasing the doping level of the fore plane gives higher output current and more linear response.
- The use of pulse doped buffer layers increases the output current too, but increasing its value degrades the device low noie performance and efficiency.
- A study of the breakdown characteristics of conventional HEMT's shows that an improvement in the breakdown potential is possible through increasing the recess offset distance. An optimal value for the $0.3\mu m$ gate device is $0.1\mu m$.

VI.4. Limitations and extensions

In its actual status the limitations of the simulator could be discussed based on three axis; physical, mathematical and computational.

As for the physical model, it seems about time to include the effect of inertial terms in the momentum conservation equation. This could result in modifying the treatment of the energy equation as well. The heat conduction gradient term ∇Q, in the energy equation seems important to study its effect upon the model performance. Also a more precise way to account for the quantification effects is necessary to be included.

The treatment of the Schottky gate needs to be modified to account for tunnelling currents as well as the 'on diode' operating conditions. In the same context the inclusion of tunnelling currents across the heterojunction is of importance specially in the case of narrow wells.

The range of application of the model could be extended to the simulation of bipolar

transistors. Another potential application, in view of the rapid performance, would be dynamic simulation of transistors having a sinusoidal signal on the gate electrode. The coupling with the external circuit which we have started is very temptative to be followed up.

From the mathematical point of view, the use of other techniques such as the conjugate gradient method seems of interest to ameliorate the model stability. Also the use of rapidly improving mathematical libraries is to be investigated to improve the simulator precision.

The computational model needs to be adapted to the rapidly improving machines, specially the massive parallel systems, this will dramatically improve the model performance. In the same time, more effort has to be done to improve the useability of the model by device engineers.

The model could then be used for the study of more complex structures not even necessarily FET's but could be extended to the study of HBT's and integrated optics devices.

VI.5. Closing remark

We think that we have fulfilled the objectives set for this work. We have developed a numerical tool for the study and analysis of sub micrometer pseudomorphic HEMT's. We used it to study the device performance and to examine the effect of varying a number of structural parameters on its performance. The results presented in this work were used in several contracts and joint publications, in itself it has led so far to the following publications and communications:

Publications and communications

1. K.Sherif, G.Salmer, O.L.El-Sayed. "Performance analysis of pseudomorphic MODFET's", Asia Pacific microwaves conference, Tokyo, JAPAN, 1990.
2. K.Sherif, G.Salmer, O.L.El-Sayed, T.Shawky. "2D Transient time simulation of pseudomorphic structures", WOCSMAD, Florida, USA, 1991.
3. K.Sherif, G.Salmer, O.L.El-Sayed. "Hydrodynamic energy modelling of MODFET's, analysis of pseudomorphic performance", WOCSDICE, Ulm, GERMANY, 1991.
4. K.Sherif, G.Salmer, I.Mouatakif. "Recent results of hydrodynamic energy modelling of FET's", GaAs simulation group meeting, Lille, FRANCE, 1991.
5. K.Sherif, G.Salmer, O.L.El-Sayed. "Parasitic oscillations in MODFET's", WOCSMAD, Texas, USA, 1992.
6. K.Sherif, G.Salmer, O.L.El-Sayed. "Simulation of delta-doped MODFET's", WOCSDICE, Madrid, SPAIN, 1992.
7. K.Sherif, G.Salmer, O.L.El-Sayed, A.Refky. "Recent results of MODFET simulation : dual gate and delta-doped structures", URSI, Paris, FRANCE, 1992.
8. K.Sherif, G.Salmer. "Optimisation des transistors HEMT pseudomorphique au moyen d'une simulation bidimensionelle hydrodynamique", JNM, Brest, FRANCE, 1993.
9. K.Sherif, G.Salmer. "Simulation bidimensionelle hydrodynamique des transistors HEMT", 25em congres d'analyse numerique, Toulon, FRANCE, 1993.
10. K.Sherif, G.Salmer, O.L.El-Sayed. "An enhanced two dimensional hydrodynamic energy model for transient time simulation of complex heterostructure field effect transistors", SISDEP93, Vienna, AUSTRIA, 1993.
11. K.Sherif, G.Salmer, O.L.El-Sayed. "Application of 2D hydrodynamic energy modelling to the optimisation of planar doped HEMT's", ESSDERC93, Grenoble, FRANCE, 1993.

References

[1] Y.Crosnier, G.Salmer *in III-V microelectronics*, p97, North-Holland publ., 1991.

