Moran, David A.J. (2004) Self-aligned short gate length III-V HEMT technology. PhD thesis

http://theses.gla.ac.uk/6577/

Copyright and moral rights for this thesis are retained by the author

A copy can be downloaded for personal non-commercial research or study, without prior permission or charge

This thesis cannot be reproduced or quoted extensively from without first obtaining permission in writing from the Author

The content must not be changed in any way or sold commercially in any format or medium without the formal permission of the Author

When referring to this work, full bibliographic details including the author, title, awarding institution and date of the thesis must be given.

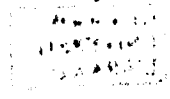
Glasgow Theses Service
http://theses.gla.ac.uk/
theses@gla.ac.uk

Self-aligned short gate length III-V HEMT technology

by

David A. J. Moran

Thesis submitted for the degree of
Doctor of Philosophy
to the
Department of Electronics and Electrical Engineering
Faculty of Engineering
University of Glasgow
January 2004



© D. A. J. Moran, 2004

BRITISH LIBRARY

IMAGING SERVICES NORTH
Boston Spa, Wetherby
West Yorkshire, LS23 7BQ
www.bl.uk

BEST COPY AVAILABLE.

VARIABLE PRINT QUALITY

Abstract

This thesis presents a new approach to the fabrication of short gate length III-V High Electron Mobility Transistors (HEMTs) that reduces the impact of external parasitic elements, and in particular access resistances, upon device performance. This was approached through the development of a self-aligned T-gate process with non-annealed ohmic contacts. The process was used to fabricate both GaAs pseudomorphic HEMT and subsequently lattice matched InP devices. In addition, a new selective recess etch was developed for cap layers containing indium.

Characterisation of the self-aligned GaAs pHEMT devices indicated good RF performance with f_T = 137GHz and f_{max} = 182GHz for devices of 120nm gate length, although DC performance was found to be restricted by the unoptimised non-annealed ohmic process. Analysis of the operation of the GaAs pHEMT devices led to the design and growth of an InP material structure incorporating double delta doping to minimise the non-annealed ohmic contact resistance. Using this optimised structure, standard and self-aligned HEMT devices with gates of length 120nm and 70nm were fabricated for comparison. The benefits and limitations of the self-aligned process were highlighted by comparing the performance of the self-aligned and standard devices. The self-aligned 120nm devices had f_T = 220GHz and f_{max} = 255GHz, which rose to f_T = 270GHz and f_{max} = 300GHz for the 70nm devices. Transconductance figures of up to 1500mS/mm were extracted for both. It is concluded that the self-aligned process, although beneficial to device performance at the 120nm, and to a lesser degree the 70nm node, would begin to degrade performance at reduced gate lengths due to increased parasitic gate capacitances.

The non-annealed ohmic technology developed in this work provides a route that minimises parasitic resistances and increases performance without the increased parasitic gate capacitances associated with a self-aligned gate approach. A possible solution for the minimisation of parasitic gate capacitances using a self-aligned approach is proposed.

"...machen Sie sich nicht lächerlich, natürlich sind Elektronen grün..."

Acknowledgments

I wish to take this opportunity to acknowledge the assistance, support and wisdom provided by many of my peers throughout this project, in particular:

First and foremost I wish to acknowledge the assistance of fellow PhD student Euan Boyd in the development of presented self-aligned GaAs pHEMT technologies, but also the entertainment provided by Euan's corridor dancing antics during those late night measurement sessions.

The experience passed on by group members which proved invaluable is also greatly appreciated, including Khaled Elgaid's big fun guide to calibration, measurement and modelling.

The support by the technical staff which I believe to be one of the key elements to the smooth operation of the Nanoelectronics research centre at Glasgow University. In particular the chocolate fuelled enthusiasm of Helen McLelland and competitive group meeting attendance which got me out of bed most Tuesday mornings. Also the dulcet tones of Susan Ferguson and memorable annual bus trip shenanigans.

Last although not least, the guidance of my supervisor and mentor Iain Thayne, who more often than not entertained if not approved of my ideas which allowed the research in this work to be performed. Also for the many curry and bonding sessions provided at the expense of Iain without which I wouldn't be the physicist/engineer I am today.

I also wish to recognise the financial support provided throughout this work by the Engineering and Physical Sciences Research Council (EPSERC) in addition to the CASE award from Bookham Caswell for the collaboration into the investigation of GaAs Metamorphic devices.

Publications

- D. Moran, E. Boyd, H. McLelland, K. Elgaid. Y.Chen, D. S. Macintyre, S. Thoms, C. R. Stanley, and I. G. Thayne, *Novel technologies for the realisation of GaAs pHEMTs with 120 nm self-aligned and nano-imprinted T-gates*, Microelectronic Engineering 2003, Volume 67 - 68, p. 769 - 774.

- D. A .J. Moran, K. Kalna, E. Boyd, F. McEwan, H. McLelland, L. L. Zhuang, C. R.Stanley, A. Asenov, I. Thayne, *Self-aligned 0.12μm T-gate $In_{.53}Ga_{.47}As/In_{.52}Al_{.48}As$ HEMT technology utilising a non-annealed ohmic contact strategy*, in *ESSDERC 2003* p. 315 – 318.

- D. A. J. Moran, E. Boyd, K. Elgaid, F. McEwan, H. McLelland, C.R. Stanley, I.G. Thayne, *Self-aligned T-gate InP HEMT realisation through double delta doping and a non-annealed ohmic process,* In print : Microelectronic Engineering, 2004

- D. A. J. Moran, E. Boyd, F. McEwan, H. McLelland, C. R. Stanley, I. G. Thayne, *Sub 100nm T-gate uniformity in InP HEMT technology*, in *GaAs Mantech 2004* p. 39 – 42.

Contents

Chapter 1

Introduction

From the dawn of the high electron mobility transistor (HEMT) in the early 1980s, extensive research has been focussed into the development of HEMT technology on a global scale. This stems directly from the continuous demand for increased performance from modern microwave/millimeter integrated circuit (MMIC) applications, and hence improvement in the performance of the transistor technology inherent to MMIC design [1.1]. The main attraction of the HEMT compared with the traditional metal semiconductor field effect transistor (MESFET) or the heterojunction bipolar transistor (HBT) lies with its unrivalled high frequency and low noise performance [1.2]. With modern MMIC applications requiring high frequency operation of 100GHz and beyond, the HEMT finds its niche in such high frequency applications. Prospective technologies in this area include broadband radio communications, automotive collision warning, and passive imaging systems to name a few, all of which have operating frequencies around 100GHz and above [1.3].

Such high frequency operation is achieved with modern HEMT technology as a result of maturing material growth technologies over the past few decades allowing the growth of more complex material structures [1.1]. This is combined with advanced lithography techniques which continue to push the limitations of ultra-small feature definition [1.4]. From the original GaAs/AlGaAs HEMT which exhibited improved performance over its counterpart, the GaAs MESFET, the III-V material system has provided a palette from which advanced HEMT devices have been realised. In particular the inclusion of indium within the GaAs device channel, which is found to improve performance considerably [1.5], has led to the development of GaAs pseudomorphic, lattice matched InP and GaAs metamorphic HEMT devices. In parallel with the development of more exotic material systems, the advance in lithography technologies has also contributed greatly to the continuous improvement in HEMT performance. The ability to define ultra-small features with electron beam (ebeam) lithography has yielded the fastest operating devices in the world to date [1.4].

Beyond the material and lithographic issues that require attention for high speed device realisation, other issues are found to influence device performance, particularly for ultra small device features (sub 100nm) [1.6]. As devices are reduced to these dimensions, effort is made to appropriately scale the various device dimensions to ensure efficient operation [1.7]. However, beyond the scaling of the intrinsic device elements, little attention is often paid to the external device parasitics and as a result these can begin to dominate performance at reduced device dimensions [1.6]. By modifying the processing technology to account for and minimise these parasitic elements, the true potential of these ultra-fast devices can be realised.

This work therefore describes the investigation into the effect of parasitic device elements, and in the particular the source and drain device resistances, and their effect upon the performance of short gate length III-V HEMT devices. This is approached through a self-aligned gate methodology, with which source and drain parasitic resistances are reduced from those associated with a standard HEMT process flow. By developing a self-aligned gate technology, the benefits associated with such a process become apparent upon comparison with standard HEMT device characterisation and can be compared with those predicted by theory.

Following this brief introduction, **Chapter 2** aims to provide insight into the detailed operation of HEMT devices and set the underlying theory for this work. The practical side of HEMT processing is then discussed in **Chapter 3** as the issues and methods of device fabrication are discussed. The various methods used to characterise devices and provide feedback as well as benchmarking the device performance are detailed in **Chapter 4**. To provide perspective for the technology presented throughout this project, a brief 'current technology review' is given in **Chapter 5**, which provides the reader with an idea of the level of similar technology published at the time of writing. **Chapters 6** and **7** then detail the relevant results from the research performed throughout this project. Finally, **Chapter 8** concludes with a summary of results and a discussion of potential future work.

References

[1.1] R. Ross, P. Lugli., *Pseudomophic HEMT Technology and Applications*, 1994, NATO ASI series. p. 1 - 10.

[1.2] P. H. Ladbrooke, *MMIC Design GaAs FETs and HEMTs*, 1989, Artech House Inc. p. 189 - 190.

[1.3] I. Thayne, *Advanced III-V HEMTs*, in *III-Vs Review*. 2003. p. 48 - 51.

[1.4] Y. E. Yamashita, A. Endoh, K. Shinohara, K. Hikosaka, T. Matsui, S. Hiyamizu, T. Mimura, *Pseudomorphic In/sub 0.52/Al/sub 0.48/As/In/sub 0.7/Ga/sub 0.3/As HEMTs with an ultrahigh f/sub T/ of 562 GHz.* Electron Device Letters, IEEE, 2002. 23(10) p. 573-575.

[1.5] R. Ross, P. Lugli, *Pseudomophic HEMT Technology and Applications*, 1994, NATO ASI series. p. 28.

[1.6] P. J. Tasker, B. Huges, *Importance of source and drain resistance to the maximum f_T of millimeter-wave MODFETs.* Electron Device Letters, 1989. 10(7) p. 291 -293.

[1.7] K. Kalna, S. Roy, A. Asenov, K. Elgaid, I. Thayne, *Scaling of pseudomorphic high electron mobility transistors to decanano dimensions.* Solid-State Electronics, 2002. 46 p. 631 - 638.

Chapter 2

The High Electron Mobility Transistor, (HEMT)

The High Electron Mobility Transistor, or HEMT, is generally favoured over the traditional Metal Field Effect Transistor, or MESFET, due to its intrinsic low noise and high frequency capabilities [2.1]. This stems directly from the incorporation of dissimilar semiconductor layers within the HEMT material structure which form heterojunctions that act to confine charge carriers in a two dimensional plane, (often referred to as a two dimensional electron gas, or 2DEG, in the case of electrons). These notable characteristics have given rise to various other acronyms for the device including the Heterojunction Field Effect Transistor, or HFET and the Modulation Doped Field Effect Transistor, or MODFET both of which refer to the same type of device.

Throughout this chapter the basic theory and operation of the HEMT is discussed extensively. As the HEMT utilises electrons as opposed to holes as its major charge carrier, examples are presented with electrons as the dominant charge carrier and hence the main contributor to current flow. The chapter opens with formation and properties of heterojunctions, followed by Schottky contact formation. Using these concepts a simple GaAs/AlGaAs HEMT structure is then considered, providing a basic model for further discussion. Device characterisation at both D.C. and RF is then discussed with relevant figures of merit, followed by the various, more complex material systems available for HEMT realisation and their benefits.

2.1 Heterojunction Formation

With the development and maturity of growth technologies in the 1980s such as Metal Organic Chemical Vapour Deposition (MOCVD) and Molecular Beam Epitaxy (MBE), the potential to create multi semiconductor layer stacks became a reality. This in turn allowed the growth of differing semiconductor layers with similar lattice constant and differing band gap energies. Whenever two dissimilar semiconductor materials are brought into intimate contact, the resulting interface is known as a *Heterojunction.* Heterojunctions exhibit many beneficial properties which have been explored in the development of various modern semiconductor devices, including the HEMT. Optoelectronic devices such as lasers, modulators and photodetectors based upon heterojunction based quantum well systems also demonstrate significant performance advantages, underpinning the expansion of the fibre-optic communication system market [2.2].

Limitations occur as to the choice of particular adjacent layers within the material layer stack however, as the lattice parameter of each layer must remain constant to a degree of tolerable strain to minimise defects throughout the crystal [2.3].

Figure 2.1.1 shows the lattice parameter and band gap energy for various III-V materials used in current HEMT technology [2.3]:

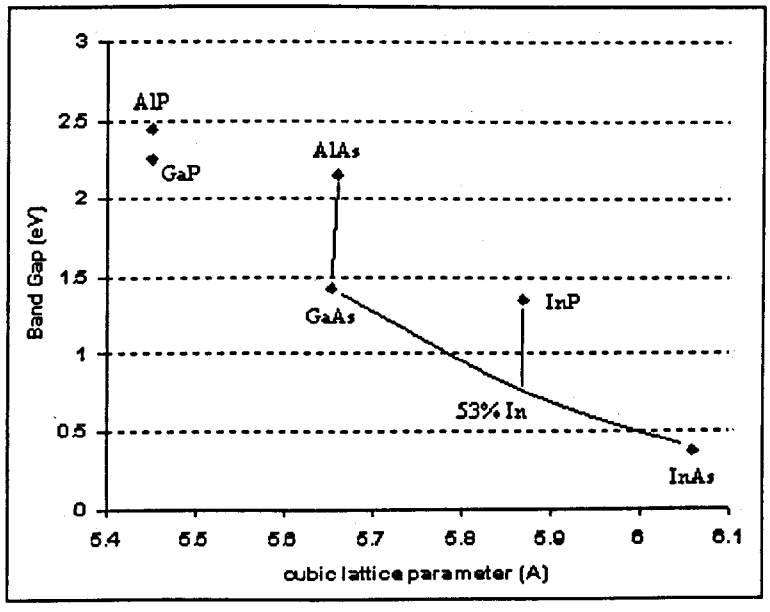


Figure 2.1.1 - Bandgap energy vs. Lattice parameter for III-V materials

As shown in **Figure 2.1.1**, as aluminium is introduced into the compound GaAs, replacing the gallium, i.e. $Al_x Ga_{1-x} As$, the lattice parameter of the structure remains similar. The difference in lattice parameter between GaAs and AlAs amounts to less than 1%, though the difference in band gap energies between the two compounds GaAs and AlAs is of the order 0.73eV.

The GaAs/AlGaAs Heterojunction (an example)

To investigate the process of heterojunction formation, a GaAs/AlGaAs heterojunction is considered. The energy band diagrams for each material in isolation are given in **Figure 2.1.2**. In this instance the AlGaAs compound has been n^+ doped, so an additional material such as silicon has been introduced to supply negative charge carriers. The GaAs is intrinsic and is hence dopant free.

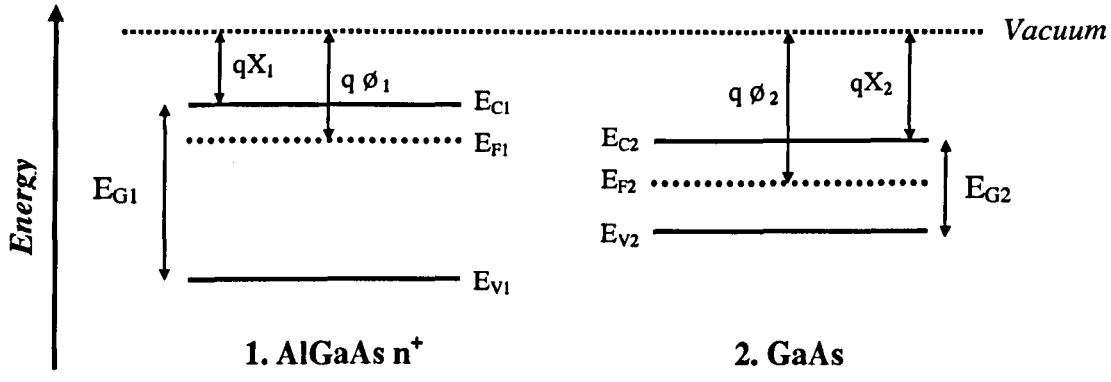


Figure 2.1.2 Energy band diagrams for isolated 1.AlGaAs n^+ and 2.GaAs

From **Figure 2.1.2**, the bandgap energy E_{Gi} separates the conduction and valence band edges E_{Ci} and E_{Vi} for both materials. Intrinsically the value of the bandgap energy for AlGaAs is larger than that for GaAs, and increases linearly with increased aluminium content. The Fermi energy E_{Fi} is also illustrated for each material. The electron affinity qX_i and the work function $q\,\phi_i$ are defined as the energies required to remove one electron from the conduction band edge E_C and the Fermi level E_F to the vacuum level respectively, (the vacuum level being defined at a position outside the material where the potential that acts to confine

carriers is effectively zero). From the Fermi distribution function, the Fermi level is defined at an energy where the probability of a state being full is exactly 1/2 [2.4]. For an intrinsic semiconductor, (such as GaAs in the above example), the Fermi level can be defined in terms of the conduction and valence band energies, the density of states in each of the bands, and the material temperature [2.4]:

$$E_F = \frac{E_C + E_V}{2} + \frac{kT}{2} \ln\left[\frac{N_V}{N_C}\right] \qquad \textbf{Eqn 2.1.1}$$

From **Eqn 2.1.1**, k is the Boltzmann constant, T the material temperature, and N_V and N_C the effective density of states in the valence and conduction bands.

For many materials at room temperature, the second term in **Eqn 2.1.1** is much smaller than the bandgap energy and so the Fermi level lies almost equidistant from conduction and valence band edges, i.e. mid bandgap. Such is the case with GaAs at room temperature in which the second term equals 0.035eV, compared to the band gap energy of 1.42eV [2.5].
With the introduction of a dopant to the semiconductor, this relation no longer holds. Typically for III-V materials, silicon is used as an n-type dopant, which supplies additional electrons to the conduction band. The energy level of these 'donor' atoms lies near to the conduction band edge of the material and are ionised at room temperature, filling the conduction band with additional electrons.

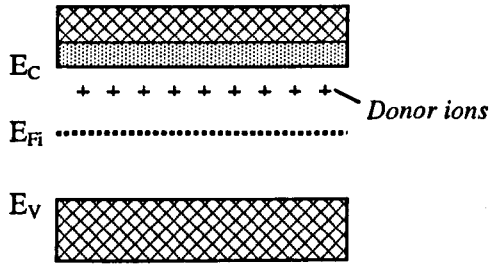


Figure 2.1.3
Band diagram with n-type doping

The expression for the Fermi level can now be rewritten to include the effect of the doping [2.4]:

$$E_F = E_C - kT \ln\left[\frac{N_C}{N_D}\right] \qquad \textbf{Eqn 2.1.2}$$

Assuming complete ionisation of the donor impurities, the energy difference between the Fermi level and the conduction band edge can therefore be related to the material temperature T, the effective density of states of the conduction band N_C and the donor concentration N_D. From **Eqn 2.1.2**, at a constant temperature, as the donor concentration is increased, the value of the second term decreases, pulling the Fermi level towards the conduction band edge.

Upon heterojunction formation, the Fermi level must remain continuous and flat across the interface between the two semiconductors. In addition the vacuum level must also remain continuous, but parallel to the band edges to maintain the values of qX_1 and qX_2. Considering again n^+ AlGaAs and GaAs as in **Figure 2.1.2**, if the two materials are brought into intimate contact a heterojunction is formed. Diffusion of higher energy electrons occurs from the n^+ AlGaAs to the GaAs. This in turn uncovers positive charge in the form of the donor ions near the interface within the AlGaAs, and gives rise to an electric field between these ions and the accumulated electrons within the GaAs. Once a specific amount of negative charge has accumulated within the GaAs, this field opposes the transfer of further electrons from the AlGaAs to the GaAs and an equilibrium is reached. In addition this field acts to confine diffused electrons within the GaAs to a narrow, almost triangular shaped well. This process is illustrated in **Figure 2.1.4**.

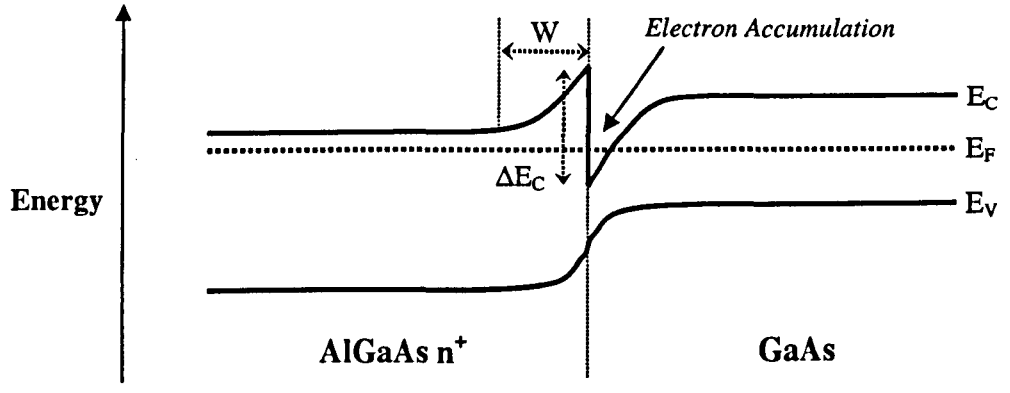


Figure 2.1.4 The AlGaAs n^+ / GaAs heterojunction

Figure 2.1.4 demonstrates the resultant energy band structure for a typical n^+ AlGaAs / GaAs heterojunction. The conduction band offset ΔE_C represents the magnitude of the discontinuity between the two conduction band edges at the interface. For non-degenerate semiconductors, such as GaAs and AlGaAs, this discontinuity is independent of doping and is simply the difference between the electron affinities for the two materials, i.e.$\Delta E_C = q (X_2 - X_1)$, [2.6]. The field created between the electrons accumulated in the GaAs and the uncovered dopant ions in the AlGaAs is represented by a ramp in the conduction band edge within the AlGaAs. The width of this 'ramp' is effectively the size of the depleted region into the AlGaAs, (represented by W in **Figure 2.1.4)**, the magnitude of which can be determined from solution of the Poisson equation.

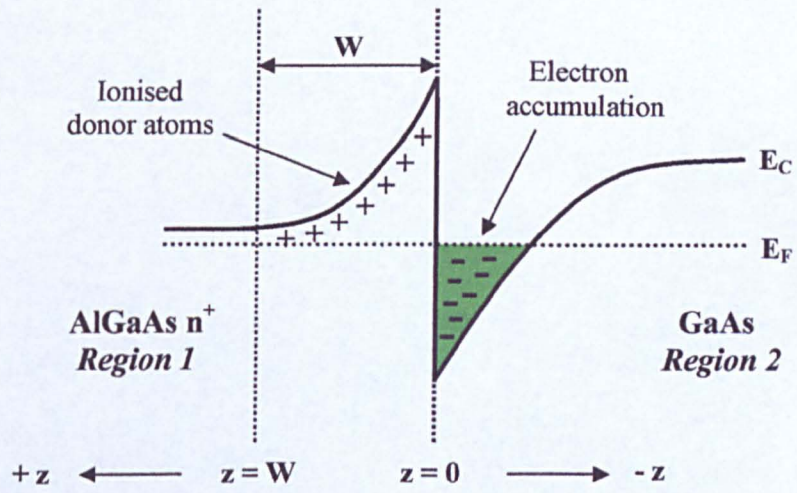


Figure 2.1.5 The AlGaAs n^+ / GaAs heterojunction conduction band profile

The magnitude of the dimension of the well formed within the GaAs in which the electrons accumulate is generally of the order of the electron wavelength. Boundary conditions imposed upon the electron population within the well then result in quantisation of the electron momentum perpendicular to the interface and discrete energy levels or states are formed throughout the well. **Figure 2.1.6** demonstrates the formation of discrete energy levels E_1 and E_2 below the Fermi level E_F within a triangular shaped well similar to that formed at an AlGaAs n^+ / GaAs interface.

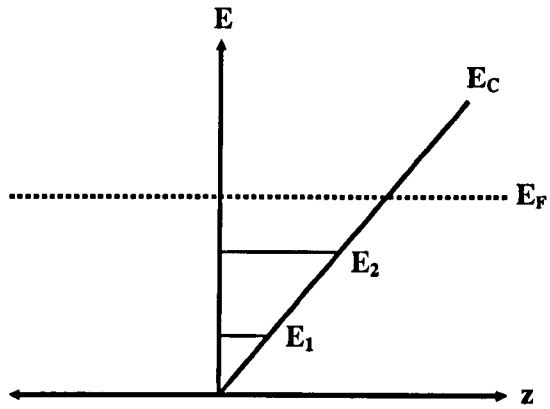


Figure 2.1.6 Energy levels within a triangular well

Figure 2.1.5 presents the details of the conduction band profile of the AlGaAs n^+ / GaAs interface. By defining a 1D co-ordinate system perpendicular to the interface as z, and taking the origin of this to be at the interface, the electric field E at the interface i.e. z = 0 in the GaAs, can be related to the carrier concentration of the 2DEG N_S, and the permittivity of the GaAs ε_{GaAs} by Gauss's law as [2.7]:

$$E_{interface} = \frac{-q N_S}{\varepsilon_{GaAs}} \qquad \textbf{Eqn 2.1.3}$$

In *Region 1*, the electric field will taper off from its maximum value at the interface until at z = W it becomes zero, i.e. at the end of the depleted region within the AlGaAs. From the 1D Poisson equation, the rate of change of field with distance within the AlGaAs can be defined as:

$$\frac{dE}{dz} = \frac{q N_d}{\varepsilon_{AlGaAs}} \qquad \textbf{Eqn 2.1.4}$$

where N_d is the uniform doping concentration within the AlGaAs and ε_{AlGaAs} the permittivity. This can be integrated to produce an expression for the field E at a point z:

$$E_{(z)} = \frac{q N_d z}{\varepsilon_{AlGaAs}} + \text{Constant} \qquad \textbf{Eqn 2.1.5}$$

If the two permittivities are assumed to be equal, i.e. $\varepsilon_{GaAs} = \varepsilon_{AlGaAs}$, from **Eqn 2.1.3**, E is defined at z = 0, and **Eqn 2.1.5** becomes:

$$E_{(z)} = \frac{q}{\varepsilon_{AlGaAs}}(N_d z - N_S) \qquad \textbf{Eqn 2.1.6}$$

At $z = W$, $E_{(z)} = 0$, and **Eqn 2.1.6** can be re-written:

$$N_d W = N_S \qquad \textbf{Eqn 2.1.7}$$

Hence the depletion depth W can be expressed as the ratio of the 2D carrier concentration in the GaAs triangular well N_S, and the 3D doping concentration in the AlGaAs N_d.

2.2 Metal - Semiconductor interfaces - (The Schottky contact)

The ability to realise complex heterostructures with current growth techniques has led to the design and development of various modern semiconductor devices. However in a real device, beyond the ability to include such heterostructures, there must exist the potential to create a connection between the device and the outside world. These connections or contacts generally allow the passage of current into and from, or the biasing of the device. In HEMT technology, examples of these can be found with the source and drain contacts which are ohmic in nature to allow the passage of current in both directions, and with the gate contact which is rectifying and acts to modulate the source-drain current typical of any field effect transistor. Both of these types of contact are crucial to the operation and performance of III-V HEMT devices, and therefore will be discussed thoroughly in this section.

The Schottky Contact

A Schottky contact is formed whenever a metal comes into intimate contact with a semiconducting material. This process is not dissimilar to the formation of a heterojunction between two semiconducting materials, and can indeed be thought of as a heterojunction with one of the semiconductors being extremely highly doped to take the role of the metal. The energy band diagrams for an isolated metal and n^+ semiconductor are presented in **Figure 2.2.1a**. In this example, the n^+ doping in the semiconductor results in the Fermi level being close to the conduction band edge. The electron affinity qX_{SC} and the work function $q\phi_{SC}$ for the semiconductor are shown together with the work function for the metal $q\phi_M$ with respect to the vacuum level.

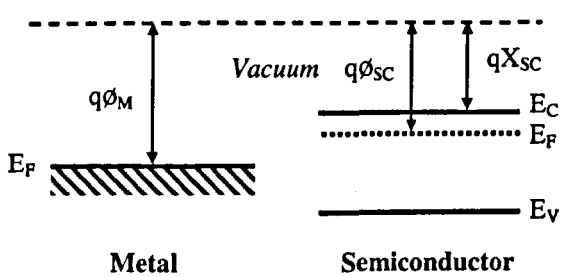


Figure 2.2.1a Isolated metal and semiconductor energy band diagrams

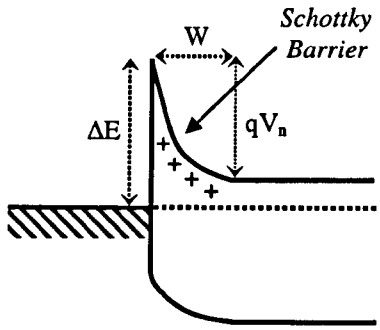


Figure 2.2.1b Schottky barrier formation

If the two are brought into intimate contact the Fermi levels of each join as electron diffusion occurs between the semiconductor and the metal. A thin accumulation layer of negative charge then forms within the metal next to the interface as electrons migrate from the semiconductor to the metal. Again the loss of negative charge near the interface within the semiconductor uncovers the positive ion cores from the dopant atoms. The combination of the two charged regions creates an electric field which opposes further electron flow from the semiconductor. As before this field is represented by a potential barrier in the conduction band at the interface which tapers off further into the semiconductor. The end result is presented in **Figure 2.2.1b**, where the resultant barrier, often referred to as the *Schottky Barrier,* rises to a height ΔE above the Fermi level, and to a depth W into the semiconductor. The built in voltage V_n separates the flat conduction band edge and the top of the barrier within the semiconductor. The depletion depth W into the semiconductor can be determined in a similar fashion as in the analysis of the AlGaAs n^+ / GaAs heterojunction presented in **Section 2.1**:

Again, the field E at the interface can be expressed as:

$$E_{interface} = \frac{q N_d W}{\varepsilon_{SC}} \qquad \textbf{Eqn 2.2.1}$$

Where N_d is the doping concentration of the semiconductor and ε_{SC} the permittivity.