[2] R.Allam, unpublished results.

[3] The first report on HEMT was given by T.Mimura et al,Jappan J. of Appl. Phys, vol.19, pp-L-225, 1980.

[4] G.Salmer et al., Rapport final de contrat DAII-CNET.

[5] T.J. Drummond et al. IEEE Electron Devices Letters, vol. EDL-3, pp.338-341, 1982

[6] T.J. Drummond et al., Proceedings of IEEE, June, 1986.

[7] P.Godts et al., Elect. Lett. vol.25, no.15, p937, 1988.

[8] P.Saunier, IEEE EDL, vol. EDL-6, no.9, 1986

[9] F. Temcamani, *PhD dessertation*, 1990, Lille, France. (Available on request)

[10] See for example : T.N.Theis, Appl. Phys. Let., vol.49 no.22, 1986

[11] J.L.Thobel, *PhD dessertation*, 1987 Lille, France.

[12] N.H.Sheng et al,IEEE Trans. Elect. Dev. letters vol.EDL-6, no.6, 1985.-

[13] Nguyen et al., IEEE Trans. Elect. Dev. vol ED-35, p139, 1988

[14] These values were reported through internal communications, the seem however a little bit exagerated.

[15] H.Fawaz et al. Device research Symposium, Dec. 1993.

[16] I.Mouatakif *memoire DEA* 1988, Lille, France.

[17] J.W. Matthew and J. Blakeslee, J. Vac. Sci. Tech., vol14, p989, 1977

[18] H.Morkoc Principles and Technology of MODFETs, J.Wiley, 1991.

[19] U.Ravaioli and D.Ferry, IEEE Trans. Elect. Dev. vol. ED-33, pp.677-681, 1986.

[20] R. Fauquemberg et al *Quantum Monte-Carlo device simulation*, 25em Congres d'analyses numerique 1993, Toulon, France.

[21] M. Mouis *PhD Dessertation*, 1988, Orsay, France. (Available on request).

[22] B. Vinter et al, Apllied Phys. Lett. vol 44, p307, 1983.

[23] J. P. Nougier, *in III-V microelectronics*, p46, North-Holland publ., 1991.

[24] Delagebeaudeuf and Linh, IEEE Trans. Elect. Dev., vol. 29, p955, 1982.

[25] Drummond et al., IEEE Trans. Elect. Dev., vol. 30, p338, 1983.

[26] Moon et al., IEEE Trans. Elect. Dev., vol. 37, p908, 1990.

[27] H. Happy et al, proceedings of SISDEP 1993, Vienna, Austria.

[28] A. Cappy, *PhD dissertation*, Lille University, 1986, Lille, France. (Available on request).

[29] C.M. Snowden et al, IEEE Trans. Elect. Dev., ED-36 p1564, 1989.

[30] S. Selbherer et al, IEEE Trans. Elect. Dev., ED-27, p1540, 1980.

[31] C.M. Snowden and D.Loret, IEEE Trans. Elect. Dev., ED-34, p212, 1987.

[32] M. Saadoon, *MSc. Dissertation*, 1983, Cairo University, Cairo, Egypt., (Available on request).

[33] S. El-Ghazaly *MSc. Dissertation*, 1983, Cairo University, Cairo, Egypt., (Available on request).

[34] F.Heleiodore et al., Trans. Elect. Dev., vol. 35, p824, 1988.

[35] Tarek Shawky et al., Trans. Elect. Dev., vol. 37, p38, 1990.

[36] Tarek Shawky, *PhD dissertation*, Lille University, Lille, France, 1990.

[37] I. Mouatakif, *PhD dissertation*, Lille University, 1993, Lille, France. (Available on request).

[38] S. Selbehrer, *Simulation of Semiconductor Devices*, Springer-Verlag publ. 1981.

[39] T.Shawky *MSc Thesis* 1987, Cairo University, Egypt.

[40] K.Sherif *MSc Thesis* 1990, Cairo University, Egypt.

[41] Zhou and Ferry, IEEE Trans. Elect. Dev. Feb. 1993.

[42] J.P.Nougier in III-V Electroniques, 1992.

[43] J.P. Nougier and M.Rolland, J. Appl. Phys. vol.48, p1683, 1977.

[44] E.M.Azoff, Solid State Electronics, vol.30, pp913-917, 1987.

[45] E.Constant, in Hot-electron transportin Semiconductors, L.Reggiani, Ed., Springer-Verlag, 1984.

[46] K.Blotekjaer, IEEE Trans. Elect. Dev. vol.17, pp38-70, 1970.

[47] O.L.El-Sayed et al, Solid State electronics, vol.30 pp643-654, 1987

[48] J.P.Nougier et al, J.Appl. Phys., pp825-832, 1987.

[49] U.Ravaioli and D.Ferry, IEEE Trans. Elect. Dev. vol. ED-33, pp.677-681, 1986

[50] H.Morkoc Principles and Technology of MODFETs, J.Wiley, 1991.

[51] D.Lippens, *These d'Etat* 1984, Lille, France.

[52] O.Vanbesien, *Memoire de DEA* 1987, Lille, France

[53] Y.Crosnier et al, Proceedings of IEE, vol. 34, no. 1 1987

[54] G.D.Hachtel et al, IBM J. Res. and Devel., p232, 1981.

[55] H.Happy *Memoire DEA* 1989, Lille, France.

[56] There is a full bibliography on this topic found in: *Principales and Technology of MODFETs* H. Morkoc, publ. J.Wiley, 1991.

[57] J.Alamkan et al., Europ.Trans on Telecom., vol.1, no.4 1990.

[58] G.Dambrine, results published in the Classic contract, 1993.

[59] J.B. Gunn, Solid State communications, p88, 1963.

[60] M.R.Friscourt *These d'Etat*, 1987, Lille, France.

[61] K.Yamagushi et al.,IEEE Trans. Elect. Dev., ?ED-33, p1283, 1976.

[62] H.Willing and P. de Santis, Elect. Lett., vol. 13, no.18, 1977.

[63] K.Sherif et al, GaAs simulation group meeting, Lille, Frnace, 1992.

[64] A.Tachafine, *PhD dessertation* 1994, Lille, France

[65] G.Salmer and K.Sherif, to be published

[66] K.Sherif, *M.Sc. Thesis* 1990, Cairo University, Egypt.

[67] R.T.Kaneshiro et al.,IEEE Trans. Elect. Dev. ED-9, no.5, 1988.

[68] P.Lugli et al, Elect. Lett. Feb 1991.

[69] A.Cappy, unpublished results.

[70] J.C.Dejaeger,unpublished results.

[71] K.Yokoyama et al.,IEEE Trans. Elect. Dev. Lett. EDL-6, p536, 1985.

[72] T.J.Drummond et al., J. Appl. Phys., p1028, 1982.

[73] F. Heliodore et al., IEEE Trans. Elect. Dev. vol. ED-35, pp.824-830 1988.

[74] T.Shawky et al, Europ. Trans. on Telecomm. July 1990

[75] I.C.Kizilyalli et al, IEEE Trans. Elect. Dev., ED-33 p142, 1986.

[76] T.J. Drummond et al, Appl. Phys. Lett. p262, 1983.

[77] A.Cappy, *PhD Dessertation* 1986, Lille, France.

[78] T. Goupez *PhD dessertation* 1993, Lille, France.

[79] F.Temcamani *PhD dessertation* 1990, Lille, Frnace.

[80] G.Salmer et al, IEEE Trans. MTT, 1987.

[81] G.Dambrine, results published under Classic contract.

[82] Y. Jin *PhD Dessertation* 1993, Universite de Paris Sud, Orsay, France.

[83] F. Heliodore *PhD Dessertation* 1987, Lille, Frnace.

[84] Chau H., IEEE Trans. Elect. Dev., ED-38, p213, 1991.

[85] M.Lefebvre private communications.

[86] J.C. Dejaeger unpublished results

[87] D.Lippens and O. Vanbesien, Intl. Symp. on GaAs and related compounds, Herakleon, Greece, 1987.

[88] F.Temcamani *PhD dessertation* 1990, Lille France.

[89] J.Ashworth p561, proceedings of IEDM-1990

[90] J.C.Dejaeger unpublished results.

[91] IBM FORTRAN77 Vector Programmer's guide.

This page is intentionally left blank

Appendix A

Vectorisation and Multitasking

A major factor in the speed up in the program is due to the use of two features of the super computer IBM-3090-600E;vVector processors and the multitasking facility.