Assuming an abrupt interface model where the charge density regions end abruptly, the built in bias within the semiconductor V_n can be related to the field at the interface $E_{interface}$ and the depletion depth W as:

$$V_n = \frac{E_{interface}\, W}{2} \qquad \textbf{Eqn 2.2.2}$$

From **Eqn 2.2.1** and **2.2.2** the depletion depth can be expressed as:

$$W = \left(\frac{2\,\varepsilon_{SC}}{q\,N_d} V_n\right)^{1/2} \qquad \textbf{Eqn 2.2.3}$$

Hence very highly doped materials will have a much narrower Schottky barrier as W decreases with increased N_d.

The Schottky barrier height ΔE, using this Schottky model, is simply the difference in the semiconductor electron affinity qX_{SC} and the metal work function $q\varnothing_M$, i.e. $\Delta E = q(\varnothing_M - X_{SC})$. This is however a simplistic model and in reality the true barrier height is often found to be less dependent on the metal work function. This results from the formation of mid bandgap surface states at the metal - semiconductor interface which can act to raise or lower the barrier height and in the extreme case, pin the Fermi level [2.8].

Surface States and Fermi level pinning

The interruption of the periodicity of the crystal lattice of a semiconductor and metal at their mutual interface gives rise to surface states existing within a thin region at this interface. The concentration of such states will depend on various factors including the crystal orientation of the two materials, the quality of the interface between them and the semiconductor doping [2.8]. It has been argued that if the surface state concentration is high enough, the Fermi level will be 'pinned' at a particular energy at the interface, where the net surface charge due to the filling of such states is neutral. This extreme model therefore predicts, in opposition to the Schottky model, that the resultant barrier height will be independent of the metal work function.

In reality the measured Schottky barrier heights of various metal - semiconductor interfaces by different methods have indicated that depending on the materials in question, the Schottky or Fermi Pinning models provide upper and lower limits to the actual value for the Schottky

barrier height [2.9]. It is generally accepted that for III-V materials, the Fermi pinning model better describes barrier formation, as the barrier height is far less dependent on the metal work function [2.10]. However the exact nature of the formation of such metal - semiconductor interfaces still remains a point of great debate.

Current transport

The nature of current transport across a Schottky barrier as with any potential barrier will depend on the height and width of the barrier in question, as well as the temperature of the material system. The mechanisms by which current flows through a metal-semiconductor contact can be categorised into either *Thermionic Emission* (TE) or *Field Emission* (FE) (Direct Tunnelling), processes. The former involves the probability of a carrier possessing enough energy to pass over the barrier, while the latter relies on quantum tunnelling through the barrier at the Fermi level.

At energies above the Fermi level where tunnelling still occurs albeit by thermally excited carriers, the process is known as *Thermionic Field Emission* (TFE) and is merely a combination of the two.

By choosing the materials and doping concentration accordingly, the Schottky contact can be tailored to behave as rectifying or ohmic in nature. To examine this process in more detail, a Schottky contact is presented under both forward and reverse bias.

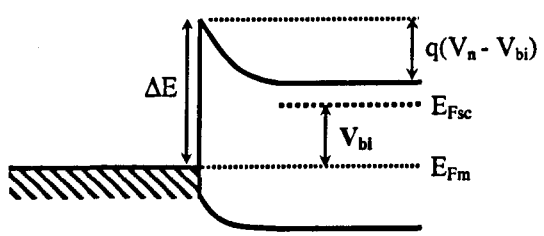


Figure 2.2.2a - Schottky contact under forward bias ($+V_{bi}$)

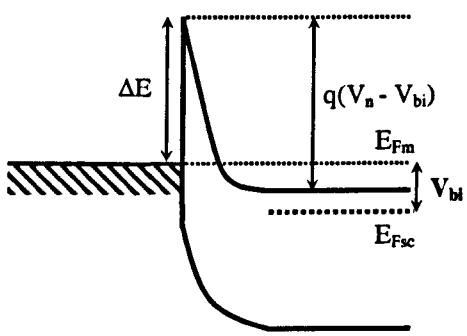


Figure 2.2.2b - Schottky contact under reverse bias ($-V_{bi}$)

Considering again an n^+ semiconductor in contact with a metal, **Figure 2.2.2a** shows the case of the contact under a forward bias V_{bi} while **Figure 2.2.2b** demonstrates the same contact under a reverse bias of similar magnitude.

In the case of the forward bias, a positive potential is introduced between the semiconductor and metal, separating the Fermi levels by V_{bi} and raising the conduction and valence bands above their equilibrium levels by a similar amount. The opposite is found under reverse bias conditions where the bands are pulled down from their former positions by the potential difference qV_{bi}. Figures **2.2.2a** and **2.2.2b** demonstrate the significant change in the size and shape of the barrier that can occur under both forward and reverse biases, which significantly affect the transport processes across it. These are examined by considering both thermionic emission and field emission processes for both forward and reverse bias conditions. Firstly, thermionic emission is discussed.

Thermionic Emission

As the name suggests, the process of thermionic emission depends largely on the temperature of the medium of transport. As the temperature of the material increases from 0K, the discrete Fermi-Dirac distribution function becomes 'smeared' around the Fermi level [2.4]. This results from thermal energy exciting carriers to energies of the order of a few kT, (k being Boltmann's constant and T being the absolute temperature) beyond that of the Fermi level. Depending therefore on the barrier height, carriers occupying states above this 'height' will readily drift over the barrier under the influence of an external electric field. From a classical point of view, this process simply represents the carriers that possess enough kinetic energy to pass through the high field region at the interface and hence pass over the barrier. To better explain this, the process of thermionic emission across a Schottky barrier from semiconductor to metal is illustrated in **Figure 2.2.3**:

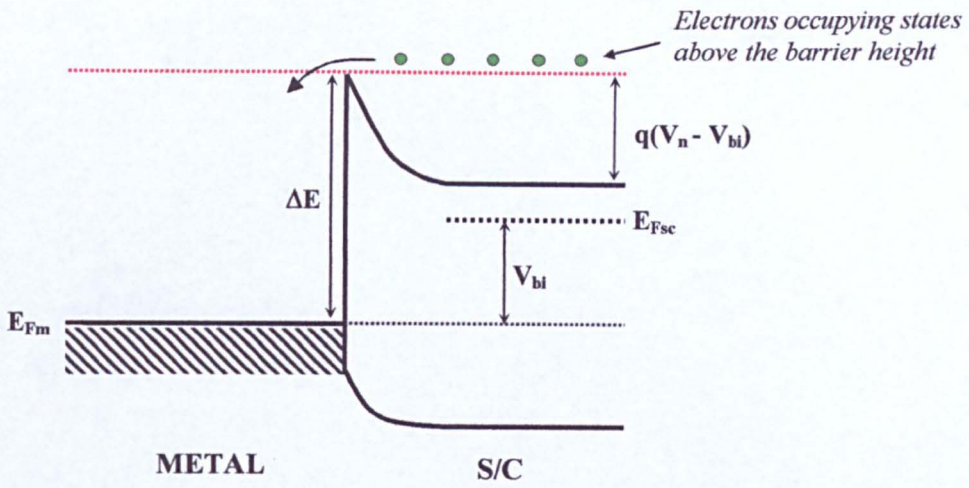


Figure 2.2.3 - Thermionic emission between semiconductor and metal under forward bias

In this example the case of a Schottky contact under forward bias is re-examined. Upon the application of the external bias V_{bi}, electrons occupying states above the height of the barrier, (defined as $q(V_n - V_{bi})$ from the conduction band edge in **Figure 2.2.3**), will drift over the barrier into the metal.

For a particular energy level in the conduction band, the concentration of electrons occupying states above this level with respect to the Fermi level is related to the system temperature and density of states by the Boltzmann distribution [2.4]:

$$n = N_C \exp\left(-\frac{E - E_F}{k_B T}\right) \qquad \textbf{Eqn 2.2.4}$$

where: n is the carrier concentration, N_C is the density of states beyond the level in question, E is the energy level of interest, E_F is the Fermi energy, T is the system temperature.

Eqn 2.2.4 demonstrates the strong dependence of the carrier concentration on the system temperature as the exponential term increases with temperature.

For the case of carriers in the metal crossing into the semiconductor, the barrier height will be virtually independent of applied bias. For carriers in the semiconductor crossing into the

metal this process is complicated by the barrier height being directly dependent on the applied bias. This becomes apparent in **Figures 2.2.2a & b** and **2.2.3** in which the barrier height within the semiconductor expressed with respect to the conduction band edge is defined as $E_B = q(V_n - V_{bi})$. With zero external bias, this is simply the built in potential qV_n. By re-defining the barrier height in the semiconductor as the energy difference between the Fermi level and the top of the barrier, the current density across the barrier for a specific bias can be related to the barrier height, E_B as:

$$J \sim \exp\left(-\frac{E_B}{k_B T}\right) \qquad \textbf{Eqn 2.2.5}$$

as the current density will be directly proportional to the carrier concentration above the height of the barrier. Hence at a constant temperature, the current density is found to have an inverse exponential dependence on the barrier height. This results in a thermionic emission dominated Schottky contact being rectifying or diode like in nature, as current will flow more readily from the semiconductor to the metal than vice versa as a result of barrier lowering in the semiconductor by applied bias.

Field Emission

The process of field emission relies on the quantum mechanical tunnelling of carriers through the Schottky barrier from either metal to semiconductor or vice versa for current to flow [2.11]. The magnitude of such current flow depends on the probability of electrons tunnelling through a particular barrier. Under pure field emission transport, only electrons near to the Fermi level are considered to tunnel through the barrier and contribute to the tunnelling current. This results in the field emission process being far less dependent on the system temperature than with thermionic emission processes as the electron-energy distribution is of less consequence. The tunnelling processes through a Schottky barrier are illustrated in **Figure 2.2.4**:

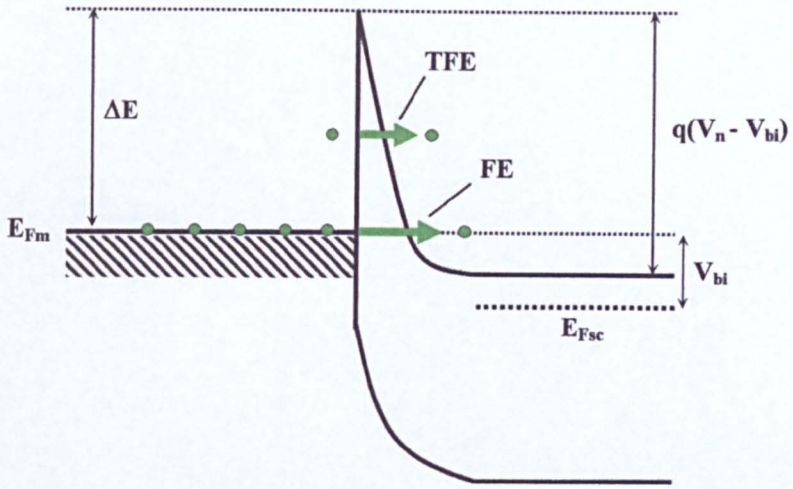


Figure 2.2.4 - Field Emission (FE) and Thermionic Field Emission (TFE) processes between metal and semiconductor under reverse bias.

In contrast to the thermionic emission example, FE and TFE processes are examined across a Schottky contact under reverse bias in which current will flow from metal to semiconductor. Tunnelling of electrons through the barrier occurs at the Fermi level and at higher exited levels defining the FE and TFE processes respectively. The details of carrier tunnelling through potential barriers has been examined by various sources [2.11-13], the details of which become considerably complex and hence shall not be discussed here. Instead the main points of relevance are summarised.

The tunnelling probability and hence tunnelling current is found to be dependent on both the barrier height and width, and is related to both as [2.11]:

$$\text{Tunnelling Prob. / Current} \sim \exp\left[-\frac{q}{\hbar}(2\,m\,\Delta E)^{1/2}\,W\right] \qquad \textbf{Eqn 2.2.6}$$

Where : m is the carrier effective mass in material (kg), ΔE is the effective barrier height (eV), W is the depletion width into the semiconductor or the barrier width (m).

Eqn 2.2.6 demonstrates the strong relation between the magnitude of the tunnelling current and the barrier width W as the current will decrease exponentially with increased barrier thickness. The height of the barrier ΔE also plays a lesser role as the current varies exponentially with the square root of the barrier height.

Incorporating the expression defined for the depletion depth W from **Eqn 2.2.3** and defining the barrier height in terms of the built in bias V_n, i.e. $V_n = \Delta E/q$, **Eqn 2.2.6** becomes:

$$\text{Tunnelling Prob. / Current} \sim \exp\left[-\frac{2\Delta E}{\hbar}\left(\frac{\varepsilon_S m}{N_D}\right)^{1/2}\right] \qquad \textbf{Eqn 2.2.7}$$

Therefore the larger the doping concentration within the semiconductor, the narrower the barrier and hence the higher the tunnelling probability.

For very highly doped semiconductors, the Schottky barrier can become very narrow to the point that field emission becomes the dominant process of transport. Generally under these conditions the current - voltage response of the contact becomes ohmic as opposed to rectifying i.e. current will flow in equal magnitude from the metal to the semiconductor and vice versa with forward and reverse bias.

2.3 The GaAs/AlGaAs HEMT (an example)

Using the concepts discussed in **Sections 2.1** and **2.2**, a model for a basic HEMT structure can now be constructed. So far the GaAs/AlGaAs heterojunction has served as an example for the discussion of heterojunction formation. This example will be taken further and used to describe the layer structure of a basic GaAs/AlGaAs HEMT. Furthermore, the nature and operation of the contacts of the device are better understood from the discussion of Schottky contact formation and current transport mechanisms from **Section 2.2**.

GaAs/AlGaAs HEMT model

The basic structure for a GaAs/AlGaAs HEMT is presented in **Figure 2.3.1**.

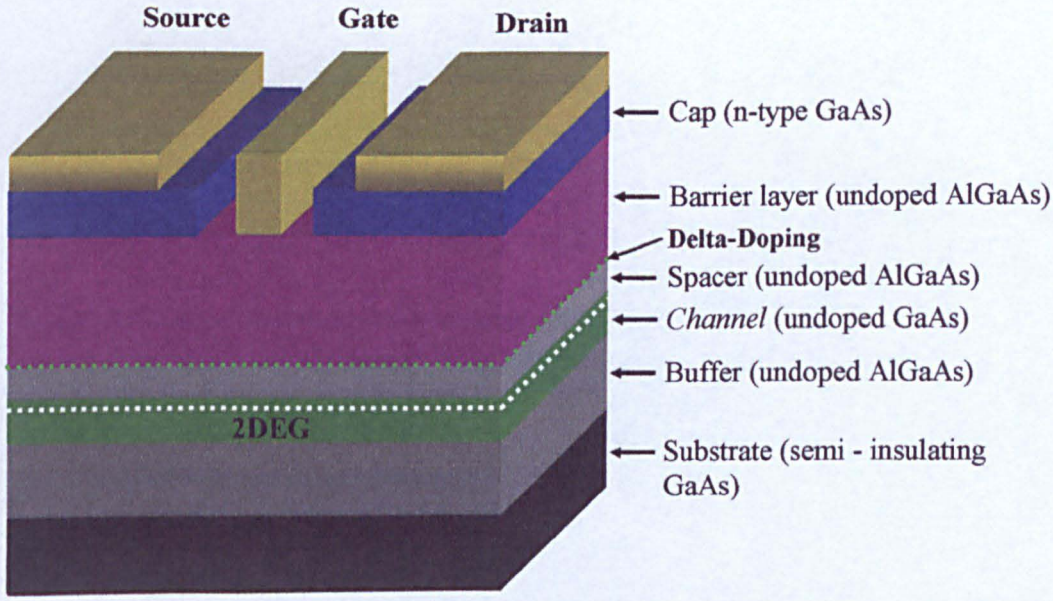


Figure 2.3.1 - GaAs/AlGaAs HEMT model

The GaAs/AlGaAs HEMT device structure is comprised of a combination of doped and undoped GaAs and AlGaAs semiconductor layers all of which are grown upon a thick semi-insulating GaAs substrate, with metal contacts forming the source, gate and drain. The GaAs *Channel* in which the 2DEG is formed is undoped and is sandwiched between *Spacer* and *Buffer* layers comprising of undoped AlGaAs. The AlGaAs buffer acts to further confine carriers within the GaAs channel by forming a heterojunction barrier at the underside of the

GaAs layer. The AlGaAs spacer separates the channel from a layer of delta doping. Delta doping is simply a very thin layer (in the order of angstroms) of highly doped material that can be located to provide doping of selective layers of a heterostructure. For the GaAs/AlGaAs HEMT example, this doping provides electrons which diffuse into the underlying GaAs layer that forms the 2DEG within the device channel. By separating the 2DEG carriers from their parent donor ions, they are less subject to coulombic scattering. As a result, the channel carrier mobility is increased, improving device performance. The specifics of carrier transport within a device channel such as carrier mobility are discussed further in **Section 2.6.** Above the delta doping a thick layer of AlGaAs known as the *Barrier* layer is grown followed by the *Cap* layer consisting of bulk n-doped GaAs. The source and drain metal contacts which are ohmic are deposited onto the cap layer and the metal gate which is rectifying, onto the undoped AlGaAs barrier layer.

The cross section structure of a GaAs/AlGaAs HEMT presented in **Figure 2.3.1** can be divided into two main regions of interest, namely the region under the gate contact in which the channel depletion process occurs, and that outside in which current enters and leaves the gate region through the source and drain contacts.
Firstly the gate region is considered.

2DEG concentration modulation

In typical FET operation, current flowing through the device channel from the source to the drain is modulated by a voltage applied to the gate contact, usually with respect to the source. For the AlGaAs/GaAs HEMT, this process is achieved by increasing the potential across the Schottky barrier into the material formed at the gate metal / AlGaAs barrier layer by applying a negative bias $-V_G$ across the interface. Forming the gate contact onto the intrinsic AlGaAs barrier instead of the doped GaAs cap provides several benefits. Firstly, the Schottky barrier will be larger with the AlGaAs - metal interface over that with doped GaAs due to the AlGaAs forming a higher barrier. This combined with the lack of doping ensures

that the majority of free charge through the structure is confined to the channel layer and minimises current leakage through the gate. Additionally, in removing a section of the cap and forming the gate in a 'trench' the shape of the effective depletion region is altered from that which would be formed on a planar surface. The shape and size of the depletion region is found to play a large role in device performance, and is discussed more extensively in **Sections 2.4**, **2.5** and **2.6**.

As the potential across the Schottky barrier is increased with applied negative voltage to the gate, depletion of the electron concentration within the 2DEG occurs with more negative gate bias. This continues until at a particular voltage V_p applied to the gate, the carrier concentration in the area below the gate is effectively reduced to zero. This process is comparable to the charging and discharging of a parallel plate capacitor, with the gate as one plate and the channel the other, and with the gate voltage applied between them. Current then flowing from source to drain under external bias will experience this reduction in carrier concentration as an increase in channel impedance in the gate region, reducing source to drain current flow.

Figure 2.3.2 illustrates by example the gate region layer structure of a typical AlGaAs/GaAs HEMT.

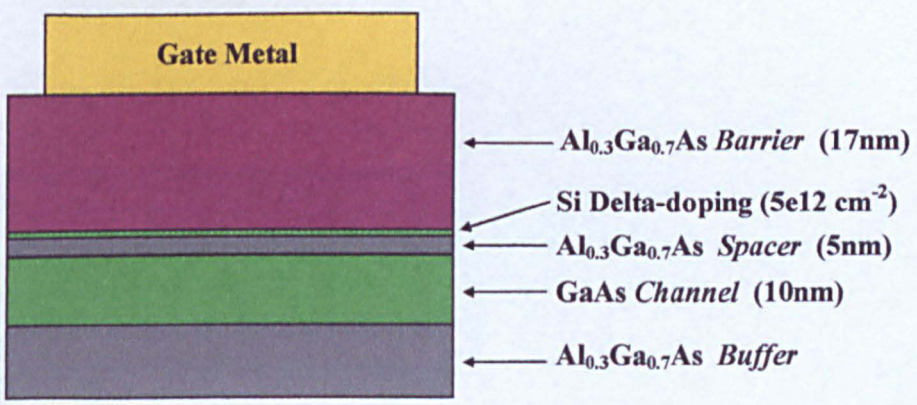


Figure 2.3.2 - GaAs/AlGaAs HEMT gate region structure

From **Figure 2.3.2**, the aluminium composition of the AlGaAs layers chosen for this example is 30%, i.e. $Al_{0.3}Ga_{0.7}As$. The delta-doping concentration has also been taken to be 5e12 cm^{-2}, a typical value for III-V HEMT structures. The dimensions of the layers are also chosen to scale with the effective length of the gate contact, which becomes more crucial at shorter gate lengths. The scaling of HEMT material structures is discussed further in **Section 2.6**. To better understand the nature of such a structure, the conduction band and carrier concentration profiles vertically through the structure are considered. These were extracted using Greg Snider's 1D Poisson / Schroedinger solver and in doing so the conduction band and carrier concentration profiles for the structure in **Figure 2.3.2** are presented in **Figure 2.3.3**:

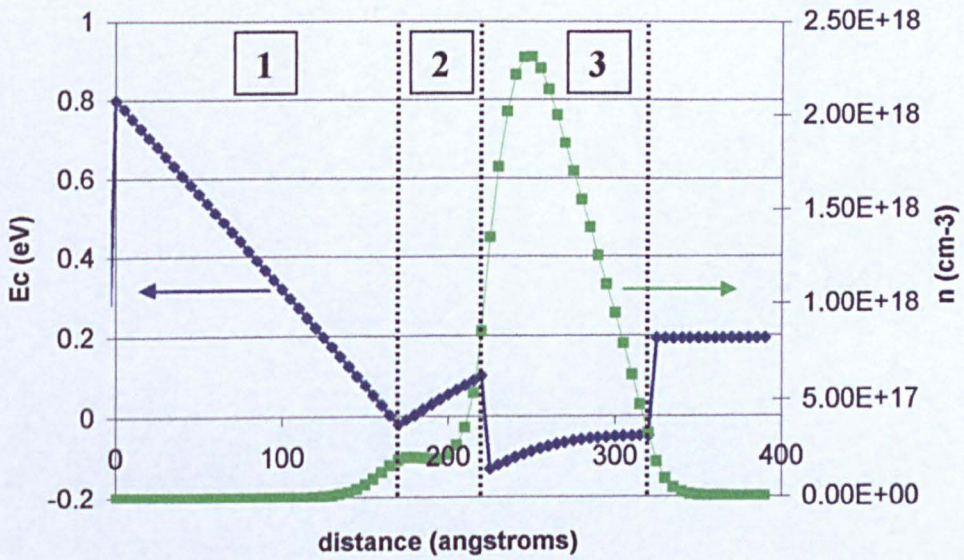


Figure 2.3.3 - Conduction Band and carrier concentration profiles of GaAs/AlGaAs HEMT gate region

In **Figure 2.3.3**, the left hand axis displays energy with the origin taken at the Fermi level, while the right axis gives the carrier concentration. Depth into the structure is given along the bottom axis in angstroms with the gate contact / AlGaAs barrier interface given at the origin. The resultant conduction band and carrier concentration profiles are divided into regions **1**, **2** and **3** corresponding to layers $Al_{0.3}Ga_{0.7}As$ barrier, $Al_{0.3}Ga_{0.7}As$ spacer and GaAs channel

respectively. The thin layer of delta doping, (taken to be no more than 5 angstroms in width), is situated between regions 1 and 2. The height of the Schottky barrier formed at the left edge of region 1 is set to 0.8eV above the Fermi level with zero external bias [2.14].The majority of the free electron population is confined to within the GaAs channel layer. This is represented by the large peak in the carrier concentration in Region 3, forming the channel 2DEG. It is of course desirable for the large majority of the electron population to be confined to the device channel. In the instance that a sizable concentration of electrons accumulates outwith the channel and contributes to the total current flow between source and drain, a parasitic channel is formed and *parallel conduction* occurs. This is a highly undesirable effect as the transport properties through these adjacent parasitic layers are greatly reduced compared to those of the specified channel, and hence the overall device performance is reduced. This process can be avoided by designing the layer structure to ensure complete depletion of the layers above the channel with zero applied gate bias, whilst maintaining a large channel carrier concentration. By applying a negative bias to the gate contact, the Fermi level within the semiconductor is raised above its equilibrium position further increasing the depletion effect into the device, reducing the electron concentration within the channel.

A HEMT device is therefore said to work in *depletion mode*, which emphasises control of the channel current through this depletion process, and where such current will flow with zero gate bias. In contrast, *enhancement mode* devices exist in which source - drain current will only flow upon the application of a positive gate voltage. These typically rely on the formation of the gate contact onto a thin film of dielectric deposited onto the semiconductor that acts to increase the depletion effect into the structure and channel while maintaining an efficient gate to channel separation.

Ohmic contact formation

As with the gate contact, the process of forming an effective ohmic contact can be better understood by considering the conduction band and carrier concentration profiles within the

ohmic contact regions of the layer structure. From a typical GaAs/AlGaAs HEMT, the ohmic contact region structure is presented in **Figure 2.3.4**:

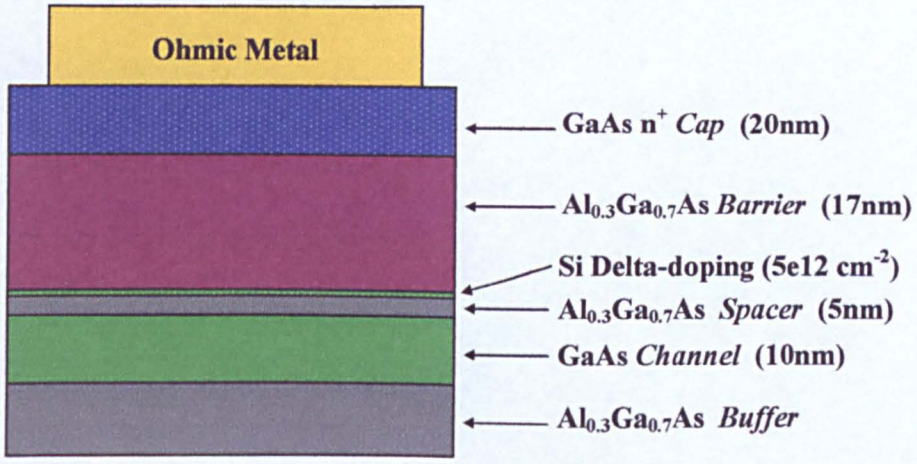


Figure 2.3.4 - GaAs/AlGaAs HEMT ohmic contact region structure

The structure is similar to that presented for the gate region, except the highly n^+ doped GaAs cap remains situated on top of the AlGaAs barrier and is in contact with the ohmic metallisation. The role of the ohmic contact is to provide low-resistance access to the device channel. Electrons must therefore be able to pass between the channel and ohmic metal with minimal resistance relative to the resistance associated through the active region of the device, i.e. the gate region. Observation of the conduction band and carrier concentration profiles from the structure and specifically the potential barriers that arise throughout the structure provide insight into the transport processes through the contact. Conduction band and carrier concentration profiles from the structure given in **Figure 2.3.4** are presented in **Figure 2.3.5**:

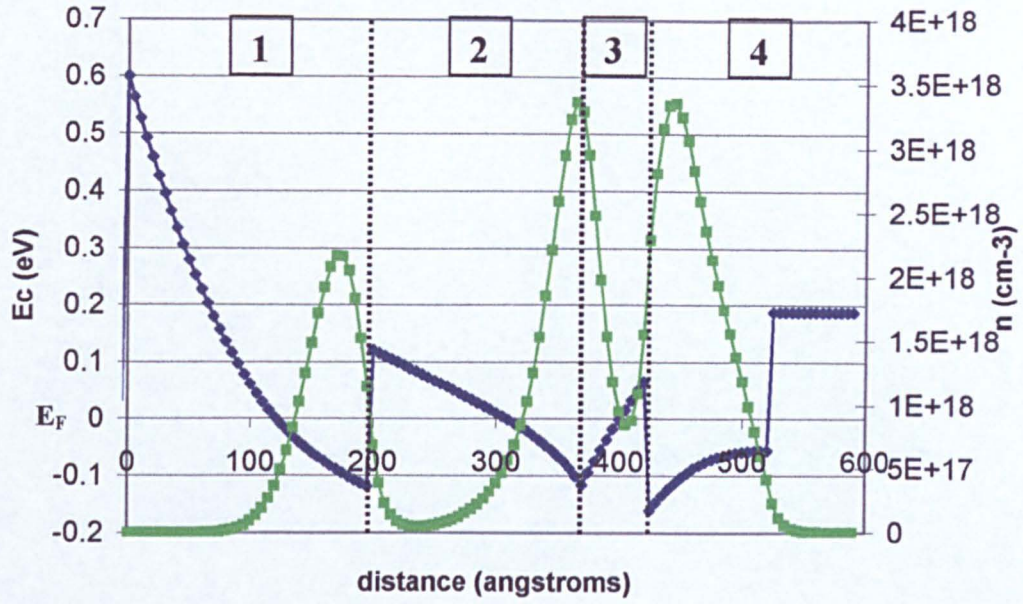


Figure 2.3.5 - Conduction Band and carrier concentration profiles of GaAs/AlGaAs HEMT ohmic region

The structure is divided into **Regions 1**, **2**, **3** and **4** which correspond to the cap, barrier, spacer and channel layers respectively. As before the horizontal axis represents depth into the structure from the metal / cap interface set at the origin, with the left and right vertical axes representing conduction band energy and carrier concentrations respectively. The Fermi level is set as zero on the conduction band axis. The resultant conduction band profile demonstrates the formation of various potential barriers throughout the structure, i.e. in **Region 1** the large Schottky barrier formed at cap / metal interface, and in **2** and **3** the barriers formed by the AlGaAs barrier and spacer. These collectively act to impede current flow to and from the channel from the ohmic metal contact and hence define the specific resistance through the contact. From **Section 2.2**, the processes by which carrier transport occur across potential barriers were discussed with reference to the Schottky contact. These concepts similarly hold for the transmission of carriers through a more complex potential structure as that for the ohmic contact, which is merely a collection of sequential potential barriers and can therefore be treated as such. At room temperature the resistance associated with a specific barrier will be related to the height and width of the barrier which will in turn

define whether thermionic emission or field emission is the dominant process of transport. To maximise vertical conduction through the structure, and hence minimise the contact resistance, the resultant potential barriers can be reduced though various means. These include larger doping concentrations to reduce barrier widths or the use of narrower band gap materials to reduce conduction band offsets and hence barrier heights. However, these have to be carefully balanced with the ability to form an effective Schottky gate contact to minimise gate leakage and parallel conduction. The choice of materials is also limited to that defined by the lattice parameter of the layer structure with a margin for acceptable strain as previously discussed. It is standard practise to anneal the ohmic contact once deposited onto the semiconductor in an attempt to encourage diffusion of the ohmic metals through the layer structure [2.15]. Annealing acts to introduce metal through the ohmic region effectively highly doping the semiconductor in this area and creating a hybrid metal/semiconductor junction between the channel and external ohmic contact metal. The process of ohmic contact annealing and related issues is discussed further under fabrication in **Section 3.5.**

2.4 HEMT D.C. characteristics

Source - Drain current-voltage response

By applying a potential difference between the source and drain of a HEMT device, current will flow between them via the ohmic contacts and through the device channel beneath the gate contact. At low voltages and hence at low electric fields, the current - voltage response between source and drain (I_d - V_{ds}) will be ohmic, i.e. $I_d \alpha V_{ds}$. (For typical HEMT operation, the source is earthed and a positive voltage is applied to the drain). As the field is increased, velocity saturation effects within the semiconductor cause the current to saturate at a particular source - drain voltage V_n, often referred to as the "knee" voltage. This saturation results in a finite output conductance of the completed device. With zero voltage applied to the gate, I_d will remain in this saturated state with increased V_{ds}, until at a critical voltage, known as the breakdown voltage V_b, the electric field through the channel will be sufficiently large to incur impact ionisation [2.16]. Under these conditions, free electrons are supplied with sufficient energy by the applied field that upon collision with an atom an electron can be knocked free from the atomic coulomb forces, creating an additional free electron and hole. The resultant sudden increase in carrier concentration within the channel leads to a sharp increase in the I_d - V_{ds} response and the device is said to 'breakdown'. These three stages are demonstrated in **Figure 2.4.1**, which shows a typical HEMT I_d - V_{ds} response with zero gate voltage.