A.1. Vector processing

It could be otherwise called a pipelining process since the data could be seen as propagating in a pipeline of instructions. The main feature of vector processors is that they process multiple array elements at the same clock instant. These processors gather and overlap different instructions on different elements in an array or group of arrays, processed in a loop cycle which results in reducing dramatically the execution time of the whole array or vector. The reduction in the execution time results from the splitting of a loop on an array or a vector into a series of stages of simpler structure. After an operand, or a pair of operands, has been processed by the first stage in the pipeline, a fetch for example, it can be process by the second stage , e.g. add the two operands, while the first stage accepts another operand or pair of operands.

Vector processing is applicable on one dimensional arrays (vectors) or two dimensional arrays where the columns are considered as vectors. However, the efficient use of pipelining, from the point of view of data ordering and handling, is more pronounced for the case of vectors[91]. The efficient use of pipelining is also strongly

linked to the type of data being processed. The existence of forward or backward dependency might result in a slow down in the processing time or in the worst case elimination of vector processing. Talking about efficient vector processing, it should be mentioned that vector processors, at least in the super computers, posses a large cash memory which is partially contributing to the speed up in the overall performance by reducing the swapping events in data recuperation. It is thus of interest to have long and few vectors, of course if the application permits, in order to increase the vectorisation efficiency which is given by:

$$\eta \propto \frac{V}{\alpha_s} \tag{A.1}$$

where α_s is the proportion of the scalar to vector code, in the same process, and V is the vector length.

To have better vector performance the following is to be satisfied:

- The scalar/vector code ratio is to be minimised by trying to vectorise the maximum of the operations in the application.
- The vector length is to be maximised.
- Data dependency is to be eliminated (or minimised).

As an example of vector processes that we applied in our program, consider the calculation of the electric field in the device

$$E = \nabla V \tag{A.2}$$

where $E = E_x + E_y$ is the electric field and V is the electrostatic potential. In a classical two dimensional finite difference scheme (fig. A.1)), each of the components of the electric field would read :

$$EX_{i,j-1/2} = \frac{V_{i,j} - V_{i,j-1}}{\Delta X_{j-1}} \tag{A.3}$$

$$EY_{i-1/2,j} = \frac{V_{i,j} - V_{i-1,j}}{\Delta Y_{i-1}} \tag{A.4}$$

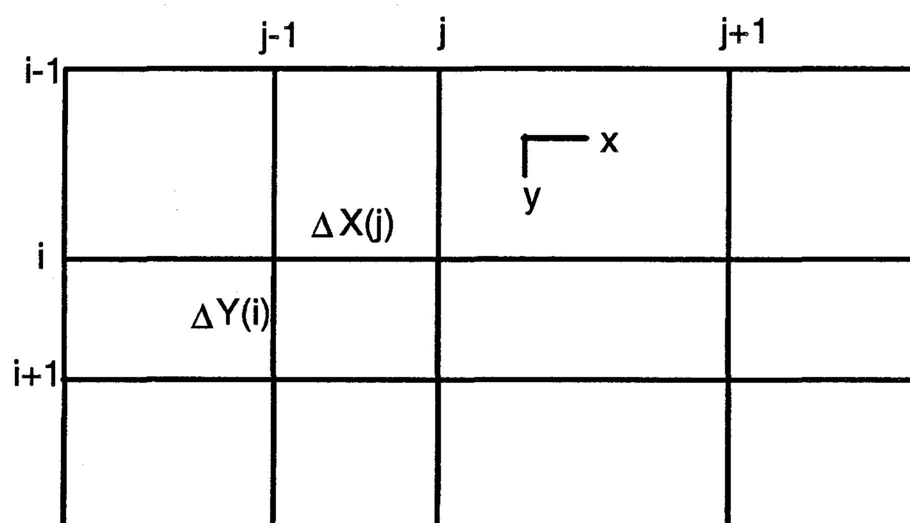


Figure A.1: Nomenclature of a classical finite difference mesh

The first step to vectorise this calculation is to transform these two array expressions of dimensions (m*n) in the y and the x directions respectively, into vectors. For this we apply the finite difference scheme shown in fig.(A.2). This would result in having:

$$EX_{K-1/2} = \frac{V_K - V_{K-m}}{\Delta X_W} \tag{A.5}$$

$$EY_{K-1/2} = \frac{V_K - V_{K-1}}{\Delta Y_Z} \tag{A.6}$$

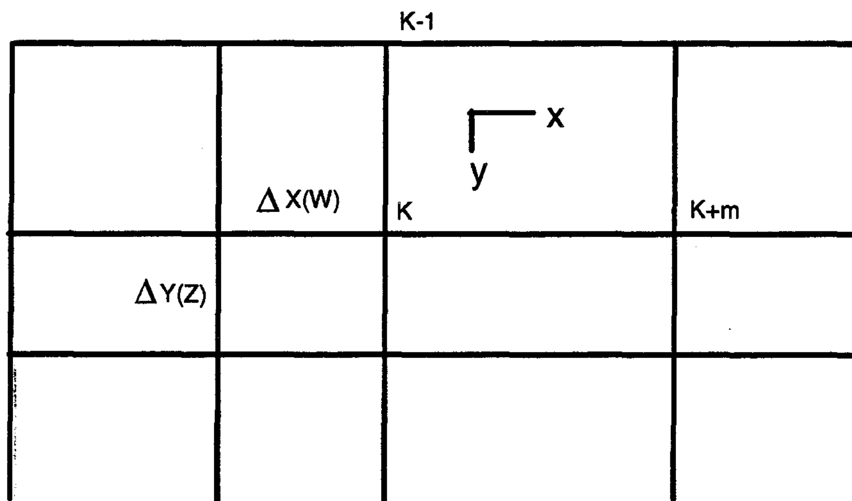


*Figure A.2: Nomenclature of a finite difference mesh where the mesh points are arranged in a vector of length m*n rather than an array*

Where K varies form 1 to mxn and the mesh steps W, Z are calculated as:

$$W = int(K/m) \tag{A.7}$$

$$Z = K - W - 1 \tag{A.8}$$

Equations A.5 and A.6 might appear already executable in a vector form but due to the existence of backward dependence in the operation, a delay, even cancelling, of vector processing is possible and it would be replaced by scalar processing. This dependence is seen in the first case (equation (A.5)) when we need to have an operand at a distance m form the first operand to perform a subtraction and the second equation an operand of a distance 1. This is remedied by an original technique introduced by the author and which resulted in a remarkable improvement in the efficiency of the calculation.

This technique is divided into three steps. The first is to eliminate the implicit dependence of the operands. This is done by creating two vectors $VS1$ which is the original vector V shifted by m steps; and a second vector $VS2$ which is V shifted by one step. The second step is to create two *separation* vectors such that $DX(W) = 1/\delta X(W)$ and $DY(Z) = 1/\delta Y(Z)$. The third step is to split the two operation so as (A.5 and A.6) become:

$$DUMMYX(K) = V(K) - VS1(K) \tag{A.9}$$

$$EX(K) = DUMMYX(K) * DX(W) \tag{A.10}$$

$$DUMMYY(K) = V(K) - VS2(K) \tag{A.11}$$

$$EY(K) = DUMMYY(K) * DY(Z) \tag{A.12}$$

Where $DUMMYX, DUMMYY$ are two transitory vectors. These four equations are then coded in four separate DO loops. Compared to scalar coding, this technique when run on a vector processor was found to be about 50 times faster although it employs two Do loops instead of one per calculation.

It is obvious that writing the simulator in vector code is an extremely complicated and delicate task. Special care had to taken specially in applying the boundary conditions which are imbedded in the vector and are not rapidly recognisable. The only

disadvantage that we found in this technique was the difficulty, but not impossibility, in modifying the model and introducing new effects.

To sum up, we have rewritten each of the model's equations in a discretised form according to a vector finite difference mesh as described above fig.(A.2). The test runs then made have proven the coherency of the code transformation and the integrity of the obtained results, not to mention the impressive speed up.[1]

A.2. Multitasking (parallel processing)

Multitasking or parallel processing is a technique aimed mainly at reducing the elapsed (turn around) time of the simulation process. It does not contribute in reducing the total CPU time consumed by the process as finally it will be the sum of the different CPU times of the different used processes.

We have used three techniques for parallel execution:

- Auto parallelism.
- Macro parallelism.
- Micro parallelism.

Whereas the first type is entirely a feature of the compiler, the other two necessitate modifying the code to benefit from them.

The primitive step in multitasking is to identify the computationally independent portions of the code. By computational independence we mean subroutines or portions of a subroutine that do not refer to each other while executing although they could use common data between them or with a third part.

Let us now consider the second type;**macro parallelism**. When we examine the simulator macroscopically for computational independence we find that we have the

[1]The cpu time necessary for one picosecond in simulating a conventional $0.3\mu m$ gate HEMT was reduced form 16 minutes to 1 minute.

following:

- The holes continuity equation which is independent of the electrons continuity and energy balance equations.
- the holes current equation which is independent of the electrons current equation.
- The updating of material parameters where we have two independent sets; field dependant and energy dependant.
- The storage of intermediate states for transient time analysis and error recovery.