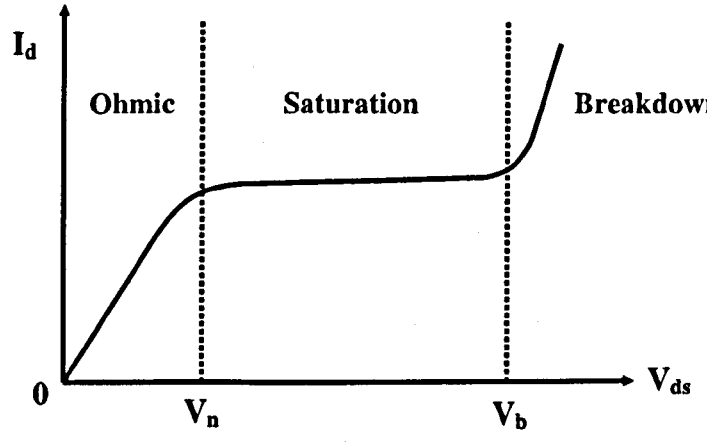


Figure 2.4.1 - Typical HEMT output characteristics with zero gate bias

Values for V_n and V_b will vary depending on the geometry and material composition of the device in question. Varying device geometry will lead to varied electric field strengths throughout the device, while differing material contents will exhibit different carrier velocities and breakdown characteristics. As a general rule the breakdown voltage will be proportional to the bandgap of the material, as the energy required by a colliding electron to free an electron-hole pair will be at least that of the bandgap energy [2.16]. In other words materials with smaller band gap energies will experience breakdown at lower electric field strengths than those with larger bandgap energies.

Source - Drain current modulation

As discussed in previous sections, the formation of the Schottky gate contact above the device channel provides a method of depleting the carrier density beneath the contact with the application of a negative voltage. This process allows the modulation of current flowing from source to drain by depleting the channel region below the contact and effectively increasing the channel resistance. The gate voltage is decreased (more negative) until at a particular voltage the channel region below the gate is completely depleted, resulting in zero current flow between source and drain. In this state the device is said to be 'pinched off', as the gate voltage has reached the pinch off voltage $-V_p$.

With the introduction of source-drain current modulation by the application of a negative gate bias, the I_d - V_{ds} characteristics presented in **Figure 2.4.1** are reconsidered under varied gate bias conditions:

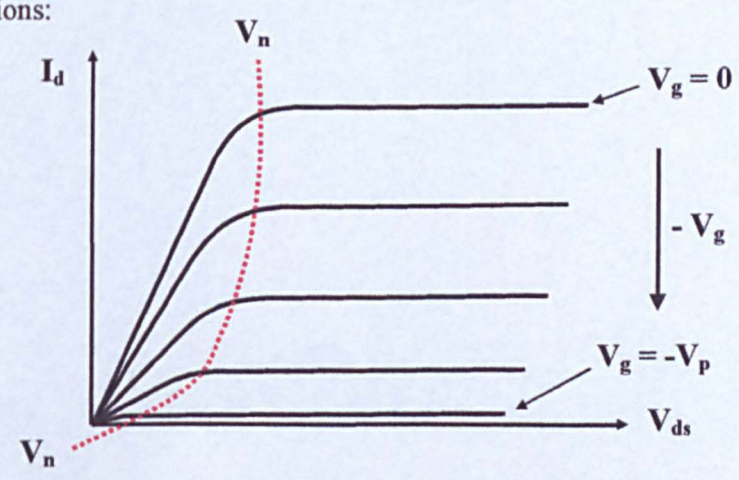


Figure 2.4.2 - Typical HEMT I_d-V_{ds} characteristics with varied gate bias ($V_g = 0$ to $-V_p$)

The I_d - V_{ds} characteristics are considered for stepped gate voltages between 0V and pinch off at $-V_p$, and below the breakdown voltage between the source and drain for each step. As the gate voltage is brought down to $-V_p$, the saturation current is reduced until virtually zero. It is also interesting to note the shift in the knee voltage V_n found with larger gate bias. As V_g approaches $-V_p$, V_n is reduced considerably. This results from a region of high electric field formed to the drain side of the gate, arising from the large potential difference between gate and drain. This process is illustrated in **Figure 2.4.3**.

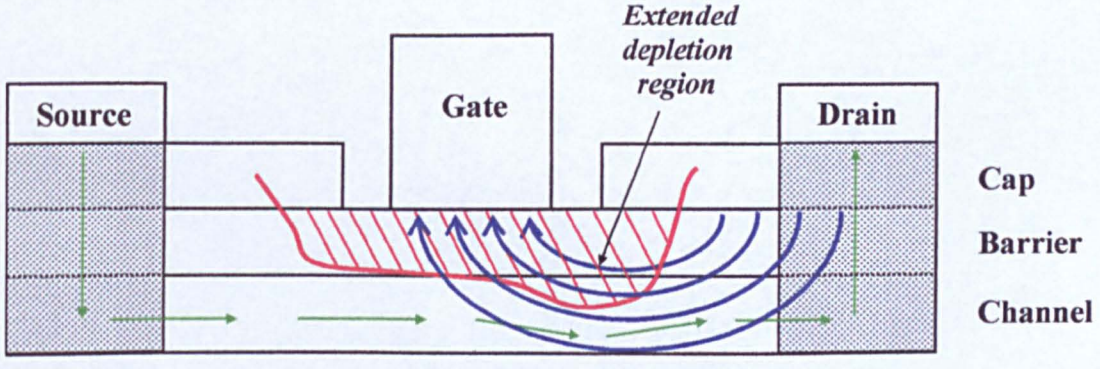


Figure 2.4.3 - Gate/Drain electric field distribution and extended depletion region for a gate and drain biased HEMT

The high electric field concentration between drain and gate regions is illustrated in **Figure 2.4.3** by the blue arrows. The increased potential difference applied between gate and drain extends the effective depletion region on the drain side of the gate as demonstrated by the red region in the figure. This in turn leads to a non-uniform carrier concentration through the channel between source and drain. Current flowing from source to drain will experience this concentrated electric field, leading to large carrier velocities and if the field is strong enough, breakdown in this area. At larger gate biases, breakdown will therefore occur at lower source-drain voltages due to the increased electric field strength supplied by the gate bias.

Intrinsic and Extrinsic models

In reality, the characteristics and performance of a completed HEMT device will not only depend on the physical processes that occur within the gate region, but also on external

parasitics, such as access resistances, inductances and capacitances that arise due to the overall device geometry [2.17]. Often the "Intrinsic" device properties are referred to, in which solely the gate region of the device is discussed. Realistically, the overall device performance will differ greatly from this intrinsic model due to the existence of such parasitics [2.17]. Hence an "Extrinsic" model can be used to accurately represent a completed HEMT device in which external parasitic elements are accounted for.

D.C. figures of merit

An important figure of merit for HEMT devices is the D.C. transconductance g_m, which is defined as the rate of change of drain current I_d with applied gate voltage V_{gs} at a fixed source-drain voltage V_{ds} [2.18], i.e.

$$g_m = \left(\frac{dI_d}{dV_{gs}} \right)_{V_{ds}} \qquad \textbf{Eqn 2.4.1}$$

The drain current can be expressed as:

$$I_d = G v n q \qquad \textbf{Eqn 2.4.2}$$

Where G is the width of the device, v the effective velocity in the channel, n the 2D carrier sheet density within the channel and q the charge on the electron.

Comparing again the carrier concentration modulation by application of a gate voltage to a parallel plate capacitor, and assuming the relative permittivity is constant and complete carrier depletion through the layers between gate and channel, the voltage between the gate and channel V_{gs} can be expressed as:

$$V_{gs} = \frac{q h n}{\varepsilon} \qquad \textbf{Eqn 2.4.3}$$

Where h is the gate - channel separation and ε is the relative permittivity of the semiconductor between the gate and channel.

By differentiating **Eqns 2.4.2** and **2.4.3** with respect to the 2D carrier concentration, n, and substituting into **Eqn 2.4.1**, g_m can be expressed as:

$$g_m = \frac{\varepsilon G v}{h} \qquad \textbf{Eqn 2.4.4}$$

Hence the transconductance is directly proportional to the carrier velocity within the gate region and inversely proportional to the gate - channel separation.

The transconductance given in **Eqn 2.4.4** however is a figure based on an intrinsic model and assumes that the potential difference between the gate and channel will be that applied between the gate and source contact. In reality this is not the case due to the voltage drop associated with the parasitic source resistance [2.19].

If the total source resistance R_s is defined, then the real voltage between gate and channel V_{gs}^* can be expressed as:

$$V_{gs}^* = V_{gs} - I_d R_s \qquad \textbf{Eqn 2.4.5}$$

By expressing the source - gate voltage V_{gs} in terms of the gate - channel voltage V_{gs}^* and the series parasitic voltage drop $I_d R_s$, from **Eqn 2.4.5**, the extrinsic transconductance can be expressed as:

$$g_m = \frac{g_m^*}{1 + g_m^* R_s} \qquad \textbf{Eqn 2.4.6}$$

where g_m^* is the intrinsic transconductance .

Hence values for the intrinsic and extrinsic transconductances will differ depending on the magnitude of the parasitic source resistance.

In addition to the transconductance, the output conductance of the device g_c is defined [2.20]:

$$g_c = \left(\frac{dI_d}{dV_{ds}}\right)_{V_{gs}} \qquad \textbf{Eqn 2.4.7}$$

This is merely the inverse of the output resistance defined between source and drain, i.e. $1/R_{ds}$ at a fixed gate voltage V_{gs}.

Other important D.C. figures of merit are summarised:

Saturation Current I_{dss} - The magnitude of current flowing from source to drain in the saturated region with zero applied gate bias.

Pinch-off Voltage V_p - The voltage required on the gate to completely deplete the channel in the gate region and reduce source - drain current to zero.

Breakdown Voltage V_b - The voltage applied between source and drain at which the device begins to breakdown through impact ionisation.

2.5 HEMT frequency response and equivalent circuit

To better understand the nature of HEMT operation at high frequency, it is beneficial to be able to identify the individual electromagnetic elements that comprise the device, and devise an equivalent circuit as a combination of these elements. An equivalent circuit also allows specific differentiation between intrinsic and extrinsic models and associated parasitic elements. First the basic HEMT structure is re-examined and equivalent circuit elements superimposed over the standard structure given in **Figure 2.5.1**.

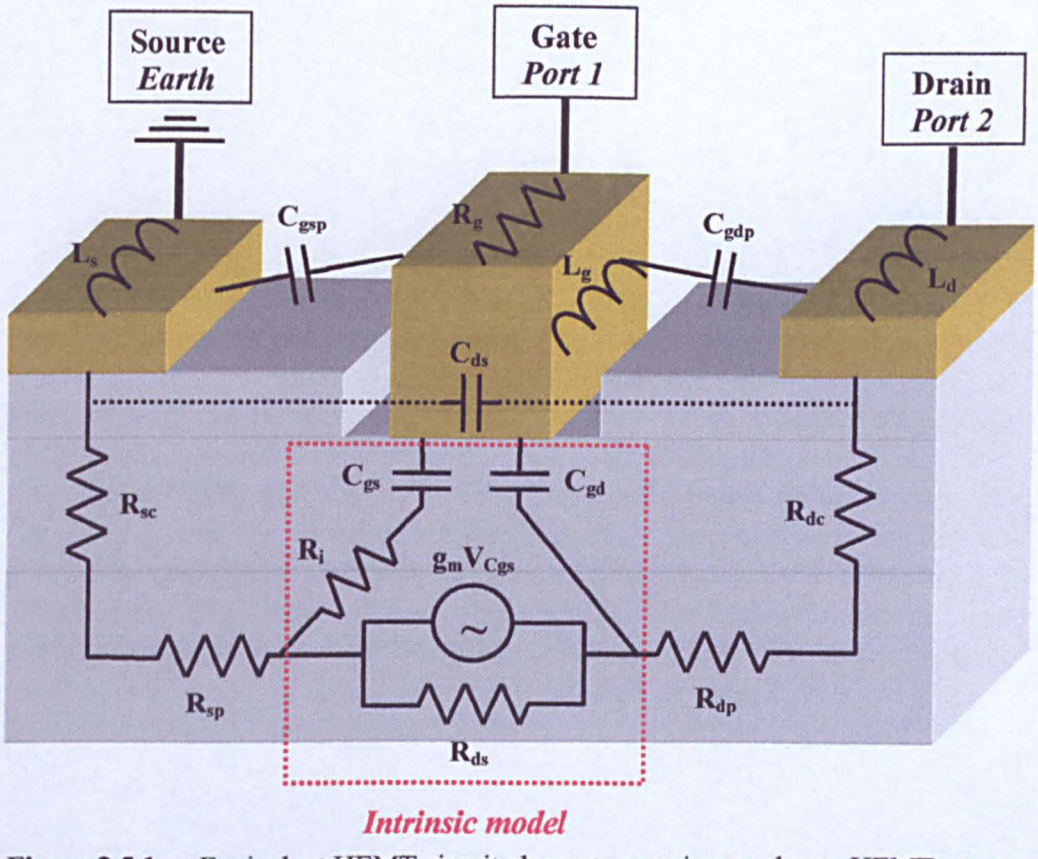


Figure 2.5.1 - Equivalent HEMT circuit elements superimposed onto HEMT structure

The structure given in **Figure 2.5.1** is a simplified version of that given previously in **Figure 2.3.1**, but is accurate enough for the means of identifying associated circuit elements. The intrinsic region of the device is marked by a red border and is effectively the region below the gate contact. The modulation of the source-drain current is represented by a current generator that produces a current equal to the intrinsic transconductance multiplied by the voltage across the capacitor C_{gs}, i.e. $g_m V_{Cgs}$. This is placed in parallel with the resistor R_{ds} which represents the finite source - drain resistance of the device. The total gate capacitance

is divided into two contributing elements, C_{gs} and C_{gd}. For reasons discussed previously, the effective depletion region below the gate will be non-uniform moving from source to drain under an applied bias, and hence the gate capacitance will vary with distance along the channel. The capacitance C_{gs} therefore represents that across the depleted region to the source side of the gate while C_{gd} represents the wider depleted region towards the drain side. The resistor R_i in series with C_{gs} accounts for resistance associated through the distributed capacitance of C_{gs}. Beyond the intrinsic model, the parasitic source and drain resistances are defined as the contact resistances R_{sc} and R_{dc} which represent the resistance of the contact through the ohmic region to the device channel, plus the parallel access resistances R_{sp} and R_{dp} which arise from conductance from the ohmic contact regions to the gate region through the device channel. Parasitic capacitance C_{ds} arises between the highly conducting ohmic contact regions in addition to capacitances C_{gsp} and C_{gdp} between the gate, source and drain metallisations. Inductance will also arise in all three terminals indicated by L_g, L_s and L_d. Finally, the gate resistance R_g represents resistance along the width of the gate i.e. perpendicular but in the plane of current flow from source to drain.

Combining each of the aforementioned elements provides an equivalent circuit as given in **Figure 2.5.2**.

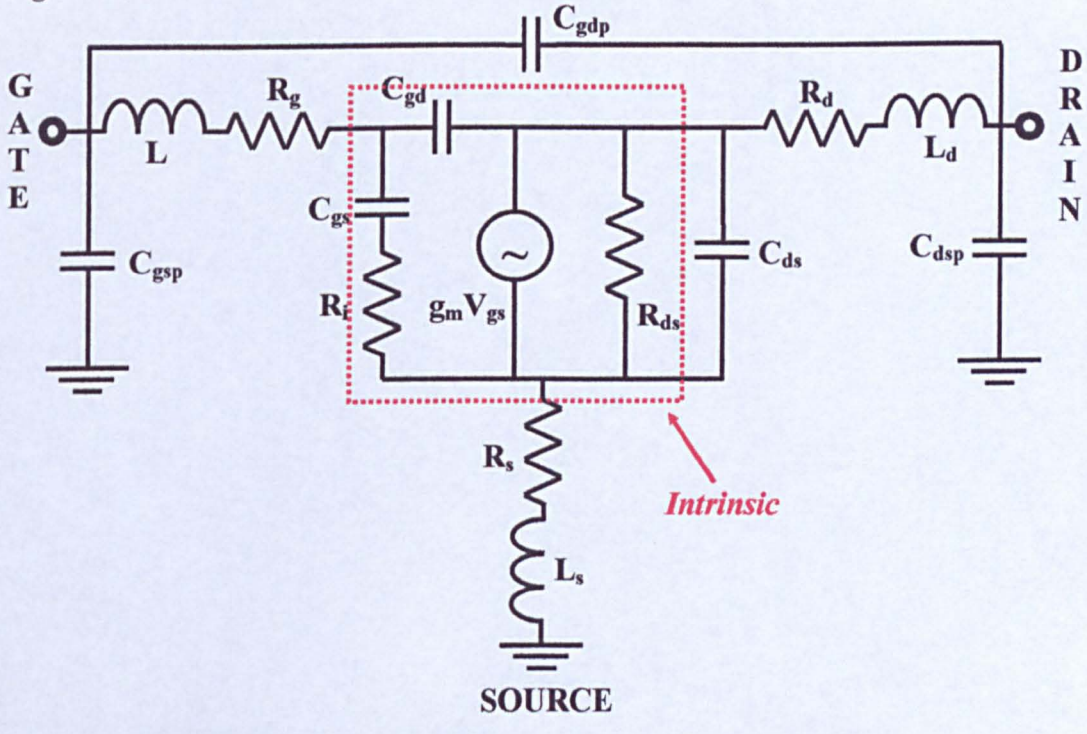


Figure 2.5.2 - Equivalent HEMT intrinsic and extrinsic circuits

The intrinsic and extrinsic equivalent circuits presented in **Figure 2.5.2** provide insight into the nature of high frequency (RF) HEMT operation and the factors that effect such operation. To allow RF characterisation, the HEMT is treated as a two port network, with the gate at port 1 and the drain at port 2. The source is earthed. An RF signal is injected into the gate at port 1 and through modulation of the source-drain current with external source drain bias, an amplified RF signal emerges from the drain at port 2.

An important figure of merit for high frequency operation known as the *transition* or *cut-off frequency* f_T [2.20], is defined as the frequency at which the magnitude of current entering the device via the gate is equal to that leaving the device via the drain, i.e. $I_g = I_d$, (or the current flowing into port 1 equals that flowing from port 2 when the output is short circuited). An intrinsic f_T is readily defined by considering the gate contact simply as the two parallel capacitances C_{gs} and C_{gd} separating the gate and channel as shown in **Figure 2.5.3**.

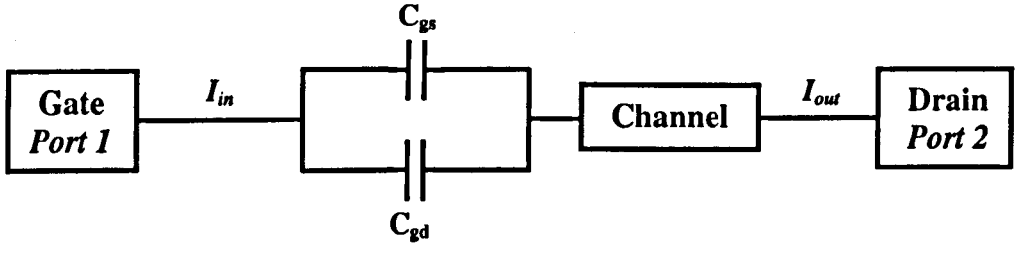


Figure 2.5.3 - Simple gate capacitance RF HEMT model

The magnitude of the current I_{in} flowing through the parallel capacitance network is defined as:

$$|I_{in}| = V_{Cg}\, 2\pi f\, (C_{gs} + C_{gd}) \qquad \textbf{Eqn 2.5.1}$$

Where V_{Cg} is the voltage across the parallel capacitance ($C_{gs} + C_{gd}$) and f is the frequency of the input signal. The current flowing through the drain I_{out} when the output is short circuited can be expressed as the product of the transconductance g_m and the voltage between the gate and channel V_{Cg}, i.e.

$$I_{out} = g_m V_{Cg} \qquad \textbf{Eqn 2.5.2}$$

When $I_{in} = I_{out}$, f_T is defined:

$$\text{Intrinsic } f_T = \frac{g_m}{2\pi\left(C_{gs} + C_{gd}\right)} \qquad \textbf{Eqn 2.5.3}$$

Hence to obtain a high intrinsic f_T, the total gate capacitance should be reduced and/or the transconductance increased.

Treating the gate - channel capacitance once again as a parallel capacitor, the gate capacitance C_g can be expressed as:

$$C_g = \frac{L G \varepsilon}{h} \qquad \textbf{Eqn 2.5.4}$$

Where L is the length of the gate region, (in the direction of source-drain current flow), G and h are the width of the device and channel gate separation respectively.

Substituting **Eqn 2.5.4** for the total gate capacitance and the expression for the transconductance given in **Eqn 2.2.4** into **Eqn 2.5.3**, the intrinsic f_T is:

$$f_T = \frac{v}{2\pi L} \qquad \textbf{Eqn 2.5.5}$$

Even though this simplistic model assumes a constant capacitance per unit length for the gate contact, it still illustrates the increase associated with f_T by maximising the velocity of carriers through the gate region and/or the reduction of the effective length of the gate region.

The second figure of merit for high frequency operation is the *maximum frequency* of oscillation f_{max} [2.20]. Unlike the transition frequency f_T which concerns current amplification through the device, f_{max} describes the frequency at which the device exhibits unity power gain, i.e. when the product of the voltage and current gains of the device equal one.

To be able to define the power amplification of the device, the output is set to drive a load of resistance R_L. A simple equivalent circuit can then be devised to allow the definition of voltage and current amplification through the device:

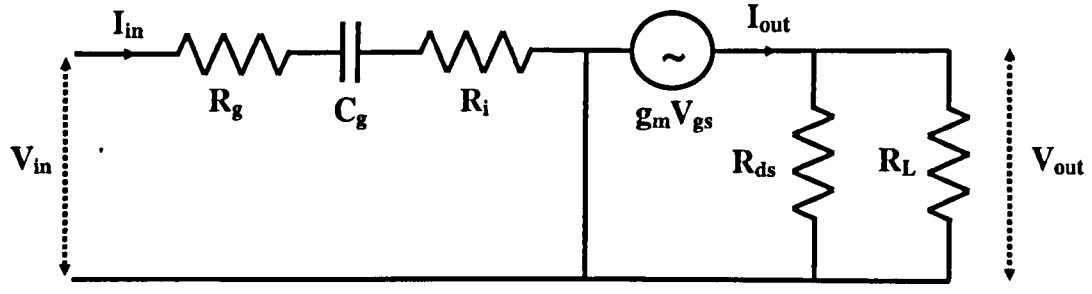


Figure 2.5.4 - Intrinsic Power amplification HEMT model with output load

where V_{gs} is the voltage between the gate and source.

From **Figure 2.5.4**, the voltage gain is defined as the ratio of output and input voltages, i.e. V_{out}/V_{in}.

For the moment, the total resistance of the parallel resistors R_{ds} and R_L is defined as R_o. Then:

$$\left|\frac{V_{out}}{V_{in}}\right| = \frac{g_m R_o}{\left(1+4\pi^2 f^2 C_g^{\,2}\left(R_g+R_i\right)^2\right)^{1/2}} \qquad \textbf{Eqn 2.5.6}$$

Assuming then the denominator of **Eqn 2.5.6** to be >> 1, the voltage gain expression simplifies to:

$$\left|\frac{V_{out}}{V_{in}}\right| \cong \frac{g_m R_o}{\left(2\pi f C_g\left(R_g+R_i\right)\right)} \qquad \textbf{Eqn 2.5.7}$$

Previously the current gain has been defined:

$$\left|\frac{I_{out}}{I_{in}}\right| = \frac{g_m}{2\pi f C_g} \qquad \textbf{Eqn 2.5.8}$$

and so the power gain G_p is defined:

$$G_p = \left|\frac{V_{out}}{V_{in}}\right| x \left|\frac{I_{out}}{I_{in}}\right| = \frac{g_m^2 R_o}{4\pi^2 f^2 C_g^2 (R_g + R_i)} \qquad \textbf{Eqn 2.5.9}$$

Substituting the expression for intrinsic f_T from **Eqn 2.5.3** into **Eqn 2.5.9** yields:

$$G_p = \left(\frac{f_T}{f}\right)^2 \frac{R_o}{R_g + R_i} \qquad \textbf{Eqn 2.5.10}$$

Looking again at **Figure 2.5.4**, for maximum power gain to occur across the load resistance, the output resistance must equal the load resistance, i.e. $R_{ds} = R_L$. Therefore the power dissipated in the load R_L will be one quarter that across the two parallel resistors R_{ds} and R_L, and **Eqn 2.5.10** becomes:

$$G_p = \left(\frac{f_T}{f}\right)^2 \frac{R_{ds}}{4(R_g + R_i)} \qquad \textbf{Eqn 2.5.11}$$

Finally f_{max} is defined when the power gain reaches unity:

$$\text{Intrinsic } f_{max} = \frac{f_T}{2}\left(\frac{R_{ds}}{R_g + R_i}\right)^{1/2} \qquad \textbf{Eqn 2.5.12}$$

Extrinsic frequency response

Up until this point the expressions derived for the transition frequency f_T and maximum frequency f_{max} have been largely based upon an intrinsic HEMT model and therefore even though they demonstrate 1st order trends in device performance with intrinsic variables, they do not tell the whole story. More detailed expressions can be derived for both by considering the parasitic resistances given in the extrinsic model from **Figure 2.5.2.**

An additional equivalent circuit can then be constructed highlighting these parasitic resistances and their effect on the frequency performance of the device:

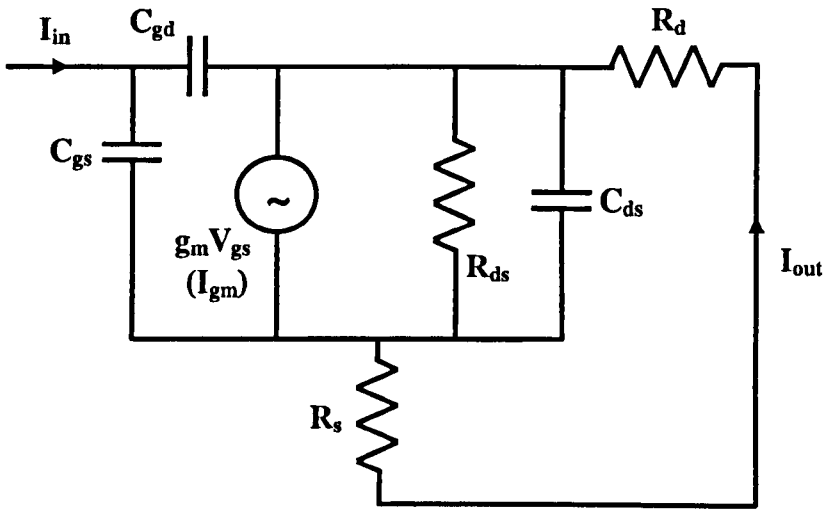


Figure 2.5.5 - Parasitic resistance HEMT model

To be able to define the current gain and hence the f_T of a FET type device, the drain is effectively shorted to the source to allow current to pass through the device. For the intrinsic expression for f_T, it is assumed that the total current produced by the current generator I_{gm} will pass through the drain. However from **Figure 2.5.5**, a parallel resistor network is set-up in the loop from the gate region in which current can flow through either the source and drain resistances R_s and R_d, or through the effective resistance R_{ds}. This resistive divider reduces the output current flowing through the drain I_{out} by a factor $(1 + (R_s + R_d) / R_{ds})$, i.e.

$$I_{out} = \frac{I_{gm}}{\left(1 + \frac{R_s + R_d}{R_{ds}}\right)} \qquad \textbf{Eqn 2.5.13}$$

In addition, the voltage across C_{gd} is greater than across C_{gs} due to the voltage developed across resistors R_s and R_d. The resulting Miller effect describes an increase in the effective capacitance C_{gd}` defined by the voltage amplification across the gate capacitance, i.e.

$$C_{gd}` = C_{gd}(1 + G_V) \qquad \textbf{Eqn 2.5.14}$$

where C_{gd} is the real gate-drain capacitance and G_V is the voltage gain.

From **Figure 2.5.5**, G_V can be defined as:

$$G_V = \frac{I_{out}\left(R_s + R_d\right)}{V_{gs}} \qquad \textbf{Eqn 2.5.15}$$

By combining **Eqn 2.5.13** and **2.5.15** into **2.5.14**, the effective capacitance C_{gd}` can be expressed as:

$$C_{gd}` = C_{gd}\left[1 + g_m \frac{R_s + R_d}{\left(1 + \frac{R_s + R_d}{R_{ds}}\right)}\right] \qquad \textbf{Eqn 2.5.16}$$

Finally using **Eqn 2.5.13** for the output current and **Eqn 2.5.16** for the effective C_{gd}, the intrinsic expression defined for f_T in **Eqn 2.5.3** is re-expressed taking into consideration the effects of the source and drain parasitic resistances as:

$$\text{Extrinic } f_T = \frac{g_m / 2\pi}{\left(C_{gs} + C_{gd}\right)\left(1 + \frac{R_s + R_d}{R_{ds}}\right) + C_{gd}\, g_m \left(R_s + R_d\right)} \qquad \textbf{Eqn 2.5.17}$$

Similar analysis yields an extrinsic f_{max} expression [2.21]:

$$\text{Extrinic } f_{max} = \frac{f_T}{2\left(\frac{R_g + R_i + R_s}{R_{ds}} + 2\pi f_T R_g C_{gd}\right)^{1/2}} \qquad \textbf{Eqn 2.5.18}$$

Eqns 2.5.17 and **2.5.18** reflect the dependence of frequency figures of merit f_T and f_{max} on associated parasitic resistance elements, portraying a more realistic model than that described by the intrinsic expressions.

The magnitude in the reduction of the frequency performance of a device by parasitic resistances will depend on the bias conditions of the device as from **Eqns 2.5.17** and **2.5.18**, g_m, C_{gd}, C_{gs} and R_{ds} are present which are all strong functions of the applied bias. In addition, for shorter gate length devices the gate capacitance is reduced, which will further accentuate the role of the parasitic resistances in reducing the frequency performance from that of the intrinsic device.

2.6 Material issues and design

Discussion of the nature of HEMT operation and performance has begun to shed light onto some of the major issues that affect and restrict overall HEMT device performance. The main factors that influence such performance are the effective gate length of the device, but also the intrinsic material and transport properties through the HEMT structure and in particular through the device channel. For modern III-V based devices, the ternary compound Indium Gallium Arsenide (InGaAs) is favoured as the device channel material for several reasons [2.22]. As indium is introduced to GaAs, the bandgap energy is reduced which provides larger conduction band offsets through the HEMT structure (or creating deeper potential wells) that better confine electrons [2.23].

For high frequency performance, higher indium content InGaAs exhibits higher peak carrier velocities and hence a larger average electron velocity through the gate region than for lower indium concentrations [2.23]. Increasing carrier velocity through the gate region provides a method of increasing device frequency response as demonstrated by **Eqns 2.5.5** and **2.5.12**.

Semiconductor carrier velocity - electric field characteristics are found to be dependent upon their energy bands as well as material doping profiles. The curvature or steepness of curve of the energy band for a material defines the effective mass for carriers at a particular energy and wave vector [2.24]. This in turn directly affects the carrier mobility which is inversely proportional to the effective mass [2.25]. As the mobility describes carrier velocity with respect to the applied electric field i.e. equals the ratio of the average carrier velocity to the field, it becomes a figure of significant interest.

Looking specifically at the energy bands for GaAs / InGaAs compounds, both are divided into Γ, X and L valleys with the low energy Γ valley centrally located where the wave vector equals zero, and the other two at higher energies at positive and negative wave vectors. Carriers within the Γ valley will exhibit a smaller effective mass and hence a higher mobility than those occupying the higher energy X and L valleys [2.26]. The overall carrier transport

characteristics through the material will therefore depend upon the distribution of carriers across these valleys. For electrons under the influence of an external electric field within an un-doped material, as the field is increased from zero, the electron velocity will steadily increase as defined by the mobility associated with the Γ valley. Once the applied field reaches a particular strength E_p, the velocity will peak, and then begin to decrease with increased field. Finally the electron velocity saturates at a field strength E_s. This process is demonstrated in **Figure 2.6.1**.

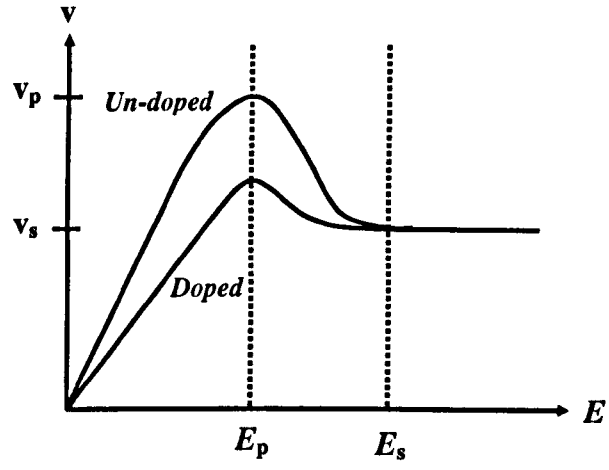


Figure 2.6.1 - Carrier velocity - electric field response for undoped III-V semiconductor

The mechanism by which the carrier velocity peaks and then decreases can be explained by considering the electron population distribution throughout the material energy band structure. Initially at low field strengths (i.e. below E_p) the majority of electrons will reside within the central Γ valley, and hence the average carrier mobility will be the peak value for the material. As the field increases, the average kinetic energy of the electrons increases until at the peak field E_p they begin to possess sufficient energy to transfer to the L valley located at an energy $E_{\Gamma L}$ above the Γ valley. The mobility associated with the L valley is considerably less due an increased effective mass [2.26] and hence electrons will exhibit a reduced velocity. As the electric field is further increased beyond E_p, the average electron velocity therefore decreases as more electrons relocate to the L valley and a negative differential mobility is observed.

As the field is yet further increased beyond E_s, the majority of carriers will occupy the higher L and X valleys in which their transport properties are dominated by intervalley scattering processes and hence the average carrier velocity saturates [2.27].

For transport through a material that is doped, the presence of ionised dopant atoms leads to coulombic scattering of the carriers, randomising their trajectories and reducing the low field mobility. This leads to a reduced peak velocity compared to that observed with the undoped material. However, the field strength at which the migration of electrons from the Γ to the L valleys occurs remains unaltered to that for the undoped material due to the scattering process conserving the total electron energy. Hence the peak velocity albeit reduced is observed at a similar electric field strength in both the doped and undoped material.

As the indium content of the compound InGaAs in increased, the Γ - L valley energy separation $E_{\Gamma L}$ becomes larger [2.28] as well as the effective mass being reduced. As a result electrons must possess a greater energy before transferring from the Γ to the L valley and therefore reach higher velocities through accelerating for longer time periods within the Γ valley while experiencing a higher mobility defined by the low effective mass. As the valley transfer energy increases with increased indium, so does the electric field strength at which the peak velocity within the material is observed.

These large peak velocities combined with other attractive attributes mentioned previously make high indium content InGaAs based devices highly desirable for high frequency and low noise operation. For power-orientated devices however, a higher indium concentration is less desirable due to lower breakdown voltages resulting from a narrower bandgap. A trade-off therefore exists in III-V HEMT technology between RF and power performance.

Material limitations

Ideally, as higher indium content InGaAs provides higher carrier peak velocities as well as better carrier confinement and lower noise figures, using InAs as the device channel material would provide the ultimate device performance with respect to the device material system.

As discussed in **Section 2.1** however, limitations exist in the growth of such structures due to the lattice constant of InGaAs varying considerably with indium content. GaAs / AlGaAs based HEMT devices are readily grown upon GaAs substrates as the lattice constant for AlGaAs remains similar for all aluminium compositions. Introducing indium into GaAs increases the lattice constant significantly however, placing a limit on the indium concentration that can be used on a GaAs substrate. By allowing an element of strain into the structure, some freedom is provided into the choice of the indium content, without creating extremely brittle material or thread dislocations through the structure. Pseudomorphic HEMT (pHEMT) devices adopt this technique where using a GaAs substrate, an InGaAs / AlGaAs structure is grown in which the indium content typically can range from 20 - 25% [2.29]. The InGaAs channel in such device is tensile strained, and a limit is placed on the thickness of such a strained layer given from the magnitude of the strain predicted by the Mathews and Blakeslee model [2.30]. Also when a semiconducting material is placed under either compressive or tensile strain, its band structure becomes distorted from that of the un-strained state. This mainly affects the valence band of the material but is also responsible for altering the conduction band discontinuity between strained adjacent layers [2.31]. In addition to GaAs pHEMT technology, an $In_{0.53}Ga_{0.47}As$ / $In_{0.52}Al_{0.48}As$ based HEMT structure that is lattice matched to an InP substrate provides a method for producing a higher indium content device. $In_{0.53}Ga_{0.47}As$ lattice matches to both $In_{0.52}Al_{0.48}As$ and InP and hence the structure is relatively strain free [2.32]. Beyond this, to achieve higher indium concentrations, a pHEMT InP based system can be used similar to the GaAs pHEMT system where the latticed matched 53% indium InGaAs has been increased up to 80% placing the channel again under tensile strain [2.33]. A technology parallel to these exists which provides the ability to grow a wide range of indium concentrations upon a GaAs substrate. By growing a thick indium graded InAlAs buffer layer between a GaAs substrate and device channel, i.e. $In_xAl_{1-x}As$ from $x = 0$ to that required for lattice matching with the channel, the strain produced by the difference in lattice constants of the substrate and channel is relaxed through the buffer. Material grown using this technique is known as *Metamorphic* due to the

graded nature of the structure [2.34]. Although metamorphic devices provide a route to producing varied indium content devices, the performance of a metamorphic device will generally be less than that for a similar lattice matched device due to imperfect relaxation through the graded buffer leading to an element of non-uniform strain in the channel. [2.34]

HEMT transport properties

The transport properties examined for bulk semiconductors provide useful insight into the transport processes that occur in a real HEMT device. However, the heterostructure that forms a HEMT complicates such transport from that in the bulk, isolated semiconductor. As previously discussed, the complex HEMT heterostructure forms a potential well within the material channel that acts to confine electrons to a two dimensional plane. This 2DEG is then spatially separated from a layer of delta doping, whose ionised cores act to scatter 2DEG carriers through coulombic interaction. The magnitude of this scattering depends on the physical separation between the doping and channel and will affect the low field mobility and hence velocity characteristics through the channel. The channel material itself is also confined by undoped material with a larger bandgap energy i.e. the spacer and buffer layers, which define the finite height of the potential well that forms the channel. It is assumed that electrons travelling through the gate region of the device will be completely confined within the channel quantum well. However as they acquire a greater energy from the applied electric field, some will scatter into the higher band gap confining material through a process known as *real space transfer* [2.35]. These "Hot" or high-energy electrons still contribute to the overall carrier transport through the device but exhibit the reduced mobility and velocity associated with the larger band gap material. Often with short gate length HEMT systems this process manifests itself as a kink in the output characteristics of the device where the drain current prematurely saturates or in the extreme decreases with increased source drain bias. This process is illustrated in **Figure 2.6.2.**

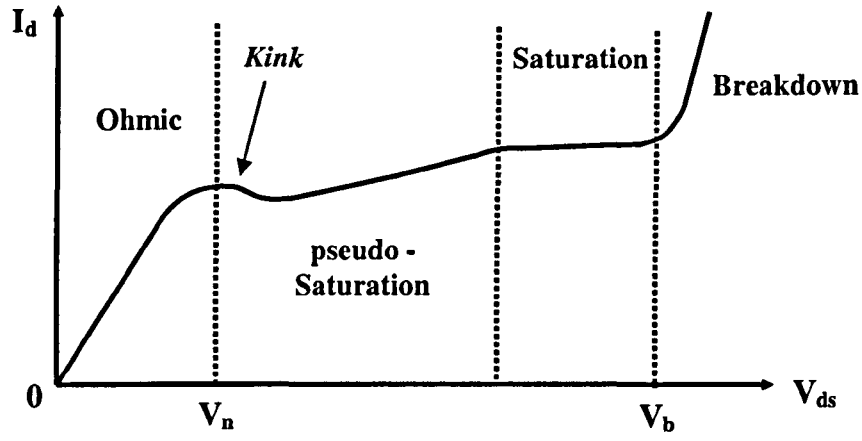


Figure 2.6.2 - I_d-V_d HEMT characteristics exhibiting the kink effect through real space transfer

Figure 2.6.2, similar to **Figure 2.4.1** displays the output characteristics of a HEMT device, but in this instance the ideal I_d-V_d response is distorted by real space transfer of electrons from the channel to surrounding layers, reducing the drain current across a small increased voltage range and creating a *kink* in the profile. Beyond the kink at higher bias, the I_d-V_d response is steeper than that for the true saturation region caused by parasitic conduction through these adjacent layers although the channel current has saturated, creating a pseudo-saturation region. Depending on the structure in question, at a large enough bias conduction through the channel and parallel layers might both saturate before impact ionisation leads to breakdown through the structure. With the application of a gate bias this process can become more pronounced within the high field concentration region formed towards the drain side of the gate due to the higher field strength providing more electrons with sufficient energy to transfer to the surrounding layers. The process by which conduction occurs through the surrounding layers can also become complex depending on the material in question. As previously discussed, the compounds AlGaAs and InAlAs are used mainly in III-V HEMT technology for the buffer and barrier layers of the structure and hence the transport properties of these materials becomes significant where real space transfer is known to occur. It is proposed that deep level 'traps' can exist in both materials depending on their growth conditions [2.36], in which electrons having attained sufficient energy to transfer from the channel, become trapped and no longer contribute to the drain current. As the bias and hence

field is increased, these trapped electrons are eventually supplied with sufficient energy to escape the trap and re-contribute to the current flow. This mechanism of trapping and releasing of electrons is time dependent as hence as the field is increased, less carriers are subject to the trapping mechanism and the current increases gradually until saturation or breakdown.

The process of real space transfer leading to the kink effect combined with intervalley transfer of electrons within the channel often leads to deviations in the ideal HEMT output characteristics portrayed earlier in **Figure 2.4.1**. Such deviations will depend strongly on the growth process for the material combined with the doping profile and the dimensions of the HEMT structure which define relative electric field strengths throughout the device.

Material Scaling

With the reduction of the gate length of HEMT devices, scaling of the material structure must also be implemented [2.37]. As shorter gates are incorporated, the magnitude of the depletion effect upon charge within the channel below the gate is minimised, and hence to maintain efficient control of the channel current by the gate, the gate contact to channel separation must be tailored accordingly. This process is illustrated in **Figure 2.6.3**.

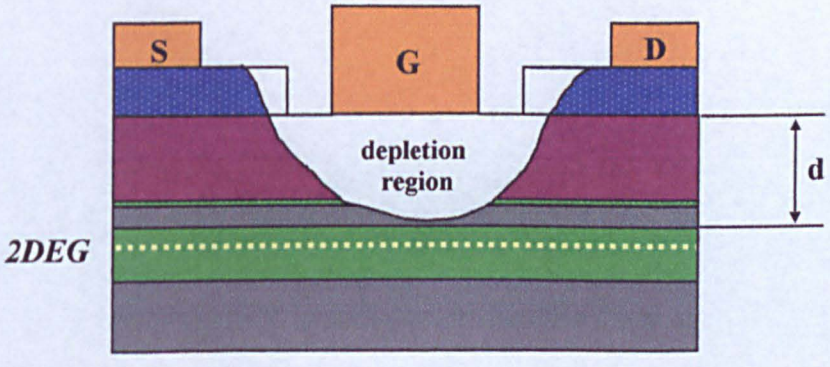


Figure 2.6.3 - Effective gate length and depletion region

The structure presented in **Figure 2.6.3** is similar to those given earlier showing the source, gate and drain contacts upon the material structure comprising of cap, barrier, spacer and channel layers. A layer of delta doping is situated between the barrier and spacer layers. The

gate is metallised upon the barrier layer and forms a Schottky contact the magnitude of which ensures depletion of the barrier layer of the structure with zero gate bias. This depletion effect is illustrated in **Figure 2.6.3**, and for this example is represented by an effective depletion region which is used to help visualise the two dimensional influence of the Schottky contact into the structure. The magnitude of this depletion effect must be tailored to ensure complete depletion of the layers above the channel, without infringing upon the channel itself to ensure maximum current through the channel with zero gate bias. This allows efficient modulation of the channel carrier density by extending the effect of this region into the channel with a gate bias. The dimensions of the region without bias will depend on the length of gate and size of the etch 'trench' that is formed to allow the gate to be metallised upon the barrier layer. The effective gate length as seen by electrons flowing through the channel will depend on the dimensions of the depletion region that imposes upon the channel. As the gate length is reduced to improve device performance, the size of the trench must also be reduced in order to maintain an effectively short gate length. The depth of the depletion region into the structure will also be reduced with shorter gate lengths and hence to maintain efficient control of the channel current, the gate to channel separation given by **d** in **Figure 2.6.3**, must be scaled accordingly.

It is proposed that the ratio of gate length to gate-channel separation be approximately 3:1 for efficient gate control [2.38]. However recent research also suggests that for ultra-short gate lengths, (sub 50-nm), it is more advantageous to reduce this ratio to ~ 2:1 to promote higher carrier velocities [2.39]. For such ultra-short gate lengths systems an unfortunate side effect is a reduction in the carrier concentration within the channel due to a smaller gate to channel separation [2.40]. This can lead to reduced device performance through increased sheet resistance, resulting in higher extrinsic resistances and lower transconductance figures. In addition, it has been suggested that as the gate length is reduced considerably, the ability to efficiently modulate the 2DEG concentration by the gate becomes questionable due to the depletion process becoming less dependent on the length of gate. The voltage range with which the channel current is modulated then becomes too large to be practical for device

application [2.41]. Such 'short channel' effects are speculated to begin to saturate the performance of short gate length devices, though recent results have demonstrated excellent performance with properly scaled systems [2.42].

Velocity Overshoot

An interesting effect that is argued to occur in shorter gate length devices, which further increases device performance is known as V*elocity Overshoot* [2.43]. Due to finite time periods associated with intervalley transfer of electrons within III-V materials, upon the onset of a sudden strong electric field, the average carrier velocity is found to peak at a greatly increased value for a short time interval before settling down to the lower saturated value associated with the strength of the applied field. If the time period at which carriers within the gate region of a HEMT device are subject to a strong field is equal to or less than that associated with this overshoot duration, carrier velocity will increase substantially over that for the static field. For shorter gate length devices, the time duration for which electrons travel through the gate region is reduced, (considering also the average electron velocity when entering the gate region which should be independent of gate length), and hence velocity overshoot is more likely to occur in these reduced gate length systems [2.44]. In addition, the velocity overshoot mechanism should increase for larger $E_{\Gamma L}$ and lower Γ valley effective mass materials, therefore also favouring a higher indium content system [2.44].

2.7 · Chapter Summary

Throughout this chapter the physics behind the operation of HEMT type devices has been discussed in an attempt to highlight the various elements that influence the completed device performance. Such performance is found to depend largely upon the transport properties through the device material, but also upon the specific device geometry. For a properly scaled device, the RF performance will increase as the effective gate length is reduced. The role of parasitic elements such as external resistances again inherent to the device geometry will also influence the performance of shorter gate length devices. With these issues in mind the process of HEMT device fabrication is considered in the following chapter along with the challenges associated with the realisation of short gate length devices.

References

[2.1] P. H. Ladbrooke, *MMIC Design GaAs FETs and HEMTs*, 1989, Artech House Inc. p. 189 - 190.

[2.2] H. P. Zappe, *Introduction to semiconductor integrated optics*. 1995: Artech House Inc.

[2.3] S. M. Sze, *High-Speed Semiconductor Devices*, 1990, John Wiley and Sons Inc. p. 24 - 33.

[2.4] S. M. Sze, *Physics of Semiconductor Devices*, 1981, John Wiley and Sons Inc. Chap. 1.4.

[2.5] S. M. Sze, *Physics of Semiconductor Devices*, 1981, John Wiley and Sons Inc. p. 19.

[2.6] S. Tiwari, *Compound Semiconductor Device Physics.* 1992, Academic Press Inc. p. 92 - 94.

[2.7] S. Tiwari, *Compound Semiconductor Device Physics.* 1992, Academic Press Inc. p. 417.

[2.8] S. M. Sze, *Physics of Semiconductor Devices*, 1981, John Wiley and Sons Inc. p. 273 - 276.

[2.9] R. T. Tung, *Schottky barrier height-do we really understand what we measure?* Journal of Vacuum Science & Technology, B Microelectronics Processing and Phenomena, 1993. 11(4): p. 1546-52.

[2.10] N. Newman et al., *On the Fermi level pinning behavior of metal/III-V semiconductor interfaces.* Journal of Vacuum Science & Technology, B Microelectronics Processing and Phenomena, 1986. 4(4): p. 931-8.

[2.11] S. Tiwari, *Compound Semiconductor Device Physics.* 1992, Academic Press Inc. p. 214 - 219.

[2.12] F. Podevin, O. Vanbesien, D. Lippens, *Quantum calculations of conduction properties of metal/InAlAs/InGaAs heterostructures.* Journal of Applied Physics, 2001. 89(11): p. 6247-52.

[2.13] M. Jang et al, *Simulation of Schottky barrier tunnel transistor using simple boundary condition.* Applied Physics Letters, 2003. 82(16): p. 2718-20.

[2.14] K. Kajiyama, Y. Mizushima, S. Sakata, *Schottky barrier height of n-InxGa1–xAs diodes.* Applied Physics Letters, 1973. 23(8): p. 458-459.

[2.15] S. M. Sze, *Physics of Semiconductor Devices*, 1981, John Wiley and Sons Inc. Chap. 5.7.

[2.16] J. Singh, *The tailoring of impact ionization phenomenon using pseudomorphic structures-applications to InGaAlAs on GaAs and InP substrates.* Semiconductor Science and Technology, 1992. 7(3B): p. 509-11.

[2.17] P.H. Ladbrooke, *MMIC Design GaAs FETs and HEMTs*, 1989, Artech House Inc. p. 222 - 224.

[2.18] S. Tiwari, *Compound Semiconductor Device Physics*. 1992, Academic Press Inc. p. 506.

[2.19] P. H. Ladbrooke, *MMIC Design GaAs FETs and HEMTs*, 1989, Artech House Inc. p. 139 - 142.

[2.20] P. H. Ladbrooke, *MMIC Design GaAs FETs and HEMTs*, 1989, Artech House Inc. p. 138 - 155.

[2.21] S. M. Sze, *Physics of Semiconductor Devices*, 1981, John Wiley and Sons Inc. p. 342 - 343.

[2.22] J. K. Zahurak et al, *Electron transport in InGaAs/AlInAs heterostructures and its impact on transistor performance.* Journal of Applied Physics, 1994. 76(11): p. 7642-4.

[2.23] R. Ross, P. Lugli, *Pseudomophic HEMT Technology and Applications*, 1994, NATO ASI series. p. 11 - 19.

[2.24] S. Tiwari, *Compound Semiconductor Device Physics*. 1992, Academic Press Inc. p. 20.

[2.25] S. Tiwari, *Compound Semiconductor Device Physics*. 1992, Academic Press Inc. p. 59 - 60.

[2.26] W. T. Masselink, *Electron velocity in GaAs: bulk and selectively doped heterostructures.* Semiconductor Science and Technology, 1989. 4(7): p. 503-12.

[2.27] M. Lundstrom, *Fundamentals of carrier transport*. 2000, Cambridge University Press. Chap. 7.

[2.28] S. M. Sze, *High-Speed Semiconductor Devices*, 1990, John Wiley and Sons Inc. p. 320.

[2.29] R. Ross, P. Lugli, *Pseudomophic HEMT Technology and Applications*, 1994, NATO ASI series. p. 28.

[2.30] J. W. Matthews, A. E. Blakeslee, *Defects in epitaxial multilayers. I. Misfit dislocations.* Journal of Crystal Growth, 1974. 27(1): p. 118-25.

[2.31] J. H. Davies, *The physics of low-dimensional semiconductors*. 1998, Cambridge University Press. p. 96 - 100.

[2.32] P. M. N. Smith, K. Nichols, W. Kong, L. MtPleasant, D. Pritchards, R. Lender, J. Fisher, R. Actis, D. Dugas, D. Meharry, A. W. Swanson, *Advances in InP HEMT technology for high frequency applications, 2001 International Conference on Indium Phosphide and Related Materials 2001*

[2.33] L. D. B. Nguyen, A. S. Brown, M. A. Thompson, L. M. Jelloian, *50-nm self-aligned-gate pseudomorphic AlInAs/GaInAs high electron mobility transistors.* Electron Devices, IEEE Transactions on, 1992. 39(9): p. 2007-2014.

[2.34] K. Van der Zanden, M. Behet, G. Borghs, *Comparison of metamorphic InGaAs-InAlAs HEMT's on GaAs with InP based LM HEMT's. 1999 GaAs MANTECH Conference.* 1999.

[2.35] T. W. Masselink, *Real-space-transfer of electrons in InGaAs/InAlAs heterostructures.* Applied Physics Letters, 1995. 67(6): p. 801-3.

[2.36] A. S. M. Brown, U. K. Mishra, C. S Chou, C. E. Hooper, M. A. Melendes, M. Thompson, L. E. Larson, S. E. Rosenbaum, M. J. Delaney, *AlInAs-GaInAs HEMTs utilizing low-temperature AlInAs buffers grown by MBE.* Electron Device Letters, IEEE, 1989. 10(12): p. 565-567.

[2.37] M.Van.Hove *et al, Scaling behaviour of delta-doped AlGaAs/InGaAs high electron mobility transistors with gatelengths down to 60 nm and source-drain gaps down to 230 nm.* Journal of Vacuum Science & Technology, 1993. B 11(4): p. 1203.

[2.38] Thayne, I., *PhD Thesis*, Uni. of Glasgow. Chap 2.

[2.39] Y. E. Yamashita, A. Endoh, K. Shinohara, K. Hikosaka, T. Matsui, S. Hiyamizu, T. Mimura, *Pseudomorphic In/sub 0.52/Al/sub 0.48/As/In/sub 0.7/Ga/sub 0.3/As HEMTs with an ultrahigh f/sub T/ of 562 GHz.* Electron Device Letters, IEEE, 2002. 23(10): p. 573-575.

[2.40] K. Kalna, S .Roy, A. Asenov, K. Elgaid, I.Thayne, *Scaling of pseudomorphic high electron mobility transistors to decanano dimensions.* Solid State Electron, 2002. 46: p. 631-638.

[2.41] Y. K. Awano, M. Kosugi, K. Kosemura, T. Mimura, M. Abe, *Short-channel effects in subquarter-micrometer-gate HEMTs: simulation and experiment.* Electron Devices, IEEE Transactions on, 1989. 36(10): p. 2260-2266.

[2.42] T. B. Parenty, S. Bollaert J. Mateos, X. Wallart, A. Cappy, *Design and realization of sub 100 nm gate length HEMTs.* 2001 *International Conference on Indium Phosphide and Related Materials.*

[2.43] M. Lundstrom, *Fundamentals of carrier transport.* 2000, Cambridge University Press. p. 332 - 334.

[2.44] K. S. Yoon, G. B. Stringfellow, R .J. Huber, *Transient transport in bulk Ga/sub 0.47/In/sub 0.53/As and the two-dimensional electron gas in Ga/sub 0.47/In/sub 0.53/As/Al/sub 0.48/In/sub 0.52/As.* Journal of Applied Physics, 1988. 63(4): p. 1126-9.

Chapter 3

HEMT Fabrication

The processes by which HEMT devices are fabricated become particularly intricate, specifically for short-gate length devices in an attempt to increase their performance. The various techniques by which such devices are fabricated are reviewed throughout this chapter in a bid to provide an understanding into such general processes and their influence on device performance. This provides a setting for the discussion of the specific processes and techniques developed throughout this research as described in later chapters.

Initially electron beam lithography (ebeam) and photolithography processes are discussed including the benefits and restrictions of both with respect to HEMT fabrication. Metallisation and 'lift-off' processes, which are crucial to the formation of HEMT contacts follow, preceding a discussion of wet and dry etching techniques and their applications to HEMT technology. Combining each of these, a generic HEMT process flow is then presented, highlighting the key issues for high performance device fabrication. The self-aligned gate process in which the standard process is altered to minimise device parasitic resistances follows. Finally, alternative lithography techniques such as nanoimprinting and their potential in short gate length HEMT realisation are discussed.

3.1 Electron Beam Lithography

The ability to define ultra-small features on semiconductor layer structures is a crucial factor in the realisation of high specification HEMT devices, as reduction of the gate length of such devices directly affects the device performance. The process by which patterns are transferred onto the semiconductor material relies upon the patterning of a layer of 'resist', which covers the top surface of the semiconductor material. For photolithography, the resist used is typically sensitive to the UV range of the spectrum and hence is patterned by exposing selected areas to UV radiation. The areas of exposure are selected by placing a mask plate between UV source and substrate that allows the transmission of radiation onto selected areas of the resist, as defined by the transparent pattern on the mask. This process is illustrated in **Figure 3.1.1**:

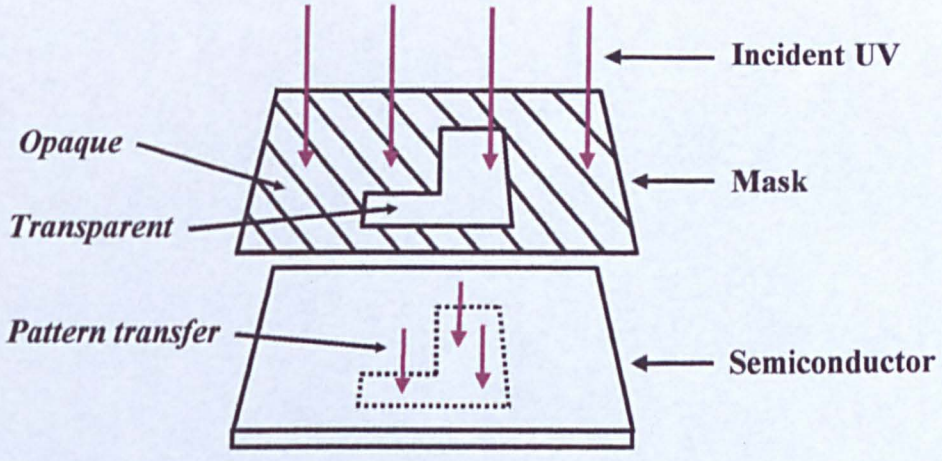


Figure 3.1.1 - Photolithography process

UV radiation will pass through the transparent areas of the mask and expose the corresponding areas of resist on the semiconductor. A chemical reaction then occurs within the exposed areas of resist which alters its composition and allows its selective removal with appropriate solutions. The process of removing the exposed areas of resist is referred to as *developing*. In the instance where the resist that has been exposed is removed, the resist type is said to be positive, while other resists exist with which the un-exposed areas will be removed by developing, and are labelled negative resists.