In each of the above cases, simultaneous execution could be done for the different subroutines without violating the results coherency. The dispatching of the different subroutines will be managed by the main program which in turn could be used for processing a subroutine itself. Since the execution times, in this type could differ significantly (e.g. while solving the holes continuity equation which takes about as three times the time of the electrons continuity in the case of avalanche multiplication) a synchronisation stop is mandatory prior to the suite of execution to guarantee data integrity.

The **micro parallelism** is used in to forms and is applied to the DO loops constituting a subroutine or a function. Consider for example the case of solving for the electric field which was discussed earlier. We ended up by having a set of four equations written again here for coherency:

$$DUMMYX(K) = V(K) - VS1(K) \tag{A.13}$$

$$EX(K) = DUMMYX(K) * DX(W) \tag{A.14}$$

$$DUMMYY(K) = V(K) - VS2(K) \tag{A.15}$$

$$EY(K) = DUMMYY(K) * DY(Z) \tag{A.16}$$

These four equations could be regarded as two sets of computationally independent systems having a common data vector V. Each set could be then calculated separately and both sets in parallel so that the elapsed time will be equal to (assuming a unique processor occupation) the CPU time of processing one set and half what we had before. The CPU total time is then the sum of the CPU time of each processor i.e. almost unchanged. Here we have full control upon the number of processors used (2) and the number of dispatched tasks is independent of the data size.

The second form of micro parallelism is more efficient but less guaranteed. To understand what we do in this type, let us consider one of the previously discussed loops. A coding of the first equation would look like in FORTRAN:

```
DO 10 K=m+1,m*n
10 DUMMYX(K)=V(K)-VS1(K)
```

This could be seen as :

```
DO 10 K=m+1,lim1
10 DUMMYX(K)=V(K)-VS1(K)

DO 20 K=lim1+1,lim2
20 DUMMYX(K)=V(K)-VS1(K)
.
.
.
DO 90 K=limn,m*n
90 DUMMYX(K)=V(K)-VS1(K)
```

where each loop is a subdivision of the main loop. These loops are strictly computationally independent and could be executed in parallel with an increased efficiency with an increased number of processors. Thus, the number of processors, is not fixed and is fetched before the start of the execution through a special instruction.

The limits of the difference sub loops are determined according to this number of available processors with a minimum guard number of two is precised so in the worst case we will be able to divide the loop by two.

Unfortunately there are limits set upon the execution in parallel of different tasks. The major and the most annoying is that only a main task could dispatch parallel tasks. Hence, the two types of the user defined parallelism are not usable in a nested form. To optimise our program efficiency we introduced a technique of decision support that helps the simulator to determine the best path for executing the different tasks.

First, the program is run in a sequential mode during 10 time steps. The time consumed by each sub task is recorded per time step and averaged over the whole. The task that takes the longest time is considered in the main program while the other shorter ones are dispatched in parallel (macro parallelism). The long task appended to the main processor is then executed in a micro parallelism scheme. This differentiation technique is applied to the conservation equations mainly. As for the equations that are solved in all cases in a sequential manner, they use systematically the micro parallelism dispatching. Accordingly we succeeded in reducing dramatically the elapsed time to about 60% of the total CPU time. It should be mentioned that the fastest turn around time in sequential processing would be 100% of the total CPU time.

The increase in the number of processor units would increase significantly the simulator efficiency. The massively parallel machines present such a possibility through the application of SIMD (Single instruction Multiple Data) technique or evenmore MIMD(Multiple Instruction Multiple Data). We believe that extraordinary performance and speed up in the simulator would be obtained. This speed up will permit by consequence the increase in the structural complexity of the treated devices and to consider for example dynamic simulation of the devices.

To conclude, we have to point out that a vectorised code, with its magnificent performance on a vector processor, presents awfully bad performance on a scalar one. The reason for this drop in the performance is the split of the DO loops into

smaller ones with larger number of operands, longer vectors and fewer instructions. This presents a heavy load on the scalar processor. To give examples, while the program gave a performance record of 1cpu minute/1picosecond for a conventional HEMT in double precision on the IBM3090, it gave *9 hours* on an HP700 machine. A scalar code gave 16cpu minutes/1picosecond for the same device on the IBM3090 and *2 hours*/picosecond on the HP700.