A similar process occurs with electron beam lithography in that selective areas of resist upon the semiconductor surface are exposed, although the process of exposure is very different. Instead of sensitivity to a particular wavelength of light, ebeam resists will be exposed upon the bombardment of high energy electrons and similarly developed out using suitable developing solutions. Also differing from photolithography, instead of using a mask to define a specific pattern from an isotropic source, ebeam lithography uses a concentrated beam or 'spot' of electrons which under computer-controlled software, draws the desired pattern onto the resist. Unlike UV radiation, the ability to generate a stable source of high-energy electrons and focus them to a usable spot size in conjunction with controlling their trajectory to within an acceptable error margin for small feature definition becomes a complicated task. In general, photolithography provides an efficient and cost effective lithographic process that provides a rapid device turn around. However, the direct photolithography process as that described in **Figure 3.1.1** limits the minimum defined feature size attainable to that of the wavelength of the light in addition to limiting diffraction effects from the masking material. These limitations can and have been exceeded in modern photolithography defined IC processing by introducing various resolution enhancement techniques (RETs) that continuously allow the minimisation of feature size definition using set wavelength optical lithography processes [3.1]. These include "Mask Proximity Correction" with which the mask geometry is altered from the desired pattern by dimensions below the resolution of the imaging system. In addition, to account for loss of contrast with overlapping or closely spaced features, "Phase Shifting" or "Off-axis Illumination" techniques help to boost image contrast and improve feature definition. The former relies on the creation of destructive interference regions by passing the light through a mask of varied topography resulting in adjacent waves being out of phase, while the latter utilizes various shaped optical elements placed between source and mask to direct light at various angles at the mask. Using these techniques combined with the wavelength of the light source of 193nm used for modern IC processing, sub 100nm features are readily defined and have already been incorporated into large scale device manufacture [3.1].

Beyond the cost efficiency and fast turn around of devices fabricated by an optical lithography process, the lack of versatility for defining small feature sizes for the purpose of research points towards ebeam lithography as a more versatile albeit slower and generally more expensive lithography technology.

The Glasgow electron beam lithography system

Although the process of ebeam lithography is considerably more complex than photolithography, ebeam tools exist in which very small spot sizes can be achieved combined with very high alignment accuracy [3.2]. The Leica Cambridge Electron Beam Pattern Generator 5 (EBPG5) used in the Nanoelectronics Research Centre at the University of Glasgow on which all the lithography work was performed on this project, has a minimum beam spot size of 12nm and can be operated at acceleration voltages of 20, 50 and 100kV. Here follows a brief synopsis of its operation:

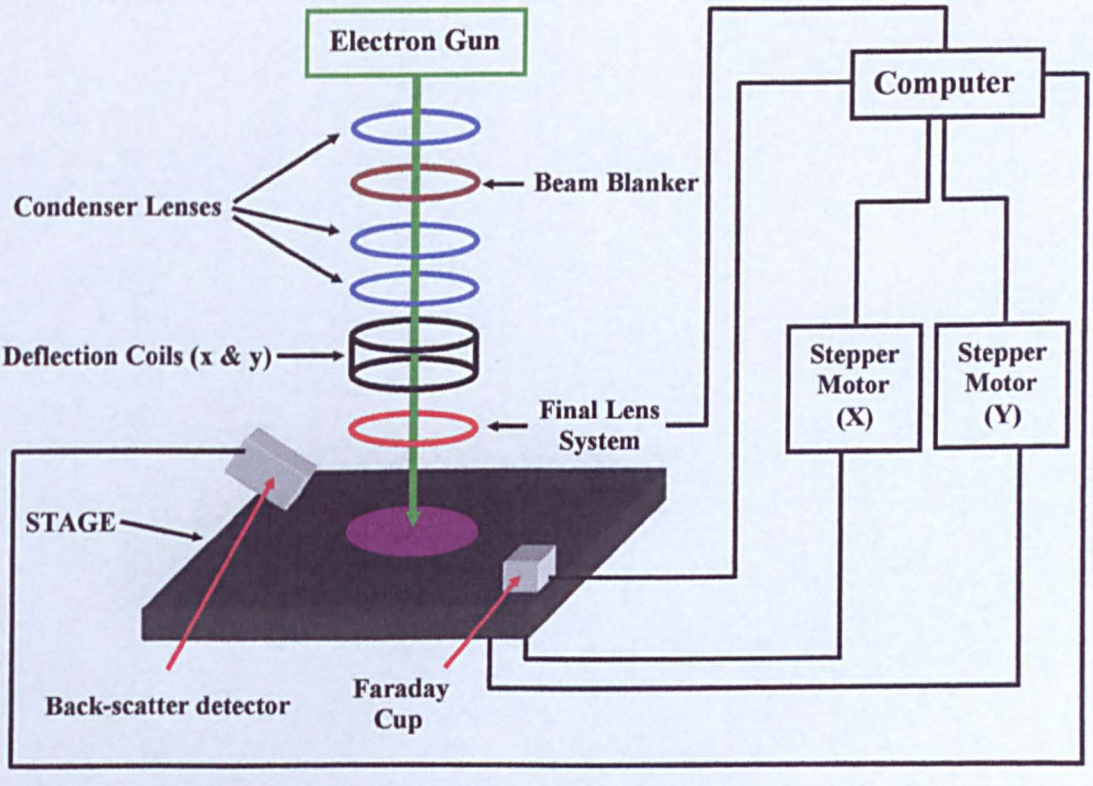


Figure 3.1.2 - EBPG5 ebeam system set-up

With reference to **Figure 3.1.2**, an electron gun generates a stream of electrons which are focused to a spot by three separate condenser lenses. The beam blanker acts as an on/off

switch for the beam to avoid unwanted exposure. Deflection of the beam in x and y directions is achieved with the two deflection coils. Due to the limited field of deflection of the beam, (~ 800 μm square at 50kV) the stage upon which the substrate is placed is mechanised with stepper motors operating in x and y directions allowing movement of the wafer and hence a greater field of writing. Prior to exposure and after the beam is deflected, a final set of lenses under computer control auto focuses the beam to the desired spot size. Also, mounted on the stage are a Faraday cup which monitors the effective beam current and an electron back-scatter detector. The ability to alter the accelerating voltage with the EBPG5 becomes extremely useful, particularly when defining very small feature sizes. At the maximum acceleration voltage of 100kV, the incident electrons exposing the resist and penetrating the semiconductor will be less subject to back scattering [3.3]. This back-scattering often results in greater exposure of the resist and can lead to an unwanted increase in feature size.

Ebeam resists

In addition to the operation of the ebeam tool used for lithography, the specific resist system that is to be written also plays a large role and is generally chosen depending on the lithographical requirements. Various types of resist exist and are currently in use within the Nanoelectronics Centre at The University of Glasgow, however only those used throughout this project are considered here. The three main types of resist used for all fabrication performed include; 1. Poly-Methyl Methacrylate (PMMA), 2. Copolymer of Methyl Methacrylate and Methacrylic acid P(MMA/MAA) and 3. DUV Shipley Ultra Violet III (UVIII). PMMA and copolymer resists are both solely ebeam resists and hence will only be exposed by an electron beam. UVIII however will be exposed under both ebeam and UV exposure and is hence suitably versatile to be popular with a number of lithographic processes [3.4].

The process by which resist is coated onto a semiconductor substrate involves initially dissolving it into a solvent, followed by spin coating onto the substrate at high rpm to

achieve a uniform coating. PMMA is typically dissolved in Oxylene, in which the thickness of the resist layer can be tailored by altering the ratio of resist to solvent, combined with altering the speed at which the resist is spun. Copolymer and UVIII resists are dissolved in ethyl lactate. As with PMMA, the thickness of both can be altered by changing the solvent - resist ratio or the spin speed.

Following spinning of the resist and prior to ebeam exposure, the substrate is baked typically at either 120°C or 180°C for a few hours depending on the resist strategy and the particular process level. This acts to drive off the casting solvent. When using UVIII, a short 120°C post exposure bake is required which is essential to controlling the resist profile.

The process by which ebeam resist is exposed and developed occurs by the breaking of the long chain polymers that compose the resist. On exposure by the electron beam, the smaller molecular weight of these broken chain polymers allows their selective removal when immersed in suitable developing solutions [3.5]. Typically PMMA and copolymer resists are developed in a combination of 4-methyl pentan-2-one (MIBK) and Isopropyl Alcohol (IPA), while MF CD26 is used to develop UVIII. Again, the ratio of the MIBK to the IPA in the developer will alter the developing time and resist profile and can therefore be chosen to suit a particular process. The temperature of the developer solution also becomes critical for smaller feature size definition.

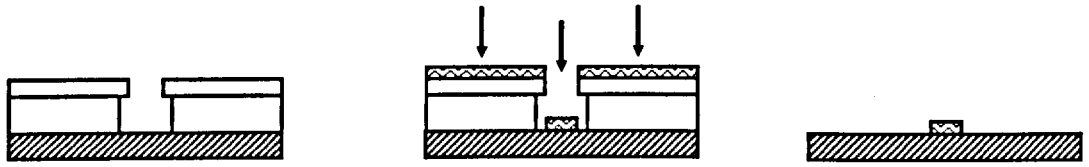
Finally a process known as "ashing" is often used post development which involves exposure to a low power oxygen plasma to remove any residual resist that might remain in the developed areas on the substrate.

Pattern Design and Software

For pattern design and multi-layer Gds II file creation, the Wavemaker (WAM) software package was employed throughout this project. The Gds II hierarchical pattern files are subsequently fractionated by CATS software into cflt file format which are converted to IWFL files readable by the ebeam tool by software known as "belle", written in the department. The belle software also permits the designation of writing variables such as spot size, dose and resolution, all of which are carefully chosen to suit the desired lithographic requirements. By incorporating a ccfa file at the fractionation stage, various layers defined within the Gds II file can be allocated multiple values of a base dose allowing different exposures to be allocated to selective areas of the substrate in a single lithography session. This becomes particularly useful for patterning multi-resist layers to produce varied height geometry patterns.

3.2 Metallisation and Lift-off

Following pattern definition in the resist using ebeam lithography, metal can be deposited across the substrate, forming the metallic contacts required for the source, drain and gate of a HEMT device. As the metal will only come into contact with the semiconductor in the exposed areas, the metal deposited upon the remaining resist can be "lifted-off" by removing the resist through submersion in acetone. This leaves a metallised pattern on the substrate defined by the original lithography. This process is illustrated in **Figure 3.2.1**:



1. Lithography and development **2. Metallisation** **3. Lift-off**

Figure 3.2.1 - Metallisation and lift-off process

When performing a metallisation stage, it is common practise to use a bi-layer of resist as illustrated in **Figure 3.2.1**, as opposed to a single layer. The top layer of resist is chosen to be less sensitive to exposure from the lithography than the underlying layer and hence an undercut in the profile is produced. Once metal has been deposited, this undercut aids the lift-off process by producing a cleaner discontinuity between the metal deposited onto the substrate and that on the resist. Often without this undercut profile, metal can accumulate on the side walls of the resist which remain after lift-off, leaving "flags" or large metal peaks at the edge of the resultant metallisation.

All metal deposition throughout this work was performed on one of two metallisation units. The first, a Plassys MEB 450 Electron Beam Evaporator (Plassys I) allows the deposition of various ohmic and gate metals including Au, Ti, Ge, Ni, and Pd as well as exhibiting a low base pressure of $1x10^{-7}$ Torr, and was used for the majority of metallisations throughout this research. The second, a Plassys MEB 550S (Plassys II) which allows the metallisation of

similar metals to Plassys I, but with the replacement of Pd with Pt provides a lower base pressure than Plassys I of $1x10^{-8}$ Torr, was used later in the project for gate metallisations replacing the Pd with Pt. This lower pressure provided by Plassys II results from the cryogenic based pumping system compared with the diffusion based system for Plassys I.

For the formation of effective ohmic and gate contacts, it is prudent to incorporate a de-oxidation (de-ox) treatment post development and ashing, and prior to metallisation to remove any native oxide formed on the semiconductor surface. The source and drain ohmic contacts rely on the low-resistive passage of current through the metal-semiconductor contact into the device which will be degraded under the presence of an insulating oxide layer. Similarly, control of the channel current by the gate will also be reduced through a voltage drop across any dielectric material between gate metal and semiconductor layers. Typically for ohmic contacts a hydrochloric acid (HCl) based de-ox solution is sufficient to remove any native oxide, while for the gate contact various processes such as hydrofluoric (HF) or orthophosphoric acid based cleans can be utilised, depending on the material in question. Further details on the processes used in this work are given in **Appendix A.**

3.3 Etching

Following pattern definition by lithography, it is possible to selectively etch areas of the device material through either dry or wet etch processes, using the resist as an etch mask. The process of dry etching involves using plasmas from selective gases within a controlled environment to etch the semiconductor material. The gases used will depend on the material that is to be etched while etch rates are generally controlled through temperature, pressure and concentration monitoring as well as etch duration. As suggested, wet etching involves the submersion of the material into specific etching solutions that act to dissolve the material. Again particular solutions are used depending on the material, but fewer variables are involved than with a dry-etch process. Typically the etch rate will depend primarily on the solution composition, concentration and the duration of the etch but also with temperature and pH leading to second order effects. Both etching processes often offer the ability to selectively etch one material over another in which the etch rate for one will be a significantly faster than the other. This process becomes extremely useful for the gate recess etch in HEMT technology, where the cap layer of the material is selectively removed and the gate metallised onto the subsequent barrier layer.

Throughout this work wet etch processes were developed and used exclusively for all of the main process etch requirements. This was mainly due to potential damage to the material that can arise using dry etch combined with a lack of etch processes available for the selective etching of indium based compounds fundamental to many III-V HEMT material structures [3.6].

3.4 Surface profiling and Atomic Force Microscopy (AFM)

Often it is desirable during fabrication to profile the surface of the device structure at a particular point in the process. This can provide crucial feedback on quantities such as metallisation heights and etch depths, as well as resist profiles. Atomic Force Microscopy (AFM) provides a very effective route to three-dimensional surface profiling. The AFM process relies on the scanning of a fine silicon tip attached to a cantilever across the plane of the structure in a two dimensional fashion. Vertical displacement of the cantilever is detected by reflecting a laser off the reverse side of the cantilever whose position is monitored by a four-grid photodiode array. Through feedback from the photodiode signal, the laser's position is maintained in the centre of the 2x2 array, while the differential feedback provides the information as to the vertical position of the cantilever and tip. This process combined with fine stepper motors controlling movement of the stage in the plane of the surface scanned, produces a three-dimensional map of the surface. Also to ensure minimal damage to the surface scanned and to increase the lifetime of the tip, modern AFMs often incorporate a 'tapping mode' which instead of the tip being applied constantly to the surface, it is tapped upon it at a high frequency.

All the surface profiling performed throughout this research was done upon a Dimension 3100 AFM, which demonstrates a sub-nm height resolution. Generally the limitations of the AFM will be related to the ability to fit the tip into the area that is to be scanned. For example measuring the depth of a 30nm wide trench in resist that is suspected to be 100nm thick would be challenging if not impossible simply due to the tip dimensions being too large to reach the bottom of the trench. As a general rule, for effective sub-100nm trench profiling, the ratio of the height to width of the trench that can be scanned effectively should be about 1:1.

3.5 Standard HEMT process flow

The basic process flow by which a HEMT device can be fabricated using the aforementioned techniques can be categorised into five main levels:

1. Markers
2. Mesa Isolation
3. Ohmic Contacts
4. Gates
5. Bondpads

The general order of the process flow is also as above, with some potential variation between the isolation and ohmic levels which is highlighted later. Each of these levels will now be discussed in detail, highlighting the key issues of each and where applicable the relevance to or impact on device performance. The details of each specific process flow are given in **Appendix A**.

1. Markers

Initially metallic markers are defined upon the blank substrate through ebeam lithography, metallisation and lift-off techniques as previously discussed. These generally consist of an array of 20x20μm squares that allow the alignment of subsequent levels by ebeam, (**Figure 3.5.1**).

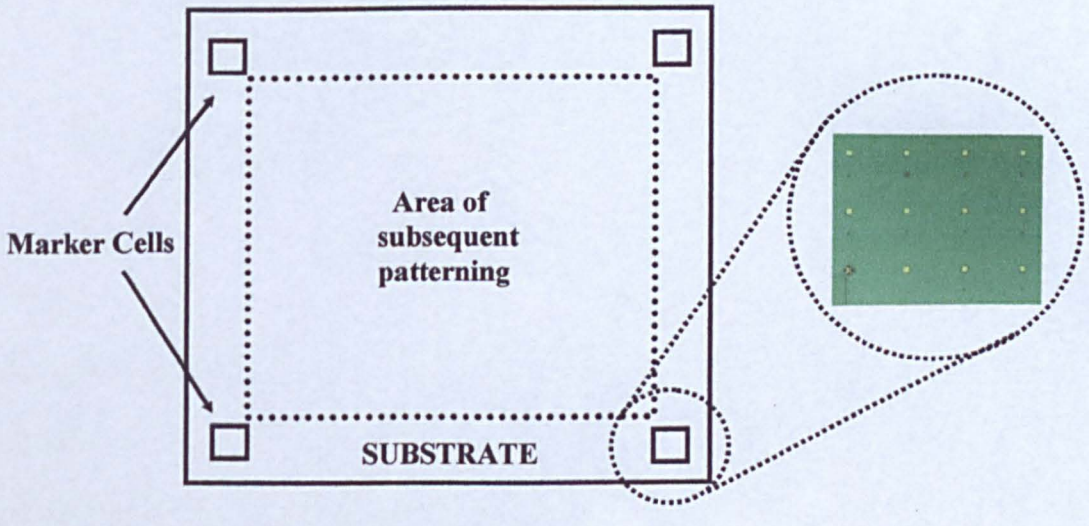


Figure 3.5.1 - The Marker Level

Typically marker cells that consist of an array of individual markers are located at the corners of the area within which the devices will be fabricated. When performing a subsequent lithography level that requires alignment (referred to as a registration level) the electron beam that is used to perform the lithography can also be used to detect the marker position. By monitoring the intensity of the backscattered electrons, the contrast between the metal that makes up the markers and the substrate material can be detected if the marker metallisation is sufficiently thick (at least 100nm Au is recommended). The location of each of the markers can therefore be identified with great precision and the registration level aligned accordingly. In addition, by using an array, a different marker can be used for each level of subsequent alignment as the ebeam will expose any markers previously located which can lead to their deformation.

2. Mesa Isolation

The Mesa Isolation level provides a method of electrically isolating regions or islands of material with an aim to defining the active geometry of the device. This is achieved by defining areas of resist with ebeam that act as a mask to a non-selective wet etch that etches to below the upper conducting layers of the layer structure. This process is illustrated in **Figure 3.5.2**:

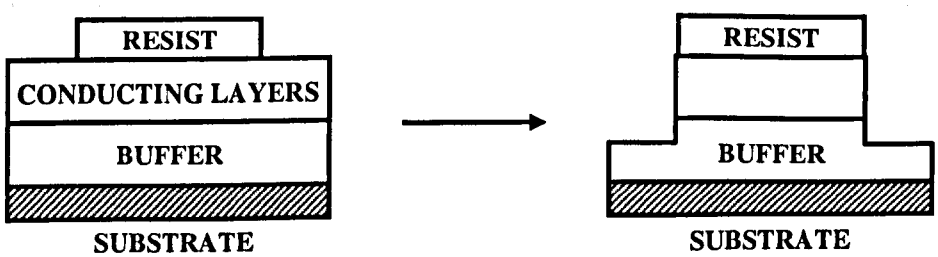


Figure 3.5.2 - Mesa Isolation process

Through depth characterisation with the AFM, the resist patterned substrate is subjected to repeated submersion in the etch solution until the desired etch depth is reached. The etch solution used will also depend on the specific material that is to be etched. For GaAs and InP

based materials, ammonia and orthophosphoric based isolation etches are used respectively. The desired depth of etch is typically just beyond the channel layer into the non-conducting buffer. The flexible order of the standard device process also allows the ohmic contact level, which is described next, to be performed prior to mesa isolation. This allows the depth of etch to be characterised electrically by monitoring the current between two adjacent isolated regions of material with ohmic contacts. It is also desirable to minimise the depth of the isolation etch, as later in the process the gates are metallised onto the isolated region and up onto the active mesa. Too deep an isolation etch can result in a discontinuity in the gate at the mesa edge. The typical geometry of the mesa isolation is given in **Figure 3.5.3**.

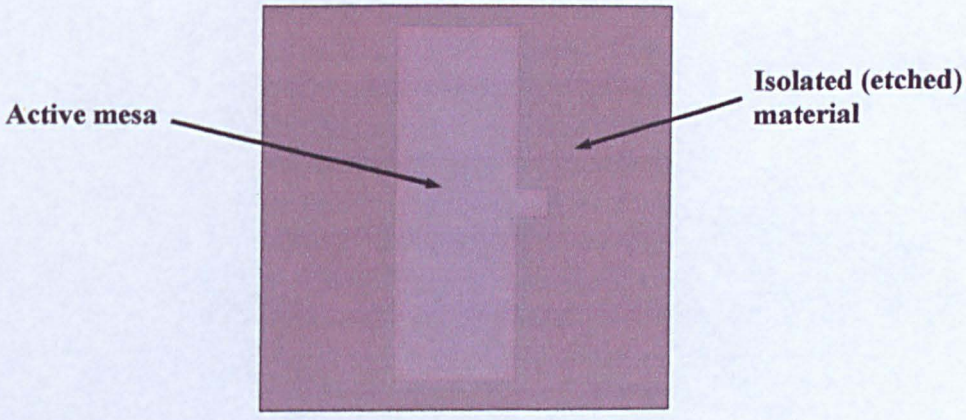


Figure 3.5.3 - Isolated mesa geometry

The mesa is typically rectangular in shape whose width defines the effective width of the resultant device. The protruding section of mesa on the right makes contact with the drain pad metallisation at the bondpad level of the process.

3. Ohmic Contacts

Similar to the marker level, the metallic ohmic contacts are defined through patterning of the resist followed by metallisation and lift-off. These are metallised within the active mesa region defined by the isolation etch and form the source and drain contacts of the device. **Figure 3.5.4** shows the typical geometry of the ohmic metallisation.

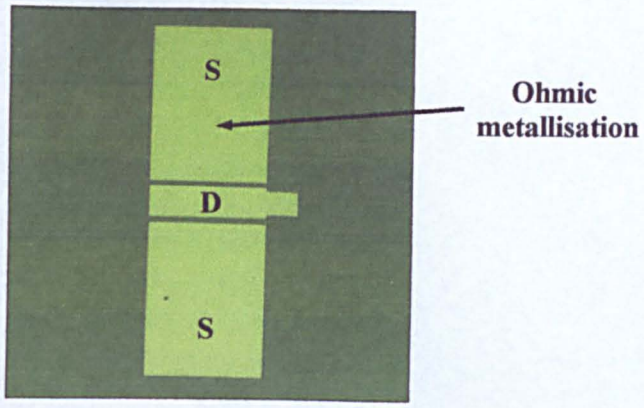


Figure 3.5.4 - Mesa plus ohmic contacts

The source and drain contacts are typically set 1.5μm - 2μm apart to allow the subsequent definition of the gate between them. Prior to metallisation, ashing again removes any residual resist in the developed areas.

After the ohmic metal has been deposited and lifted-off, the substrate is typically annealed up to temperatures of close to 400°C for 30 seconds to 1 minute using a FAV4 rapid thermal annealer (RTA) in a nitrogen atmosphere. By annealing, the ohmic metals diffuse into the semiconductor layers to help create a low resistance region for conduction between contact and channel. The choice of metals that compose the contact vary depending on the particular material structure. Often for GaAs based structures, germanium is favoured due to it forming a highly doped conductive region when diffused into GaAs [3.7]. Gold is also favoured due to its intrinsic high conductivity. Beyond these, various ohmic contact recipes exist, comprising of different metals and annealing strategies which are reported by various sources to produce favourable results [3.8-10].

4. Gates

The gate level is by far the most important of the entire process, as its formation will define to a large degree the overall performance of the completed device. Up until this point the gate has been considered to be rectangular in cross section akin to the ohmic or marker metallisations. However in an optimised high frequency device, it is beneficial to define the

gate having a 'T' shape. This allows the gate length to be minimised through reducing the size of the contact area to the material, while reducing the resistance through the gate along the width of the device i.e. perpendicular to the length of the gate, by including a large metallic head. In **Chapter 2** the role of parasitic resistances and their impact on device performance was discussed. By incorporating this T-gate profile, the gate resistance is minimised leading to an increase in f_{max} and to a lesser degree f_T.

This of course complicates the lithography from simply defining a rectangular type metallisation. However the problem is tackled by using a multi-resist stack combined with a ccfa process to allocate different doses to selective areas of the resist during a single exposure session. The result is given in **Figure 3.5.5**, which shows the cross section of the resist profile from a typical 120nm T-gate.

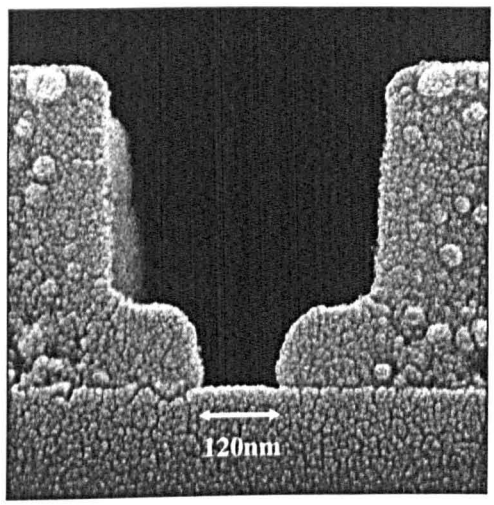


Figure 3.5.5 - 120nm T-gate profile

The image from **Figure 3.5.5** was taken using a Hitachi S900 scanning electron microscope (SEM) which exhibits a range of beam energies of 1 to 30keV and a minimum feature resolution of 5nm. The resist and substrate is sputtered with a thin layer of Pd/Au to improve the quality of the SEM image.

For 120nm gates, a tri-layer of PMMA/copolymer exposed at 50kV is used to produce a gate resist profile like that shown in **Figure 3.5.5**. For reduced gate lengths of 70nm and 50nm, a

UVIII/LOR/PMMA resist stack exposed at 100kV is used to maintain the large gate head / short gate length ratio. LOR or lift off resist provides a method of using both UVIII and PMMA without resist intermixing. The development of the profile is then performed in two stages, the first to develop out the gate 'head' in the UVIII followed by the second which develops the gate 'stalk' in the PMMA.

Following definition of the gate profile, ashing removes any residual resist within the gate foot area without damage to the material. The gate recess etch is then performed in which the highly conducting cap layer is etched, and the gate metallisation formed onto the barrier layer. For shorter gate lengths i.e. sub 100nm, and to promote higher breakdown voltages, it is desirable to move to a double recess process. This is performed by two sequential etches which first etches a window in the cap and then a narrow trench into which the gate is metallised, minimising the effective gate length. For all the recess work performed in this project, a selective succinic acid based wet etch was developed which can be applied to both GaAs and InP based HEMT material structures. The etch which is pH balanced to ensure a predictable etch rate, was found to be extremely controllable as well as repeatable. This allowed the specific size of the recess (i.e. the lateral distance of the etch) to be tailored to the specific length of gate. For a gate profile of 120nm, a recess length of around 25nm is desirable as shown in **Figure 3.5.6**:

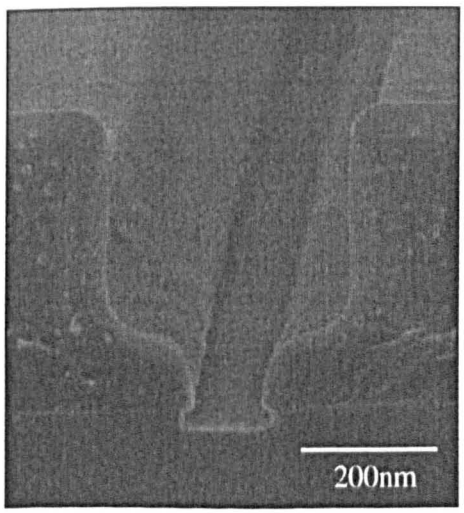


Figure 3.5.6 - 120nm T-gate profile plus recess etch

Following the succinic acid etch, a post recess clean is performed immediately prior to metallisation to remove any oxidised residue that may have formed on the barrier layer. For both GaAs pHEMT and InP based devices, the aluminium composition of the barrier layer will oxidise readily, which upon gate metallisation will hinder device performance through reduced control of the gate. The lower aluminium content typically used in the barrier layer for GaAs pHEMT structures means that a dilute hydrofluoric acid (HF) clean is sufficient to remove any oxidised layers. The higher aluminium composition of the barrier in a InP based structure can create problems when attempting to remove the oxidised layers using only HF. Throughout this work a short dilute orthophosphoric etch was performed post succinic acid etching for all InP based structures, which was found to produce a high quality Schottky gate contact.

Two types of gate metallisation were utilised throughout this research, namely the original titanium (Ti), palladium (Pd), gold (Au) recipe, which was then replaced by a similar metallisation of titanium, platinum (Pt), gold, with the introduction of the new metallisation unit, the Plassys II. Ti in both recipes acts as excellent adhesion layer to the barrier layer of the material in addition to providing a good Schottky contact. This is followed by a thin layer of Pd in the original process, and superseded by Pt in the revised process. Both Pd and Pt act as barriers to prevent diffusion of the gold upper layer of the gate into the semiconductor, which would otherwise degrade the gate contact. The benefits of using Pt over Pd would appear to include greater gate yield, most likely due to the robust nature of Pt when compared to Pd. The final thick layer of gold acts as a highly conductive medium along the width of the gate, acting to further reduce the gate resistance. A completed metallised gate of 70nm gate length with single recess is presented in **Figure 3.5.7**:

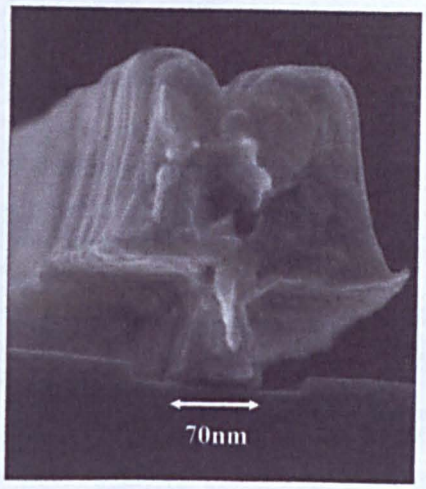


Figure 3.5.7 - Metallised 70nm T-gate including recess

In a real device, electrical contact has to be made with the small footprint T-shaped gate. At one end of the gate feature, 'gate feeds' are defined which provide a connection to measurement pads defined later. **Figure 3.5.8** continues the evolution of the device by displaying a plan view of the isolated mesa with ohmic contacts and now gates.

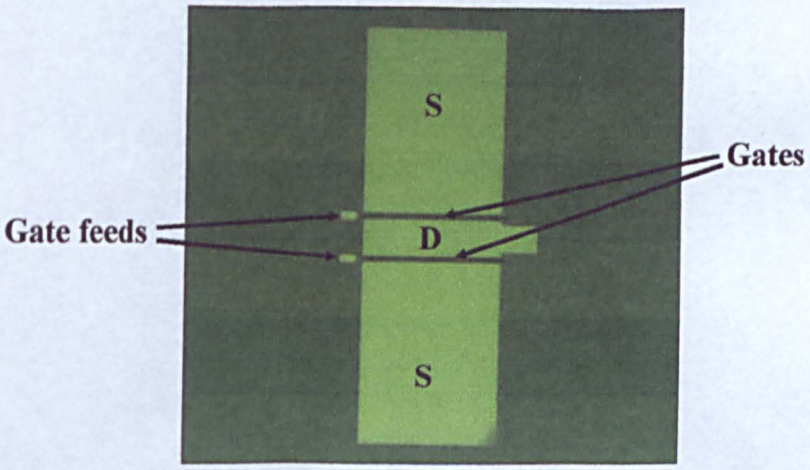


Figure 3.5.8 - Device with isolated mesa, ohmic contacts and gates

The gate feeds can be seen clearly on the left hand side of the figure and are metallised up onto the mesa. This provides a safer method of ensuring the continuity of the gate across the mesa edge by defining a larger metallic 'wedge' instead of the T-shape part of the gate. Unfortunately this additional contact results in an increased gate capacitance, which will be more significant for narrower devices with lower 'intrinsic' gate capacitance, generally resulting in better RF performance for wider devices.

5. Bondpads

To allow on wafer probe measurements of the completed device, large metallic bondpads are defined to make contact with the source, drain and gate contacts. In addition, to perform high frequency measurements they are designed to accommodate the geometry of the measurement equipment and probes. Marker lines also defined at the marker level indicate the required placement of the probes upon the pads to allow definition of the calibration reference planes (see **Section 4.3**). The metallisation used for the bondpads consists of a nichrome (NiCr) adhesion layer followed by a thick gold layer (typically 1 μm) to minimise access resistances to the three terminals. The completed device structure is now given in **Figure 3.5.9**, displaying the isolated mesa, ohmic contacts, gates and bondpads.

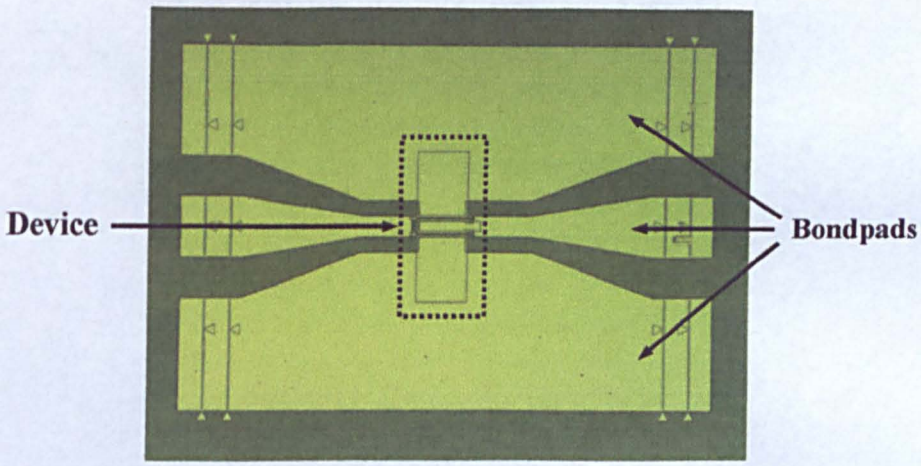


Figure 3.5.9 - Completed HEMT device geometry

The bondpads for all three terminals are deposited onto non-conducting material exposed after the mesa isolation etch, which minimises signal loss at high frequencies to the substrate. However an element of signal loss will still occur through capacitive coupling between the large area of the bondpads, and any free residual charge within the isolated area beneath the pads. These parasitic capacitances are extracted as part of the modelling process from device measurements, and allow the creation of an accurate device model. The specifics of device measurement and characterisation are discussed further in **Section 4.3**.

3.6 The Self-aligned gate process

An alternative process to that described for the standard device involves the minimisation of source and drain parasitic resistances through the reduction of the source - drain ohmic contact separation. This is achieved by defining the gate prior to the ohmic contacts, then depositing a thin layer of ohmic metal across the device structure, forming the source and drain contacts. The gate need not therefore be aligned between the ohmic contacts and is said to be *self-aligned.* This process is illustrated in **Figure 3.6.1**:

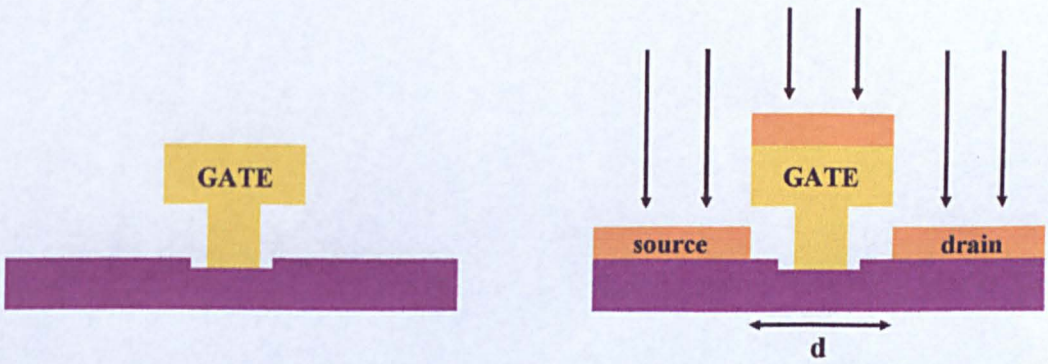


Figure 3.6.1 - The Self-aligned gate process

If the ohmic metal is sufficiently thin, i.e. less than the height of the 'stalk' of the T-gate, a discontinuity will exist between the gate and ohmic metal on the substrate. This results in an effective ohmic contact separation defined by the length of the head of the gate given by **d** in **Figure 3.6.1**. Compared with the standard process, which defines the source - drain separation to be between 1.5μm and 2μm, the typical dimensions of a T-gate result in this separation being reduced to around 400nm. This leads directly to a reduction in the parasitic access resistances compared with the standard process. In addition, ohmic metal will be deposited on top of the gate which further reduces the gate resistance.

The self-aligned process however places certain restrictions on specific elements from the standard process. In reversing the order of processing the ohmic and gate levels, the

annealing of the ohmic contacts is performed after the gate metallisation. This can result in severe degradation of the Schottky contact that forms the gate, as the gate metal can also diffuse into the semiconductor depending on the metal in question and the annealing temperature. The annealing temperature must therefore be minimised to maintain an efficient Schottky gate contact and minimise any gate 'leakage', while still producing a low resistance ohmic contact. In addition to devising an ohmic process in which the thermal budget can be kept low, the dimensions of the contacts themselves also place further restrictions on the process. To ensure a discontinuity between the ohmic contacts and the gate along the width of the device and hence eliminate any shorting between them, the height of the ohmic metallisation must be kept to below that of the stalk of the gate. This requires an ohmic metallisation height of typically less than 100nm, the average stalk height of the T-gate process. The challenges involved in devising such a self-aligned gate process therefore depend on the ability to form very thin, low temperature ohmic contacts that exhibit a low enough contact resistance to justify the use of the self-aligned process to minimise device parasitic resistances.

In addition to reducing associated parasitic resistances, the self-aligned process also provides various other benefits including the ability to lithographically define the gate level upon a planar surface. With the standard process the gates are patterned in resist between the source - drain contact metallisations. This can lead to non-uniformity of the gate resist profile along the width of the device due to resist thickness fluctuations and in the extreme result in 'patchy' gate lithography.

This process is illustrated in **Figure 3.6.2**:

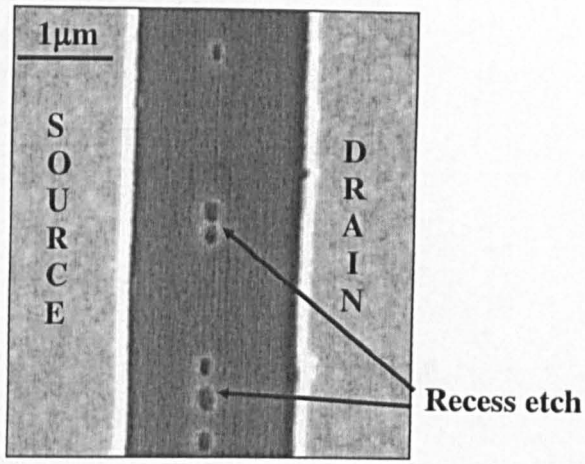


Figure 3.6.2 - Patchy gate lithography in source-drain gap shown by recess etch

By performing a gate recess etch following definition of the gate profile, the areas of underexposure as a result of resist thickness fluctuations become apparent upon inspection after removal of the resist. The example given in **Figure 3.6.2** shows the severity to which this effect can occur, and in this case only isolated areas of the profile have been exposed sufficiently to allow etching of the cap layer. Although this effect can be controlled to a certain degree by increasing the dose of exposure by ebeam, this will generally result in larger feature size, or a variation in feature size across the width of the device. Hence for shorter gate length devices this becomes a highly undesirable process, but one that can be avoided by defining the gate level upon a planar surface.

It has also been suggested that by minimising the thermal budget for the entire process, the device performance will be enhanced through reducing degradation of the abrupt MBE grown layers that can occur at high temperatures [3.11]. The heterojunctions, so crucial to device operation are found to be degraded when exposed to high temperatures and arguably this degradation can occur at typical annealing temperatures i.e. 400°C [3.11]. For shorter gate length devices and hence aggressively scaled material, such degradation could impact transport properties and hence reduce device performance.

Beyond the various benefits offered, one drawback associated with a self-aligned system is a decrease in the breakdown voltage of the completed device. In **Section 2.4** the issue of device breakdown was discussed and related to the bandgap energy of the channel material as well as the creation of the high field at the drain end of the gate. The large potential difference between drain and gate results in this high field region between the gate and drain contacts in which impact ionisation leading to breakdown will occur at a particular gate / drain voltage. This effect is exaggerated with the self-aligned system compared to that with a standard device due to a reduction in separation between drain and gate contacts, leading to much larger fields and hence breakdown at a lower source-drain voltage. One method of minimising this problem is by adopting a gamma gate process as opposed to the standard T-gate, with which the drain contact is spatially separated further from the gate than the source is from the gate. This process is illustrated in **Figure 3.6.3**:

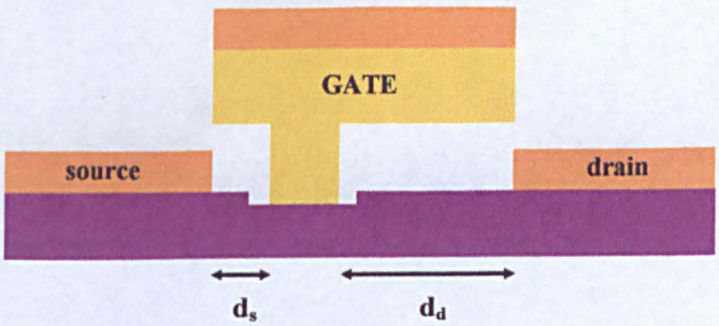


Figure 3.6.3 - The Self-aligned gamma gate (Γ) process

The gamma gate (Γ), so named because of its resemblance to the Greek capital gamma, is an asymmetric T-gate as presented in **Figure 3.6.3** that provides a method of increasing the drain - gate separation with a self-aligned process. Even though this directly increases the drain parasitic resistance in which a self-aligned process seeks to reduce, as discussed in **Section 2.5** the majority of performance degradation arises from the source parasitic resistance which remains minimised. Therefore the Γ-gate process provides a compromise

between increasing the breakdown voltage of the device while still reaping the increased performance benefits of a self-aligned system.

Often the aspect ratio of the Γ-gate is quoted which is merely the ratio of the lengths of the head of the gate on the drain and source sides, or the ratio of separation between the gate and drain and gate and source contacts given by d_d and d_s in **Figure 3.6.3,** i.e. d_d / d_s.

Beyond the issue of device breakdown, the parasitic capacitances that exist through the air between the metal gate and source/drain contacts will also be increased with the self-aligned process through reduced separation between the contacts. The extent to which this process reduces the device performance is open to speculation however when compared with the potential benefits that a self-aligned process has to offer for sub 100nm gate length devices.

3.7 Nanoimprint Lithography

In addition to ebeam and photolithography, a newer technique known as Nanoimprint lithography is in development which instead of using some form of radiation for lithographic pattern transfer, utilises a mechanical process to 'stamp' the desired pattern into resist. The dimensions of the tool "shim" used for the stamping process will then define the pattern geometry within the resist. Once the pattern has been transferred to the resist, metallisation or etch processes can be performed as with post ebeam or photolithography. An example nanoimprint T-gate process is presented in **Figure 3.7.1**.

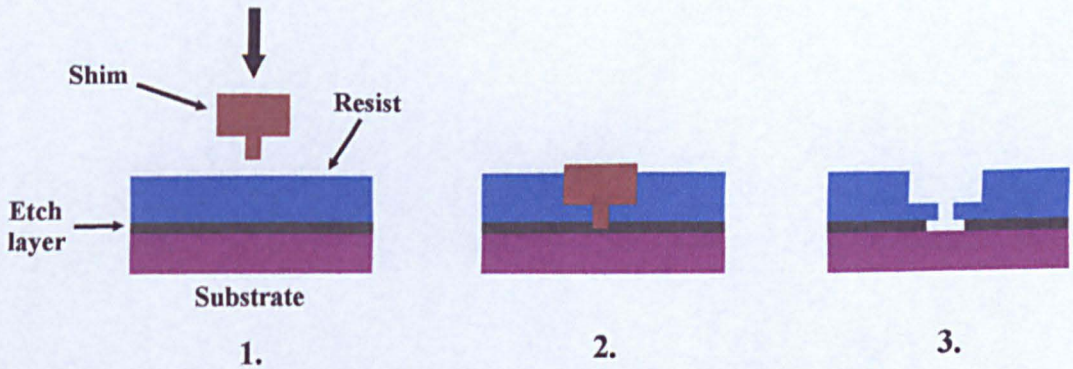


Figure 3.7.1 - Nanoimprint T-gate process.

The shim, which is shaped to produce a T-gate profile is imprinted into the resist upon the substrate. A thin layer of etchable material separates the resist and substrate in which the end of the shim stops to prevent damaging the substrate material. Following removal of the shim, this layer is etched away beneath the T-gate profile without deforming the profile and produces an undercut beneath the resist. Etching, metallisation and lift-off of the gate is then performed as with the standard ebeam process.

The benefits of such a process include repeatability, as without process error, the pattern transferred by the shim should be identical with each use. Cost efficiency also becomes a significant factor, as the cost for running ebeam tools becomes expensive, while with a stamping tool maintenance is minimal and far less costly.

Problems can arise however when integrating nanoimprint lithography into a HEMT type process, as the patterning technique generally requires that the surface to be patterned be planar. In a conventional HEMT process, the gate is defined after the source and drain metal contacts and hence incorporating nanoimprinting into a standard process flow poses a number of significant challenges. These problems are overcome with the use of the self-aligned strategies described previously.

3.8 Chapter Summary

The issues associated with the realisation of HEMT devices and in particular the challenges associated with the fabrication of those with shorter gate lengths (100nm and below) have been outlined throughout this chapter. The various technologies implemented in the fabrication of devices have also been touched upon to provide an understanding into the current practical limitations in device realisation. These generally include the lithographical limitations associated with modern lithography tools in the definition of the gate length of the device, as this dictates to a large degree the overall device performance. The potential to further increase performance for a particular gate length through the incorporation of a self-aligned gate process has also been discussed, with the challenges associated in the development of such a process also brought to light.

Following the design and fabrication of devices, their characterisation and measurement provides crucial feedback into their operation in addition to the identification of figures of merit for device performance. The complexities of material and device characterisation are therefore discussed in the following chapter.

References

[3.1] F. Schellenberg, *A little light magic*, in *IEEE Spectrum*. 2003. p. 34 - 39.

[3.2] Y.Chen, S. Yoon, *Numerical Simulation for Single Electron Logic Gates. Proc. International Symposium on IC Technology & Applications*. 1997.

[3.3] S. M. Sze, *Semiconductor Devices : Physics and Technology*, 1985, John Wiley and Sons Inc. p. 443 - 446.

[3.4] S. Thoms, D. S. Macintyre, *Process Optimisation of DUV Photoresists for Electron Beam Lithography*. Microelectronic Engineering, 1999. 46: p. 287-290.

[3.5] D. G. Hasko, S. Yasin, A. Mumtaz, *Influence of developer and development conditions on the behavior of high molecular weight electron beam resists.* Journal of Vacuum Science & Technology B: Microelectronics and Nanometer Structures, 2000. 18(6): p. 3441-3444.

[3.6] S. K. Murad, S. P. Beaumont, M. Holland, C. D. W. Wilkinson, *Selective reactive ion etching of InGaAs and InP over InAlAs in SiCl4/SiF4/HBr plasmas.* Journal of Vacuum Science & Technology B: Microelectronics and Nanometer Structures, 1995. 13(6): p. 2344-2349.

[3.7] D. G. Ivey, S. Eicher, S. Wingar, T. Lester, *Performance of Pd-Ge based ohmic contacts to n-type GaAs.* Journal of Materials Science – Materials in Electronics., 1997. 8(2): p. 63-68.

[3.8] C. S. Wu, K. K. Yu, M. Hu, H. Kanber, *Optimization of ohmic contacts for reliable heterostructure GaAs materials.* Journal of Electronic Materials., 1990. 19(11): p. 1265 - 1271.

[3.9] P. A. Verlangieri, M. Kuznetsov, M. V. Schneider, *Low-resistance ohmic contacts for microwave and lightwave devices.* IEEE Microwave and Guided Wave Letters, 1991. 1(3): p. 51 - 53.

[3.10] Hawksworth, S. J. Chamberlain, J. M. Cheng, T. S. Henini, M. Heath, M. Davies, A. J. M. Page, *Contact resistance to high-mobility AlGaAs/GaAs heterostructures.* Semiconductor Science and Technology, 1992. 7(8): p. 1085 -1090.

[3.11] Y. Yamashita, A. Endoh, K. Shinohara, K. Hikosaka, T. Matsui, S. Hiyamizu, T. Mimura, *Pseudomorphic In/sub 0.52/Al/sub 0.48/As/In/sub 0.7/Ga/sub 0.3/As HEMTs with an ultrahigh f/sub T/ of 562 GHz.* Electron Device Letters, IEEE, 2002. 23(10): p. 573-575.

Chapter 4

Characterisation and Measurement

The ability to accurately characterise the material from which HEMT devices are fabricated as well as benchmarking the completed device performance becomes crucial to the cyclic process of design, fabrication and testing that forms the foundation of such progressive research. Implicit understanding of the device operation provides feedback into the individual elements that comprise the device, allowing continuous improvement with each subsequent generation as performance is continually pushed to its limits.

Fundamentally, and as has been discussed in **Chapter 2**, the material structure from which the device is fabricated, will determine the overall completed device characteristics to a large degree. Characterisation of the material therefore provides a basis into understanding the operation of the completed device. The elements of the device structure itself are also characterised through various techniques in addition to their DC and RF characteristics.

In this chapter the techniques used for both material and device characterisation are presented together with a discussion of how these results pertain to the overall extraction of figures of merit for the device performance. The Van der Pauw (VDP) method, which provides information concerning carrier densities and mobility through material structures is initially discussed, followed by accurate parasitic resistance extraction from real devices. The process by which devices are characterised at DC follows, with a discussion concerning S-parameters and RF measurements through to model extraction concluding the chapter.

4.1 Material Characterisation

The structure and quality of the MBE grown material used for HEMT fabrication plays an extensive role in the overall characteristics and performance of the device. As discussed in **Section 2.6**, scaling of the structure must be implemented with respect to the device gate length for efficient operation. In addition, the material doping profile will largely define the carrier concentration and distribution throughout the structure, as well as strongly influencing the low field mobility through the device channel. A high channel carrier density is desirable for increasing the drain current and reducing parasitic resistances, while a high mobility will lead to higher saturation velocities through overshoot with shorter gate length systems.

Values for the carrier concentration and mobility are readily extracted through Van der Pauw (VDP) measurements [4.1], which utilises the Hall effect with a symmetric four point structure under the influence of an external magnetic field. Two typical VDP structures are shown in **Figure 4.1.1**:

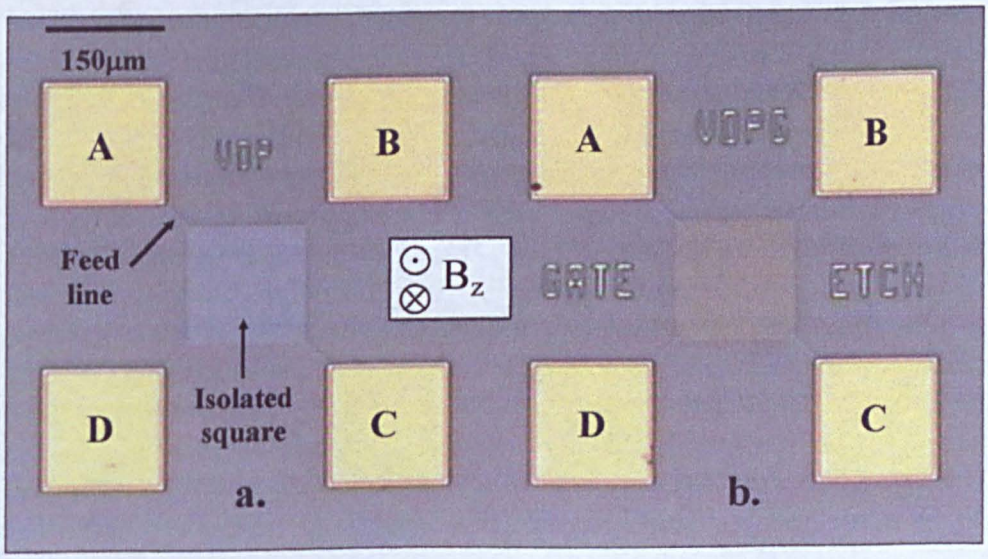


Figure 4.1.1 - VDP test structures with (a) and without cap (b).

The two structures presented in **Figure 4.1.1** demonstrate an isolated area of active material in the form of a square in the centre of each structure which is fed from four narrow 'feed' lines of mesa from four large pad areas. Metallic ohmic contacts are formed on these pad

areas to allow the structure to be tested electrically. Each pad has been labelled A, B, C and D for reference. Two similar structures are presented (a) and (b), differing only in that a selective recess etch has been used to remove the conducting cap layer from the central square region with the right hand structure (b), while the cap remains intact in the left structure (a). In removing the cap, the majority of conduction through the structure should occur solely through the channel layer and hence VDP figures extracted should relate to transport through the channel alone. Using the four ohmic probe pads, current is readily passed through the island of active material in the centre of each structure between probes, while the potential difference between adjacent probes is measured. Van der Pauw proved that between four arbitrary points A, B, C and D within a conductor, the following relation can de derived [4.1]:

$$\exp\left(-\pi\frac{d}{\rho}R_{AB,CD}\right) + \exp\left(-\pi\frac{d}{\rho}R_{BC,DA}\right) = 1 \qquad \textbf{Eqn 4.1.1}$$

where d is the thickness of the conductor, ρ is the conductivity of the material, and $R_{AB,CD} = V_{AB}/I_{CD}$, $R_{BC,DA} = V_{BC}/I_{DA}$, i.e. the ratio of the voltage between two points and the current between the adjacent two points. If the structure of the conductor is symmetric so that $R_{AB,CD} = R_{BC,DA}$ such as with the VDP structure, **Eqn 4.1.1** simplifies to:

$$2\exp\left(-\pi\frac{d}{\rho}R\right) = 1 \qquad \textbf{Eqn 4.1.2}$$

where $R_{AB,CD} = R_{BC,DA} = R$, and so the conductivity ρ can be expressed as:

$$\rho = \frac{\pi d}{\ln 2}R \qquad \textbf{Eqn 4.1.3}$$

The sheet resistance R_{sh} which is a useful figure for defining the two-dimensional resistivity of the material, is then defined as:

$$R_{sh} = \frac{\rho}{d} = \frac{\pi R}{\ln 2} \qquad \textbf{Eqn 4.1.4}$$

Hence the sheet resistance can be readily extracted from the four probe IV characteristics from the VDP structure.

Considering again the VDP structures in **Figure 4.1.1**, if a voltage V_{AC} is applied between pads A and C, current will flow between these points, the magnitude of which will depend on the electric field strength created by the applied potential and the transport of carriers through the material. If the field concentration remains sufficiently low to avoid saturation or breakdown effects, this voltage / current response should remain effectively ohmic and a linear response is observed. If a magnetic field is then introduced through the structure B_z, the carriers flowing under the influence of the applied potential will be subject to a force perpendicular to both the electric and magnetic fields. With the magnetic field normal to the surface of the structure, this causes current to flow perpendicular but in the plane to the original flow between contacts A and C, i.e. between contacts B and D. Without a route for this current to escape the structure, charge will accumulate at each of these contacts, giving rise to an electric field between them E_{BD}. The magnitude of this field soon becomes large enough to oppose the effect of the magnetic field and any current flowing between contacts B and D drops to zero. In this equilibrium state, the magnitude of the transverse field developed across contacts B and D can be related to the current flow I_{AC} and electric field E_{AC} between A and C and the magnetic field strength B_z to yield values for the carrier concentration and carrier mobility through the structure:

For equilibrium condition and zero current flow between contacts B and D, the force vectors affecting carriers in this direction must be equal, i.e.

$$q E_{BD} = q V_{AC} B_z \qquad \textbf{Eqn 4.1.5}$$

where V_{AC} is the velocity of carriers flowing between contacts A and C.

Therefore the velocity of the carriers flowing between contacts A and C can be expressed as the ratio of the transverse electric field to the magnetic field strength, i.e. $V_{AC} = E_{BD} / B_z$.

As the carrier velocity and electric field strength are known, the low-field carrier mobility μ is extracted:

$$\mu = \frac{V_{AC}}{E_{AC}} = \frac{E_{BD}}{E_{AC} B_z} \qquad \textbf{Eqn 4.1.6}$$

Furthermore, the current density J_{AC} between contacts A and C can be related to the conductivity of the material ρ (ascertained previously to extract the sheet resistance) and the electric field E_{AC} with the carrier velocity V_{AC} to produce a figure for the carrier concentration n:

$$J_{AC} = \rho E_{AC} = n q V_{AC} \qquad \textbf{Eqn 4.1.7}$$

For bi-polar structures in which both holes and electrons contribute to the overall current flow, the process of extracting the low-field mobility of each carrier becomes problematic as the figure produced by the described method would be an average value across the distribution of the two types of carrier. Fortunately for HEMT structures the dominant charge carrier is the electron and so the contribution to current flow by holes should be insignificant with field strengths below that that breakdown of the material occurs. It is therefore a reasonable assumption that under these conditions, the VDP method is an effective method of extracting figures for the carrier concentration and low field mobility of HEMT material structures.

4.2 Parasitic and contact resistance determination

The ability to form low resistance ohmic contacts to a HEMT device becomes crucial to maximise the device performance as such contact resistances contribute to the overall parasitic resistances which restrict performance, particularly with shorter gate length systems as discussed in **Section 2.4 / 2.5**. Characterising the specific contact resistance is generally achieved through the transmission line method (TLM) [4.2], which utilises incrementally spaced ohmic contacts metallised onto the isolated material with which the contact resistance can be extracted linearly from the resistance measured between adjacent contacts. A standard TLM structure is presented in **Figure 4.2.1**:

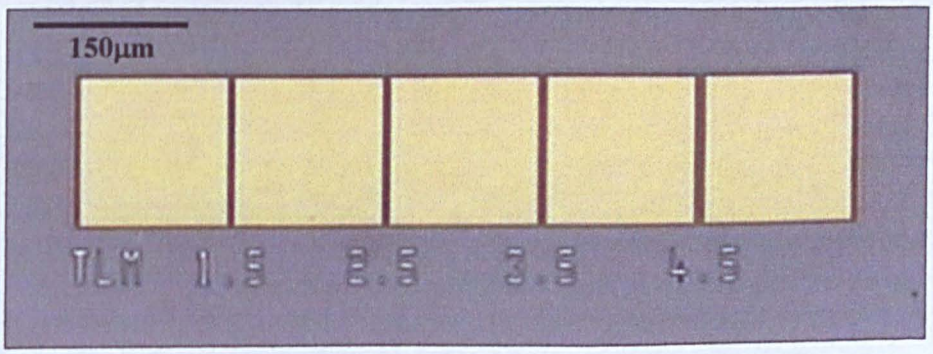


Figure 4.2.1 - Standard TLM structure with 1.5, 2.5, 3.5 and 4.5µm spacing

Initially, a rectangular area of active material is defined by isolation etching, upon which the ohmic contacts are formed at varied separations. The typical TLM structure given in **Figure 4.2.1** utilises an increment of 1µm with the smallest gap being 1.5µm. Four probe measurements are then taken by placing two probes on each adjacent contact and measuring the IV characteristics between them, (typically current is passed between two probes on separate pads while the voltage between the remaining two probes is monitored). The resistance extracted between contacts of a specific separation through linear extraction provides figures for the sheet and contact resistances for the structure, as the total resistance between pads will be the sum of the two contact resistances plus the resistance through the material.

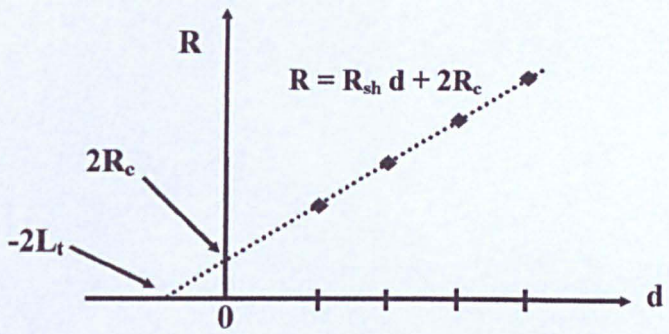


Figure 4.2.2 - TLM resistance figure extraction process

The process used to extract the contact and sheet resistances is presented in **Figure 4.2.2**. The resistances measured between each of the contact separations provides a range of data points which when plotted as resistance R against contact separation d, yields the sheet resistance as the gradient of the line, and the sum of the two contact resistances as the intercept of the y-axis, i.e. when the contact separation becomes zero.

For HEMT type material structures, this standard method can become inaccurate as parallel conduction through the cap as well as the channel distributes the current flow across a parallel resistor network. The contact resistance as extracted by the standard TLM in this instance will more likely be the resistance between the ohmic metal and the cap, as opposed to the ohmic metal and the channel which is the true resistance in a real device. This process is illustrated in **Figure 4.2.3**:

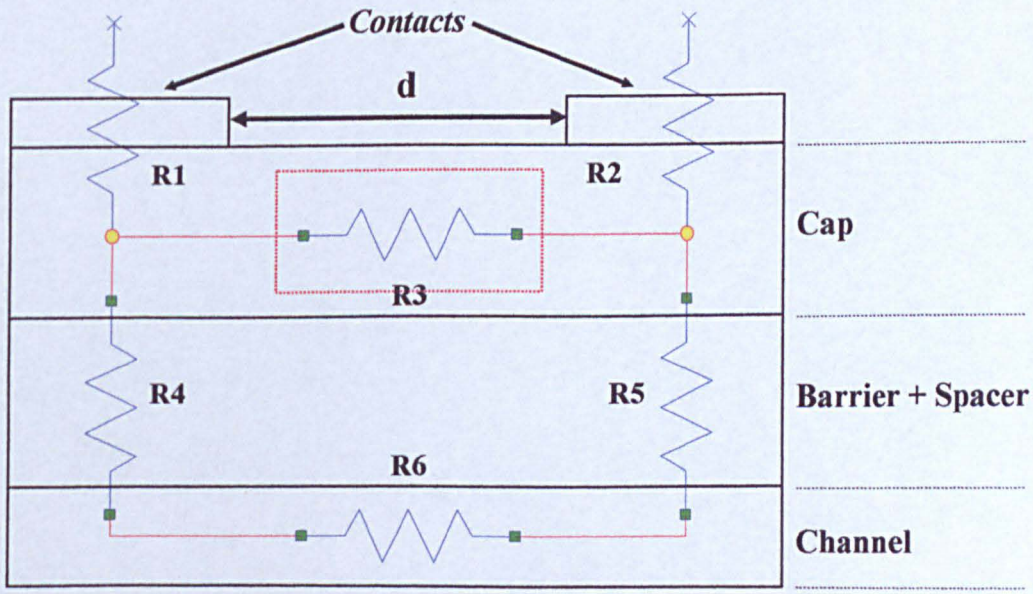


Figure 4.2.3 - The TLM HEMT resistor network

Figure 4.2.3 displays the parallel resistor network through the standard TLM structure. Each of the six resistive elements are related to active resistances within the HEMT material structure. R1 and R2 represent the resistance associated with the Schottky barrier formed at the semiconductor / metal interface, while R4 and R5 denote the resistance across the large potential barrier formed by the barrier layer between the cap and channel layers. R3 and R6 represent the horizontal resistance through the cap and channel layers respectively which will depend upon the contact separation d, as well as the sheet resistance of each of these layers. In this configuration, the magnitude of the ratio of the resistance R3 to the sum of the resistances (R4 + R5 + R6) will define the distribution of current passing between the two contacts. For real HEMT devices, the contact resistance is defined as the resistance between the metal source or drain contact and the device channel i.e. (R1 + R4) or (R2 + R5). If R3 is very large i.e. much greater than R4 + R5 + R6, the majority of current will flow through the channel as in a recessed device and the contact resistance extracted will accurately represent that in the real device. However with modern HEMT structures utilising thick highly doped cap layers, the magnitude of the resistance R3 can become much smaller than that through the channel route and the majority of current will flow through the cap. Under these conditions, the contact resistance extracted through TLM measurements will simply be that across the Schottky barrier between the metal contact and semiconductor, which although is a useful figure for analysis does not reflect the true contact resistance of a the real device.

A technique that can be adopted to extract an accurate contact resistance figure involves removing a section of the cap through etching akin to the gate recess etch in a real device. This eliminates the possibility of conduction through the cap, but also destroys the linear resistance scaling of the material with varied contact separations by introducing two separate sheet resistances, that with and without the cap. This process is described in **Figure 4.2.4**:

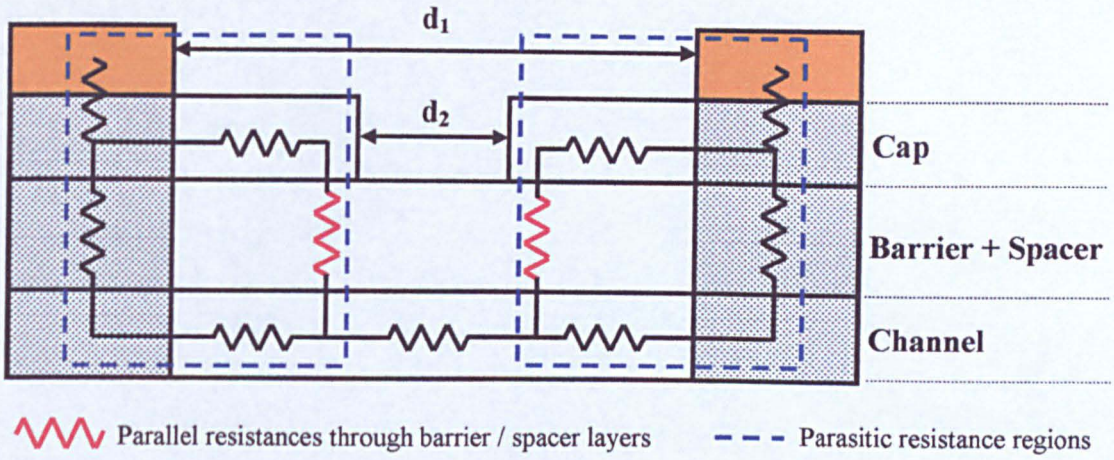


Figure 4.2.4 - The TLM HEMT resistor network with etched cap

By removing a section of the cap of length d_2, the TLM process is complicated by introducing parallel resistances through the barrier / spacer layer outwith the ohmic region beneath the contacts, denoted by the red resistance elements in **Figure 4.2.4**. The magnitude of these resistances will largely influence the overall figure extracted for the contact resistance using this method, and are defined by the potential barriers formed by the material conduction band profile. The properties of conduction through such structures have already been discussed in **Section 2.2**, but it can be noted that for GaAs and InP based HEMT structures the potential barriers formed through the barrier layer can become very large, imposing large resistances across them. This reiterates the need for ohmic contact annealing with which these barriers are minimised through metal diffusion into the material and an ohmic region is formed as highlighted by the shaded area beneath the ohmic contacts in **Figure 4.2.4**. Under these conditions the magnitude of these resistances through the barrier layer outwith the ohmic region (i.e. the red resistive element in **Figure 4.2.4**) can be taken to be sufficiently large to encourage the majority of current to flow solely through the ohmic regions and channel without any conduction through the cap. This simplifies the extraction of the contact resistance as the resistance seen between the contacts will be that defined by the two contact resistances and channel resistance alone without concern of parallel conduction

through both cap and channel layers. Under these conditions the total resistance measured between contacts will scale linearly with contact separation.

Another advantage of this particular process is the accurate extraction of the total device source and drain parasitic resistances as opposed to just the vertical contact resistances which provide far more information concerning the completed device operation. In addition to conduction vertically through the structure for the contact resistance, the resistances associated with conduction horizontally through the device layers outwith the gate region also contribute to the total parasitic resistances. The parasitic source and drain resistances are therefore defined as the total resistances between the intrinsic region of the device, i.e. the gate region, and the source and drain contacts outside the device. These parasitic regions are highlighted in blue in **Figure 4.2.4**.

If the length of section of cap that is removed from the TLM structure is scaled accordingly, these total parasitic resistances are readily extracted by verifying the dimensions of the section removed and from the known sheet resistance of the channel extracted from VDP measurements with the cap removed.

The process by which figures for the contact and total parasitic resistances were extracted throughout this work involved using the TLM test patterns as described with both the cap complete and with scaled cap removal depending on the material and devices in question. A four probe Wentworth card was used to probe the structures and using an Agilent 4145 / 4155 semiconductor analyser, four terminal IV characteristics were taken across the varied contact separations. This produced resistance figures which could then be normalised to the width of the structure. The form Ω.mm is often adopted for the quoting of contact resistances and provides a method of producing a dimension independent resistance figure. For example if the contact resistance of a 100μm wide contact is found to be 1.5Ω, this can be quoted as 0.15Ω.mm. Similarly an additional form takes into consideration the area of conduction of

the contact as opposed to just the width and is of the form of $\Omega.cm^2$. The transfer length into the contact L_t which will define the area of conduction with the contact width, is extracted through tracing the linear resistance / contact separation response to where the resistance is reduced to zero as shown in **Figure 4.2.2.** The negative value of the contact separation at this point will equal twice the transfer length, as two contacts are measured simultaneously. Again for example, if similar values are used for the contact resistance and contact width of 1.5Ω and 100μm, and the transfer length is found to be 700nm, (a typical value), the specific contact resistance can be quoted as 1.05×10^{-6} $\Omega.cm^2$.

All figures extracted for contact and parasitic resistances throughout this work are quoted in the Ω.mm format as this provides an understandable figure without the concern of calculating contact transfer lengths.

4.3 Device characterisation

The complete set-up as used for the DC and RF characterisation for all devices throughout this work is presented in **Figure 4.3.1**:

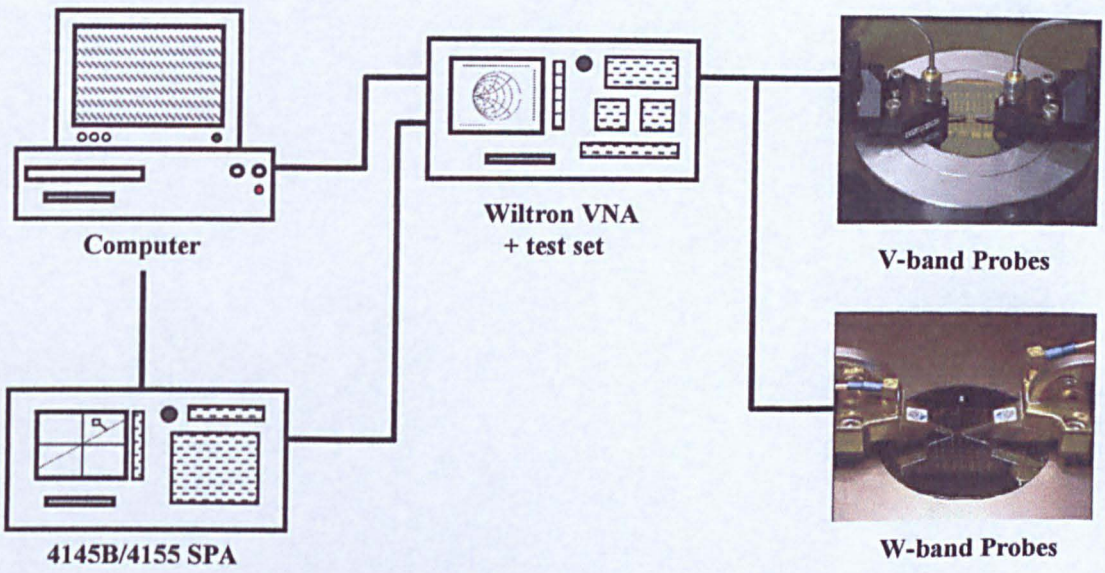


Figure 4.3.1 - DC and RF measurement set-up

Both DC and selected RF measurements from completed HEMT devices were taken using Picoprobe high frequency probes mounted upon a Karl Suss probe station as presented in **Figure 4.3.1**. An Agilent 4145B / 4155 semiconductor parameter analyser (SPA) provided a method of measuring DC characteristics of the devices via the picoprobe probes as well as measuring gate diode performance. The 4145B unit is the older model to the 4155, and hence was used for the majority of the earlier measurements presented throughout this work.

Automated extraction of multi-bias RF measurements across either V-band (24 MHz to 60GHz) or W-band (67 to 110 GHz) was achieved through PC control of the Wiltron vector network analyser (VNA) and 4145B/4155 SPA, used to bias the device. Two separate probe stations could be used for V-band and W-band measurements respectively, with the Pico probes covering DC and V-band measurements while Cascade Microtech probes were used for W-band measurements. An RF switch situated between the probe stations allowed switching between the two frequency ranges of interest while the VNA could be configured

for either range. The majority of RF measurements taken throughout this work were within the V-band range.

DC measurements

By sweeping the voltage applied to the drain of the device from zero to a positive value whilst keeping the source earthed, the drain current could be monitored with respect to this bias to produce a plot of the output characteristics of the device i.e. drain current versus source - drain bias. By then introducing a stepped negative gate bias for each drain voltage sweep, the reduction in the drain current with more negative gate bias could be observed across this source - drain voltage range. These output characteristics when plotted resemble those predicted in **Figure 2.4.2.**

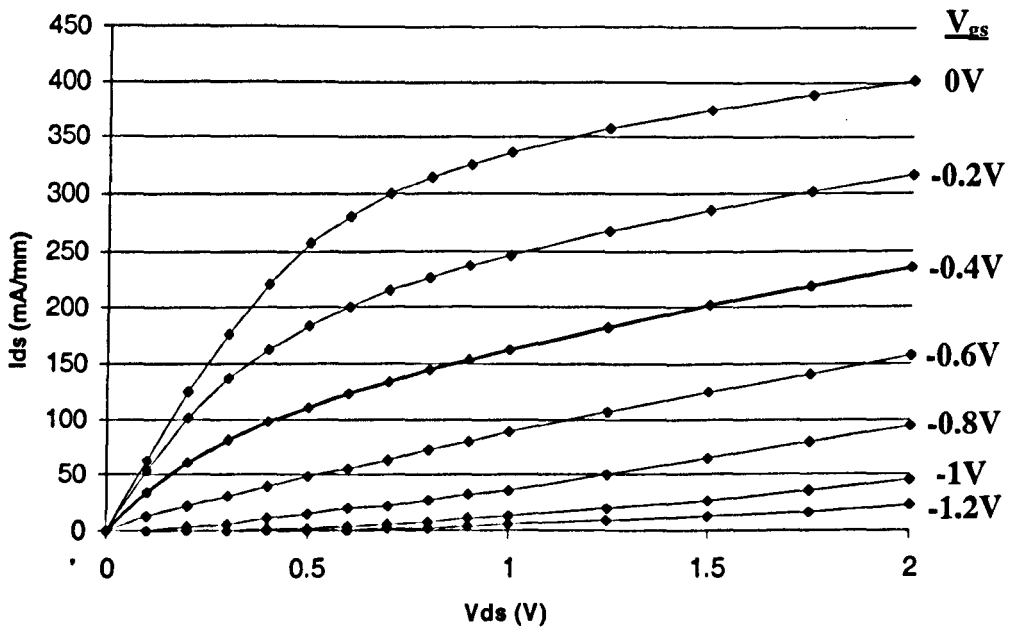


Figure 4.3.2 - 120nm GaAs pHEMT output characteristics

Figure 4.3.2 shows typical output characteristics as measured by the 4145B unit from 120nm GaAs pHEMT devices fabricated from standard GaAs pHEMT material (**Appendix BI**) and using the fabrication process given in **Appendix AI.** With zero bias applied to the gate, the source - drain bias is swept from 0 to +2V. The resultant drain current is monitored across this voltage range and exhibits the initial ohmic and then saturation response expected from

transport through the semiconductor. In this example the drain current is normalised to the width of the device and quoted in mA/mm. With a negative gate bias applied, this current is reduced across the swept source - drain bias range by mechanisms described in **Section 2.3**, until at a particular gate voltage the drain current is effectively reduced to zero across this range. In this state the device is said to be "pinched off" as previously discussed, and the voltage required on the gate to produce this effect is labelled the 'pinch off' voltage. In this example the device appears to pinch off at -1V to –1.2V V_{gs}. The magnitude of the drain current within the saturation region at zero gate voltage is also a figure of importance and is labelled the saturation current, I_{dss}. From **Figure 4.3.2**, an I_{dss} of 350 to 400 mA/mm is observed.

The DC transconductance or the rate of change of drain current with gate bias at a fixed source - drain voltage can also be extracted from device transfer characteristics:

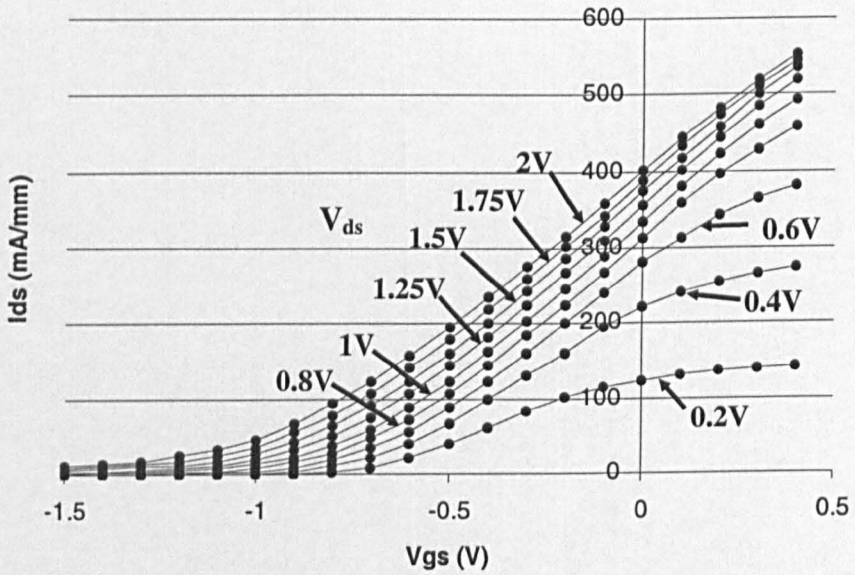


Figure 4.3.3 - 120nm GaAs pHEMT transfer characteristics

By calculating the gradient at each point of the slope for each V_{ds} sweep from **Figure 4.3.3**, the transconductance against gate bias for each source - drain bias can be plotted:

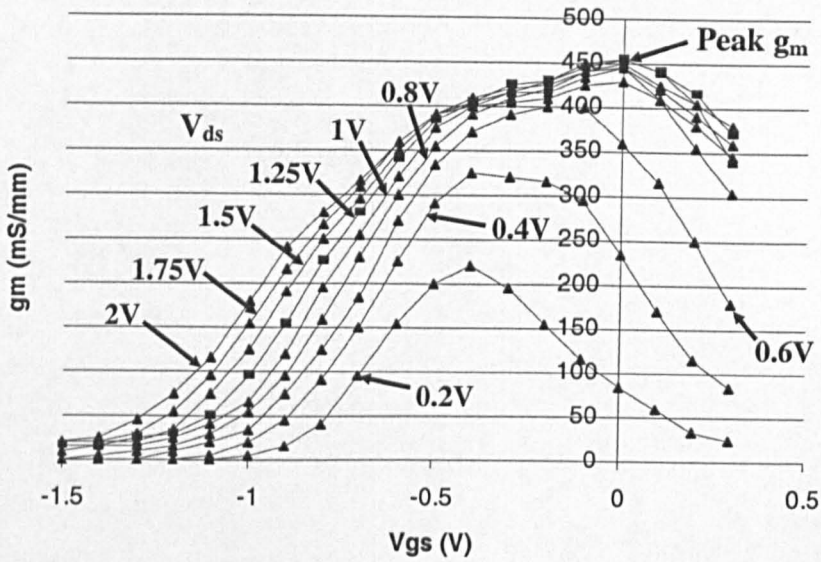


Figure 4.3.4 - 120nm GaAs pHEMT transconductance characteristics

Figure 4.3.4 shows the variation in transconductance g_m with varied gate bias and stepped source - drain voltage as calculated from the device transfer characteristics. In both **Figures 4.3.3** and **4.3.4**, the gate bias is swept from –1.5V to a positive value. This range illustrates a peak transconductance value of 450mS/mm that occurs at zero gate bias and at a source - drain bias of 1V. The position of the peak transconductance at zero gate bias combined with a substantial increase in I_{dss} with increased positive gate voltage would indicate slight over etching in the device processing. Depletion of the channel carrier concentration beneath the gate without the application of a gate bias would then result from a reduction in the gate - channel separation.

The quality of the Schottky gate contact can also be determined through DC characterisation. In an ideal HEMT device, there would exist zero current flowing between gate and channel. In reality however, with a negative applied gate bias, thermionic emission and to a lesser extent tunnelling processes will account for the magnitude of current flow in a real device observed between the gate and channel. Due to the rectifying nature of the gate Schottky contact, in reverse bias a substantial current will begin to flow through the gate contact producing diode like IV characteristics. These characteristics are readily measured by

monitoring the current through the gate contact for varied source - gate voltages, with the option of introducing a varied source-drain bias to observe the impact of these bias conditions on the gate current.

RF characterisation

Treating a HEMT as a two port device, taking the gate and drain as the two terminals and the source as earth, two port scattering parameters (S-parameters) can be measured across the specified frequency range and under various device bias conditions. S-parameter measurements provide information concerning the magnitude and phase of incident, reflected and transmitted voltage signals between the two device terminals. This process is illustrated in **Figure 4.3.5**:

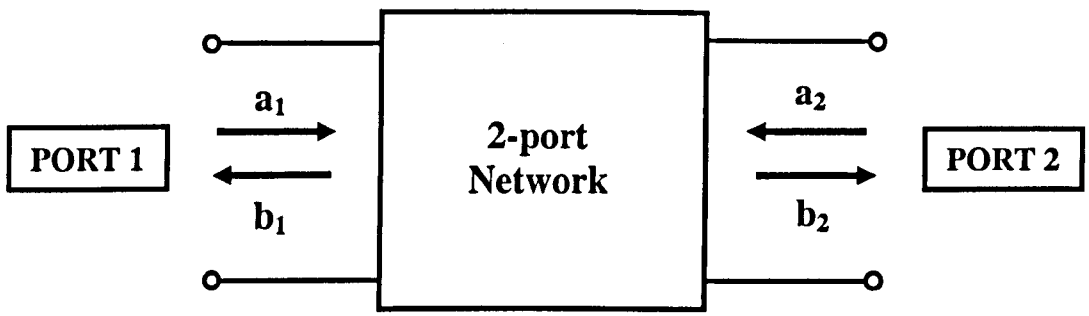


Figure 4.3.5 - 2-port S-parameter process

The four signals at the two ports are related to the corresponding S-parameters with the following expressions [4.3]:

$$b_1 = S_{11} a_1 + S_{12} a_2 \qquad \textbf{Eqn 4.3.1}$$

$$b_2 = S_{21} a_1 + S_{22} a_2 \qquad \textbf{Eqn 4.3.2}$$

Each of the four individual S-parameters can then be extracted by modifying the set-up as given in **Figure 4.3.1** to produce zero reflected signal in a particular direction, i.e. in one direction a_2 can be eliminated by creating a source at port 1, and terminating port 2 with a 50Ω load to absorb the emitted signal and ensure zero reflection. Similarly a_1 can be

neglected when the source is placed at port 2 and port 1 terminated with a 50Ω load. This yields:

$$S_{11} = \frac{b_1}{a_1} \quad \textbf{Eqn 4.3.3a} \quad \& \quad S_{21} = \frac{b_2}{a_1} \quad \textbf{Eqn 4.3.3b} \quad \textbf{(for } a_2 = 0\textbf{)}$$

$$S_{22} = \frac{b_2}{a_2} \quad \textbf{Eqn 4.3.4c} \quad \& \quad S_{12} = \frac{b_1}{a_2} \quad \textbf{Eqn 4.3.4d} \quad \textbf{(for } a_1 = 0\textbf{)}$$

S-parameters are generally favoured for network analysis due to their relative ease of measurement at microwave frequencies without the need for shorts/opens, plus their easy conversion to other significant parameters for the calculation of figures of merit. In particular, by shorting port 2 to earth, or the drain to the source, the current gain H21 can be extracted. For HEMT devices this produces an inverse logarithmic dependence with frequency that decays at a rate of 20dB/decade due to the inverse dependence of the current gain with frequency. This decay will occur until at a particular frequency, the device will demonstrate unity current gain between the gate and drain at which point the cut-off frequency of the device f_T is defined. Plotting H21 therefore provides a method of extracting the cut-off frequency for a particular device, which as discussed in **Section 2.5** is a significant figure of merit for device RF performance.

Similarly the maximum frequency f_{max} is extracted through plotting the maximum available gain (MAG) response using the following relation [4.4]:

$$\text{MAG} = \frac{S_{21}}{S_{12}}\left[K + \left(K^2 - 1\right)^{1/2}\right] \qquad \textbf{Eqn 4.3.5}$$

where K is labelled the stability factor and is defined as [4.4]:

$$K = \frac{1 + \left|S_{11}S_{22} - S_{12}S_{21}\right|^2 - \left|S_{11}\right|^2 - \left|S_{22}\right|^2}{2\left|S_{21}S_{12}\right|} \qquad \textbf{Eqn 4.3.6}$$

For $K < 1$, the device is said to be conditionally stable and in such a condition could potentially oscillate when matched to various input/output impedances. For $K > 1$, the device is described as unconditionally stable. This process is reflected in the MAG / frequency response, which can be divided into conditionally/unconditionally stable ranges of frequency.

For the conditionally stable region when K < 1, the maximum stable gain (MSG) is plotted and is defined as the magnitude of the ratio of S_{21} to S_{12} i.e. MSG = $|S_{21}/S_{12}|$. For K > 1, the MAG is then plotted as defined by the expression in **Eqn 4.3.5**. This process is demonstrated in an example HEMT MSG/MAG plot given in **Figure 4.3.6**:

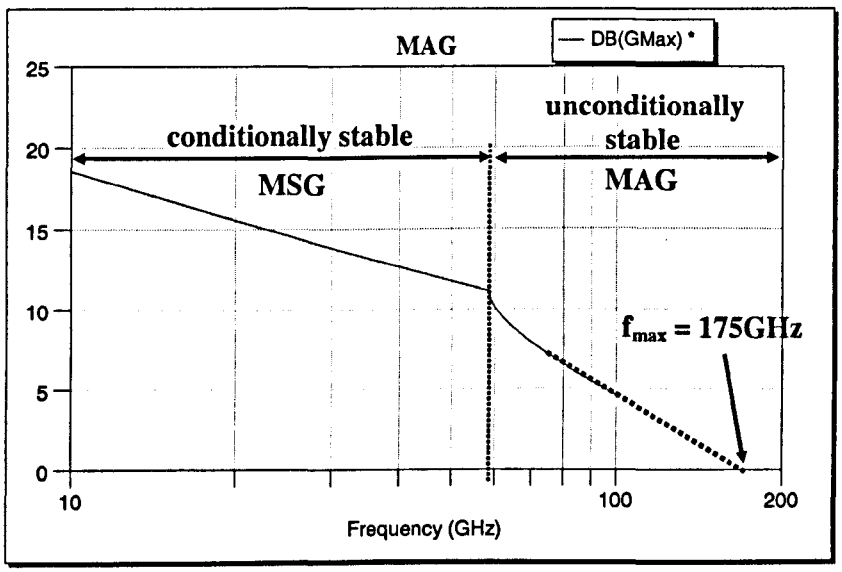


Figure 4.3.6 - MSG/MAG plot from 120nm GaAs pHEMT data

This data can be compared with the stability factor K versus frequency for the same device:

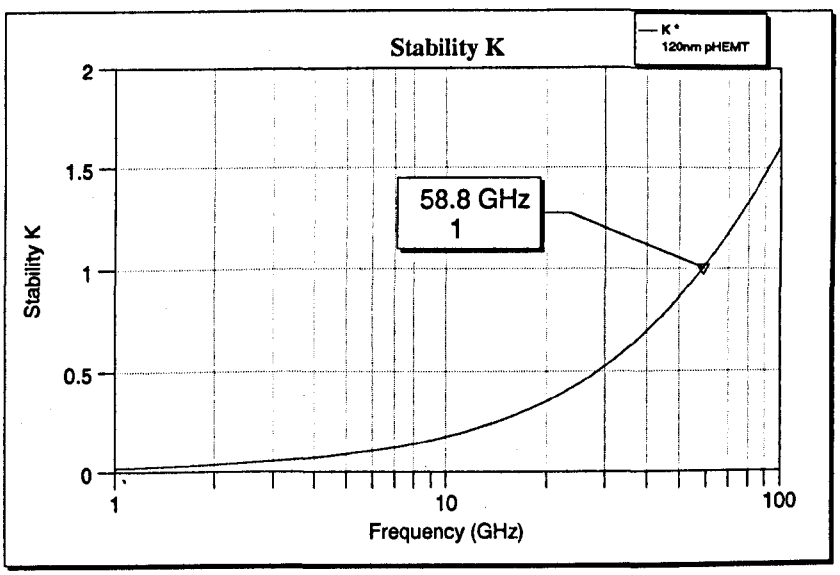
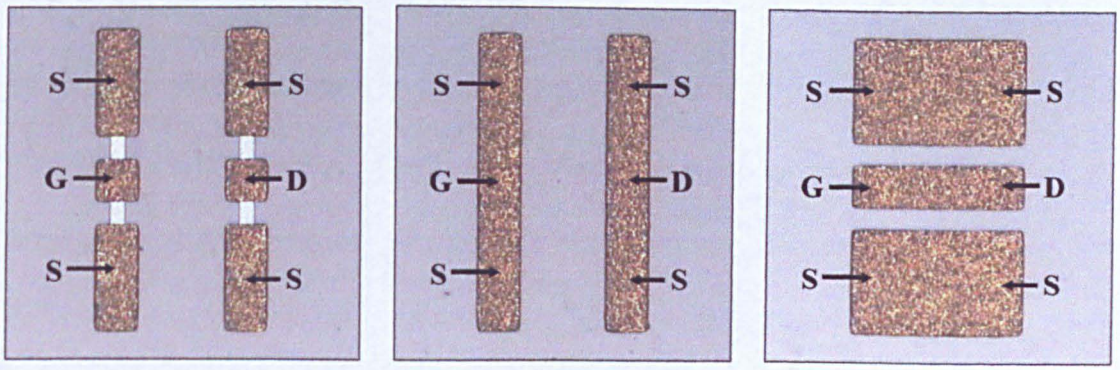


Figure 4.3.7 - Stability factor K plot from 120nm GaAs pHEMT data

The data presented in **Figures 4.3.6** and **4.3.7** was determined from an equivalent circuit that was modelled from 120nm gate length GaAs pHEMT device measurements. **Figure 4.3.6** shows the decline in MSG with frequency through the conditionally stable region until the stability factor becomes greater than 1 at a frequency of 58.8GHz as demonstrated in **Figure 4.3.7**. Beyond this frequency the device is unconditionally stable and a figure for f_{max} is determined through linear extraction of the MAG response i.e. when MAG = 0dB. This yields a figure of 175GHz for f_{max} in this example.

Vector Network Analyser Calibration

To ensure the accurate measurement of multi-bias device S-parameters, proper calibration of the measurement system is crucial. Calibration of the VNA used to perform the RF measurements, acts to de-embed the systematic errors introduced through the measurement system by the likes of the cables, probes and the system itself. Two distinct calibration techniques used for V and W band on wafer RF measurements used through this work include SOLT (short, open, line, through), and LRRM (line, reflect, reflect, match). Typically the LRRM method is acknowledged as the superior of the two for on wafer measurements due to more accurate S-parameter measurement at higher frequencies [4.5]. Both incorporate the use of various test structures to designate measurement reference planes to identify parasitic elements introduced by the measurement equipment and help exclude them from device measurements through the calibration process. These structures include a 'short', 'open', 'line' and 'through', each of which is measured in turn in either of the calibration processes. An image of each type of structure as taken from a Cascade Microtech Impedance standard substrate (ISS) is given in **Figure 4.3.8**.



a) 50Ω Load b) Short c) Through

Figure 4.3.8 - ISS standard calibration structures

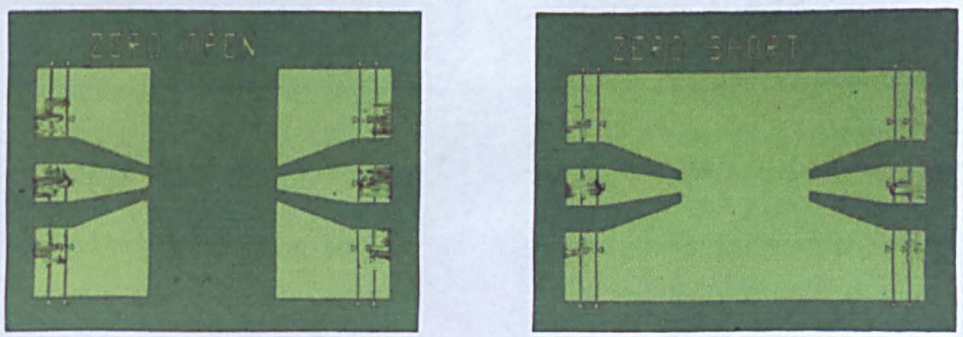
Each of these structures is fabricated precisely upon the alumina ISS substrate. For each, the probes that connect the source-gate-source from port 1, and source-drain-source from port 2, are placed as represented in each of the images in **Figure 4.3.8**. Typically when using an ISS standard for calibration, the 'open' measurement is taken simply by raising the probes from the sample surface and separating them spatially.

Using an ISS and corresponding test structures for either SOLT or LRRM calibration, the measurement reference plane is defined at the probe tips i.e. only the systematic errors and parasitic elements of the test set and probes are excluded in the measurement process. The parasitic elements that arise from signal propagation through the large metallic bond pads that feed into the actual device are therefore included in the measurements taken under these calibrations. As these will reduce the overall device performance and are not inherent to the device characteristics, but are merely a side effect of the probe pad method of measuring individual devices, their removal is desirable for the extraction of the real device performance. This can generally be achieved by one of two methods. By performing a system calibration that defines the reference plane at the probe tips and hence includes the pad parasitics within the measurement, the measured S-parameters can be used to create an equivalent device circuit that includes the effect of the probe pads. Removal of the probe pads within the circuit then allows the reconstruction of the device performance without the negative effect of the pads. The removal of these external parasitic elements is known as de-

embedding and hence defines the difference between embedded and de-embedded device results. As an alternative, if instead of using an ISS as a calibration sample, calibration structures of geometry similar to that for an actual device could be used to account for the parasitic effect of the bond pads. The S-parameters measured should then resemble more accurately those for the actual device without the pads, depending on the location of the measurement reference plane as defined by the calibration structures.

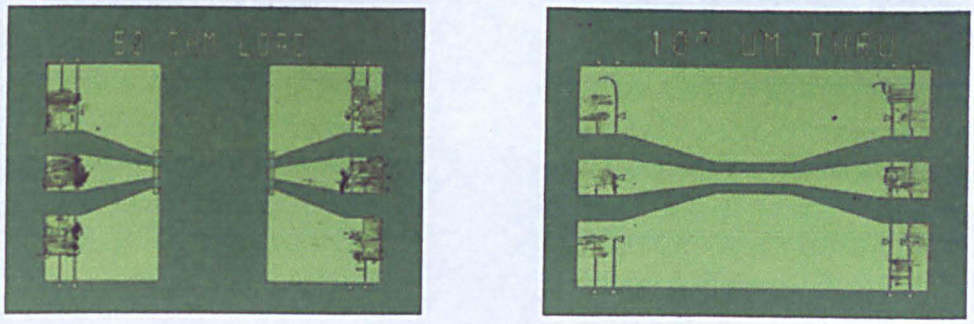
With this in mind calibration structures have been fabricated previously "in house" from a GaAs substrate for use with either SOLT or LRRM calibration processes [4.6]. These include equivalent 'short', 'open', 'line' and 'through' structures used for either calibration process. By using a GaAs substrate, the losses observed at RF from the probes pads into the substrate that occur in a real device are emulated in the calibration. Ideally, these structures would be fabricated upon the MBE grown material with the actual devices but due to the difficulty in fabricating an accurate load structure without considerable additional processing, this is impractical when realising individual HEMT devices.

Images of each type of structure from the GaAs calibration sample are presented in **Figure 4.3.9**:



a. Open **b.** Short

Figure 4.3.9 - GaAs calibration sample structures



c. 50Ω load **d.** Through

Figure 4.3.9 - GaAs calibration sample structures

The GaAs calibration sample test structures presented in **Figure 4.3.9** mirror the geometry of the probe pads defined in the real device precisely in addition to consisting of a similar metallisation of NiCr and Au. Probe placement marker lines also similar to those in the real device allow the accurate placement of probes and hence define the reference plane when using the ISS for calibration, (these can be seen to be worn by repeated calibrations). Using this GaAs calibration sample and structures, the measurement reference plane is moved from the probe tips or from the probe placement markers closer to the actual device as defined by the open structure. **Figure 4.3.10** again shows the completed HEMT structure with probe pads similar to that presented in **Figure 3.5.9**:

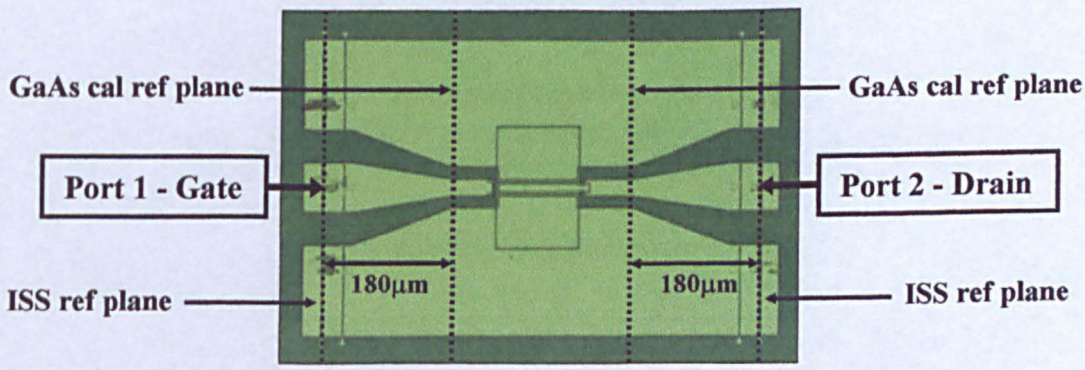


Figure 4.3.10 - Completed HEMT structure plan view

The placement of the measurement reference planes by calibration with either the ISS or GaAs calibration standards are indicated on the device structure in **Figure 4.3.10**. With the

GaAs calibration sample this is moved approximately 180μm closer to the actual device and hence de-embeds the majority of the influence of the probe pads for RF measurements.

Although this calibration process will result in the more direct, accurate extraction of the device characteristics, the quality of the 'load' test structure may not be duplicated to the accuracy of that for the ISS standard. In addition, the characteristics of the material from which the devices were fabricated can differ greatly from those for the GaAs calibration sample. This can result in different losses to the substrate from the pads depending on the quality of the material in question.

Although using the ISS yields results that are still embedded with associated probe-pad parasitics, a model of the complete device from which the effect of the probe pads can be de-embedded is easily derived. It then becomes relatively straight forward to remove the effects of these pads from the model and thereby reconstruct the performance from the device alone.

4.4 Chapter Summary

The various methods used throughout this project for the characterisation of material and completed devices have been presented in this chapter. The importance of the accuracy of such techniques has also been stressed. This becomes evident for the likes of RF device characterisation from which device results can differ greatly depending upon the calibration process used. Accurate extraction of the device performance then relies on the ability to identify and remove external elements to the device which itself can become a complicated task.

Having provided a background into the operation, fabrication and testing of modern III-V HEMT devices, a review of current technology is presented in the following chapter to provide an idea into the current level of the technology in this field.

References

[4.1] V. D. Pauw, *A method for measuring specific resistivity and hall effect of discs of arbitrary shape.* Philips Res. Rep., 1958. 13: p. 1-9.

[4.2] G. K. Reeves, H. B. Harrison, *Obtaining the Specific Contact Resistance from Transmission Line Model Measurements.* IEEE Electron Device Letters, 1982. 3(5): p. 111.

[4.3] D. Ballo, *Network analyser basics seminar.* 1998, Hewlett Packard: Santa Rosa.

[4.4] S. M. Sze, *Physics of Semiconductor Devices*, 1981, John Wiley and Sons Inc. p. 162.

[4.5] M. Nishimoto, M. Hamai, J. Laskar, R. Lai, *On-wafer calibration techniques and applications at V-band.* Microwave and Guided Wave Letters, IEEE, 1994. 4(11): p. 370-372.

[4.6] K. Elgaid, *A Ka-Band GaAs MESFET Monolithic Downconverter*, in *Electronic and Electrical Engineering*. Thesis, 1998, The University of Glasgow: Glasgow.

Chapter 5

Current Technology Review

Throughout this chapter the various HEMT technologies that have been developed and resultant device performance quoted by various groups that have bearing on the research presented in this work are discussed. This provides a route to the comparison and benchmarking of the developed processes and device technology presented in following chapters.

Generally the methodology adopted for maximising HEMT device performance as has been previously stressed lies with the minimisation of the device gate length, and hence the reduction of the associated device gate capacitance. Various publications chart the incremental process development in reducing the gate length of HEMT devices through various lithographical processes. At present the minimum gate length incorporated into a HEMT process stands at 25nm developed by Y. Yamashita *et al* [5.1]. By combining this ultra-short gate length with a pseudomorphic InP based material structure incorporating a 70% indium InGaAs channel to maximise carrier velocity beneath the gate, an f_T of 562GHz is extracted which at the time of writing is the highest transistor cut-off frequency figure ever reported. A DC transconductance of 1230mS/mm and an f_{max} of 330GHz were also extracted for these devices. Beyond an ultra-short gate length and a pseudomorphic InP system, Y. Yamashita *et al* attribute such high frequency performance to better scaling of the system through the reduction of the channel to gate distance over previous reported results [5.2].

Although these HEMT type devices at present claim the title of fastest three terminal devices in the world, progress in HBT technology, particularly in InP has led to a significant increase in high frequency device performance. Traditionally HBT technology could not compete with HEMT high frequency performance due to the large base resistance inherent to HBT design. However recent publications have quoted figures of 370GHz and 375GHz for f_T and f_{max} respectively for InP HBT devices [5.3], rivalling some of the highest frequency performance figures achieved from HEMT technology to date.

Even though the majority of HEMT performance research has been focussed on lattice matched and pseudomorphic InP systems due to their obvious performance advantages, considerable research has been devoted to the fabrication of high performance GaAs metamorphic devices. As mentioned in **Section 2.6**, the metamorphic systems utilises an indium graded InAlAs buffer layer grown upon a GaAs substrate to allow the subsequent growth of high indium content layers with minimal strain or dislocation through such active layers. The advantages of such a system from a financial point of view include the ability to

produce near InP type device performance using GaAs without the cost or the fragility of InP [5.4]. In addition, the metamorphic process provides an element of freedom to the range of indium concentrations that can be chosen for the device channel depending on the device application i.e. power or frequency performance. This can only be achieved with an InP system by straining the structure creating pseudomorphic structures. In the field of metamorphic HEMT technology, considerable research has been performed by Y.Cordier and S.Bollaert *et al* into the development and comparison of GaAs metamorphic devices. In [5.5], they demonstrate the diversity of the metamorphic system through the comparison of varied indium content 100nm gate length devices from 30 to 52%. It is concluded that the intermediate indium concentration of 40% would prove desirable for high frequency and power applications where GaAs pHEMT and lattice matched InP devices would not suffice due to the limited frequency response of GaAs pHEMTs and the inferior breakdown characteristics of lattice matched InP HEMTs. In a parallel publication they discuss the fabrication and characterisation of 60nm gate-length metamorphic devices which exhibit a DC transconductance of 850mS/mm and an f_T and f_{max} of 260GHz and 490GHz respectively [5.6]. By increasing the indium content up to 80%, A.Tessman *et al* have fabricated metamorphic devices of 70nm gate length with f_T in the range 290GHz, f_{max} of 340GHz and a transconductance of 1450mS/mm [5.7]. These can be compared with the performance of 80nm lattice matched InP devices published by M. Riaziat *et al* who quote an f_T of 260GHz and an f_{max} of 310GHz [5.8]. At the time of writing, the fastest GaAs metamorphic based HEMT to have been reported is claimed by X.Cao *et al* within the Nanoelectronics Research Centre at the University of Glasgow. By using a digital recess wet etch technology, 50nm gate length 53% indium metamorphic HEMTS with an f_T of 350GHz and a DC transconductance in the range 1500 mS/mm have been fabricated [5.9].

Beyond the reduction of the device gate length through more complex lithographical processes or the use of high indium content material systems, little research would appear to have been focussed towards the reduction of associated parasitic resistances, which as has

been highlighted in **Section 2.5**, play an ever increasing role in sub-100nm gate length devices. Without due attention paid to the minimisation of such parasitics, the true performance of these ultra-short gate length systems cannot be realised. For a standard HEMT structure, the total internal parasitic resistances will arise from both the contact resistances and the resistance through the device channel between the ohmic contact and gate regions. Considerable research has been performed into the development of efficient ohmic contact formation as is discussed in **Section 3.5**. However without a reduction in the lateral access resistance through the material, the total parasitic resistances will not scale with device gate length and hence begin to dominate for short gate lengths. The self-aligned gate process as described in **Section 3.6** provides a method of reducing these lateral parasitic resistances by reducing the physical separation between the source and drain ohmic contacts which is generally limited by lithography alignment limitations with the standard process.

Considerable research has been performed by U.K.Mishra & L.D.Nguyen *et al* into the area of self-aligned gate III-V HEMT technology and results across a range of varied gate length and material composition devices implementing such technology have been reported. An initial paper published in 1989 first describes their self-aligned gate process developed for lattice matched InP devices of 150nm gate length [5.10]. By developing a low annealing temperature ohmic contact process, degradation of the gate contact is minimised. In addition, the formation of a SiO_2 layer encompassing the gate prior to ohmic definition ensures protection of the exposed material during oxygen ashing of the ohmic level and separates the ohmic and gate metals to avoid shorting of the two. The evolution of this technology continues with the implementation of the self-aligned process into an 80nm gate length process flow as reported by L.D.Nguyen *et al* [5.11]. Again a lattice matched InP material structure is used to realise the 80nm self-aligned devices which are reported to exhibit a peak DC transconductance of 1150mS/mm and a cut-off frequency in the range 250GHz. The benefits of the adoption of a self-aligned gate process are discussed extensively in this publication, including a prediction of the performance increase associated with a self-aligned system over the conventional for devices of gate length down to 50nm, as extracted from the

scaling of equivalent circuit parameters. These figures are verified from results extracted from 50nm self-aligned gate devices fabricated with an 80% indium pseudomorphic material system again reported by L.D.Nguyen *et al* [5.12]. An extremely high DC transconductance figure of 1740mS/mm is reported in addition to an f_T of close to 350GHz. This figure is compared with that extracted for lattice matched InP (53% indium) self-aligned devices at the 50nm node of f_T =295GHz and again demonstrates a substantial increase in device performance associated with the pseudomorphic system through a 15% increase between these two cut-off frequencies.

Beyond the extensive research performed by L.D.Nguyen *et al* into the potential of adopting a self-aligned gate methodology into a III-V HEMT regime, little appears to have been reported into the investigation of such technology, particularly at performance competitive gate-lengths, i.e. 100nm and below.

References

[5.1] Y. Yamashita, A. Endoh, K. Shinohara, K. Hikosaka, T. Matsui, S. Hiyamizu, T. Mimura, *Pseudomorphic In/sub 0.52/Al/sub 0.48/As/In/sub 0.7/Ga/sub 0.3/As HEMTs with an ultrahigh f/sub T/ of 562 GHz.* Electron Device Letters, IEEE, 2002. 23(10): p. 573-575.

[5.2] Y. Yamashita, K. Shinohara, A. Endoh, K. Hikosaka, T. Matsui, T. Mimura, S. Hiyamizu, *Extremely High-Speed Lattice-Matched InGaAs/InAlAs High Electron Mobility Transistors with 472 GHz Cutoff Frequency.* Japanese Journal of Applied Physics, Part 2, 2002. 41(4B): p. L437-L439.

[5.3] M. Rodwell, *InP Bipolar Transistors:High Speed Circuits and Manufacturable Submicron Fabrication Processes.* in *11th GAAS Symposium.* 2003. Munich.

[5.4] K. Zanden, M. Behet, G. Borghs, *Comparison of metamorphic InGaAs-InAlAs HEMT's on GaAs with InP based LM HEMT's.* in *GaAs MANTECH.* 1999.

[5.5] Y. Cordier, M. Zaknoune, S. Bollaert, A. Cappy, *Charge control and electron transport properties in In/sub y/Al/sub 1-y/As/In/sub x/Ga/sub 1-x/As metamorphic HEMTs: effect of indium content.* in *2000 International Conference on Indium Phosphide and Related Materials.* 2000.

[5.6] S. Bollaert, Y. Cordier, M. Zaknoune, H. Happy, S. Lepilliet, A. Cappy, *0.06 mu m gate length metamorphic In/sub 0.52/Al/sub 0.48/As/In/sub 0.53/Ga/sub 0.47/As HEMTs on GaAs with high f/sub T/ and f/sub MAX/.* in *2001 International Conference on Indium Phosphide and Related Materials.* 2001.

[5.7] A. Tessmann, C. Leuther, C. Schwoerer, H. Massler, S. Kudszus, W. Reinert, H. Schlechtweg, *A coplanar 94 GHz low-noise amplifier MMIC using 0.07 /spl mu/m metamorphic cascode HEMTs.* in *Microwave Symposium Digest, 2003 IEEE MTT-S International.* 2003.

[5.8] M. Riaziat, C. Nishimoto, S. Silverman, Y. C. Pao, S. L. Weng, M. Glenn, S. Bandy, R. Majidi-Ahy, G. Zdasiuk, *Highest current gain cutoff frequency with 0.08 μm gate HEMT on InP.* in *Indium Phosphide and Related Materials, 1990. Second International Conference.* 1990.

[5.9] X. Cao, I. Thayne, S. Thoms, M. Holland, C. Stanley. *High Performance 50nm T-gate In0.52Al0.48As/In0.53Ga0.47As Metamorphic High Electron Mobility Transistors.* in *ESSDERC 2003.* 2003. Estoril, Portugal.

[5.10] U. K. Mishra, A.S. Brown, L.M. Jelloian, M. Thompson, L.D. Nguyen, S.E. Rosenbaum, *Novel high performance self-aligned 0.15 micron long T-gate AlInAs-GaInAs HEMTs.* in *Electron Devices Meeting, 1989. Technical Digest., International.* 1989.

[5.11] L. D. Nguyen, L.M. Jelloian, M. Thompson, M. Lui, *Fabrication of a 80 nm self-aligned T-gate AlInAs/GaInAs HEMT.* in *Electron Devices Meeting, 1990. Technical Digest., International.* 1990.

[5.12] L. D. Nguyen, A.S. Brown; M. A. Thompson, L. M. Jelloian, *50-nm self-aligned-gate pseudomorphic AlInAs/GaInAs high electron mobility transistors.* Electron Devices, IEEE Transactions on, 1992. 39(9): p. 2007-2014.

Chapter 6

Results I
Self-aligned GaAs pHEMT

The benefits of reducing the parasitic resistances that are inherent within a HEMT device have been stressed repeatedly throughout this work. The extent to which external research has touched upon this subject, and in particular the self-aligned gate process, has also been discussed. The majority of mainstream HEMT performance research would appear to have been focussed upon the minimisation of device gate length and the use of higher indium content material structures to increase device performance. Little research has therefore been devoted to the development of a self-aligned gate process which becomes crucial to the realisation of the true potential performance of ultra-short gate length devices.

With this in mind, the challenge to develop a self-aligned process for GaAs pHEMT devices was undertaken. This was initially approached by experimenting with various ohmic contact strategies, varying the ohmic metals and annealing processes used, and using slightly differing material structures, in an attempt to develop a low temperature thin (~100nm) ohmic contact. Having discovered a viable ohmic contact technology compatible with a self-aligned gate process, a suitable recess etch for the necessary material structures was investigated. This led to the fabrication of self-aligned 120nm T-gate GaAs pHEMT devices using the newly devised processes.

6.1 Ohmic contact process development

Four GaAs pHEMT type material structures were used for the development of a thin, low temperature ohmic process. These structures were virtually identical comprising of a GaAs substrate and subsequent MBE grown layers, with the exception that the cap material / doping profile varied for each. The layer structure for each used is presented in **Appendix B**. The composition of the cap layer with reference to each structure is given in **Figure 6.1.1**:

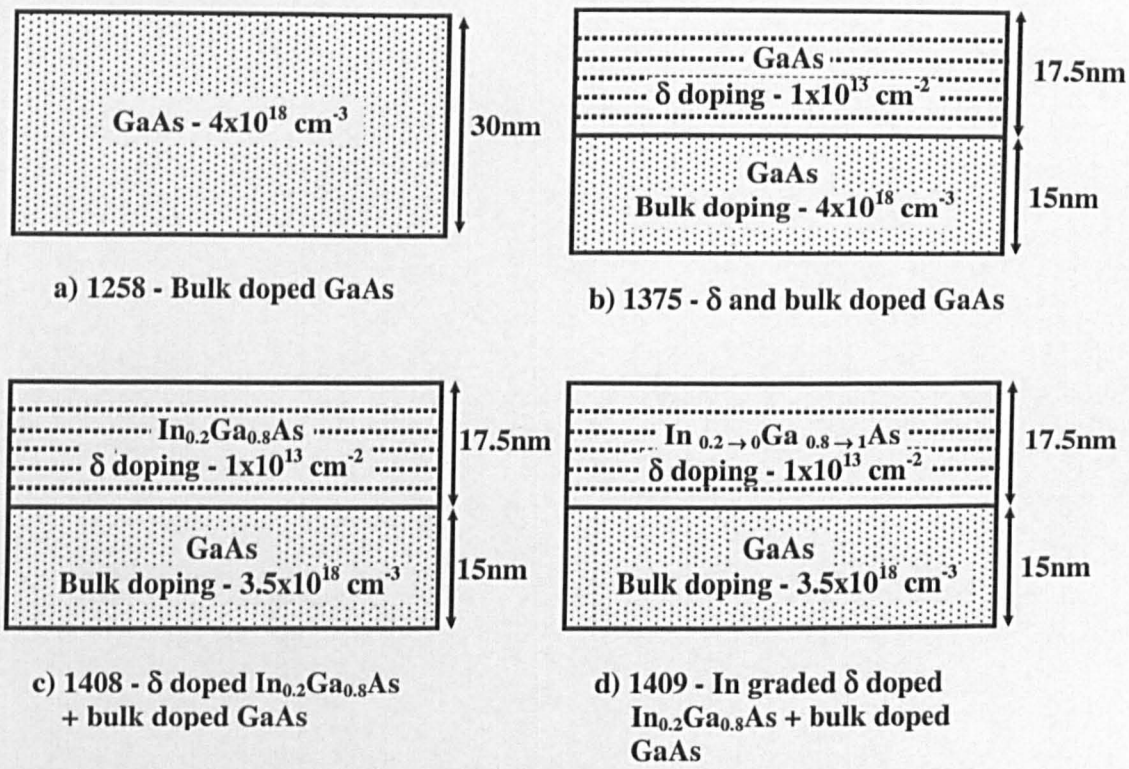


Figure 6.1.1 - Cap composition for GaAs pHEMT material structures, a) 1258, b) 1375, c) 1408, d) 1409

The four material structures are labelled 1258, 1375, 1408 and 1409 as their MBE growth reference numbers. 1258 takes the form of a standard pHEMT structure with a bulk doped GaAs cap. 1375 differs slightly from 1258 by dividing the cap into bulk and delta doped regions, with the top half of the cap incorporating 5 layers of delta doping with a period of 2.5nm. The 1408 cap structure is similar to that for 1375, except that the GaAs delta doped layer is replaced with 20% indium InGaAs. The doping profiles for the two remain almost identical apart from a small reduction in the concentration of the bulk doping of the GaAs.

Finally the cap of 1409 exhibits a similar structure to 1408, but instead the indium content through the InGaAs is graded from 20% to 0% from the surface of the material to the bulk doped GaAs.

Modifying standard ohmic recipes, three thin ohmic metallisations were devised for testing with the four material structures, namely:

Ohmic 1. 10nm Ni, 50nm Ge, 80nm Au

Total metal height 140nm.

Ohmic 2. 20nm Ge, 30nm Pd, 30nm Au

Total metal height 80nm

Ohmic 3. 14nm Au, 14nm Ge, 14nm Au, 11nm Ni, 50nm Au

Total metal height 103nm

These three recipes were to act as templates from which modification could be implemented from TLM results.

Following the metallisation of each of the three ohmic recipes onto each of the material structures, various annealing strategies were implemented to produce a range of resistance figures across a spread of annealing temperatures. These included three temperatures / time ranges for each metallisation on each material:

1. 240°C for 20s followed by 260 °C for 30s.
2. 280°C for 20s followed by 300 °C for 30s.
3. 300°C for 20s followed by 360 °C for 30s.

Contact resistances for each were then extracted using the standard TLM using typical ebeam defined TLM test structures such as those also described in **Section 4.2**. Results for each metallisation are summarised in **Figure 6.1.2**. All figures are quoted in Ω.mm.

Ohmic 1	Non-annealed	240/260 °C	280/300 °C	300/360 °C
A1258	Non-ohmic	Non-ohmic	Non-ohmic	*0.42 ± 0.04
A1375	Non-ohmic	0.67 ± 0.04	0.42 ± 0.03	*0.22 ± 0.02
A1408	Non-ohmic	0.33 ± 0.03	0.26 ± 0.03	*0.22 ± 0.02
A1409	Non-ohmic	0.26 ± 0.03	0.34 ± 0.02	*NA

*** Bad surface morphology**

Figure 6.1.2a - Ohmic contact resistances for **Ohmic 1** recipe with each material structure and varied annealing processes

Ohmic 2	Non-annealed	240/260 °C	280/300 °C	300/360 °C
A1258	Non-ohmic	Non-ohmic	Non-ohmic	0.52 ± 0.03
A1375	Non-ohmic	1.21 ± 0.07	Non-ohmic	1.15 ± 0.04
A1408	Non-ohmic	0.16 ± 0.02	0.42 ± 0.02	0.3 ± 0.03
A1409	Non-ohmic	0.24 ± 0.02	Non-ohmic	1.12 ± 0.05

Figure 6.1.2b - Ohmic contact resistances for **Ohmic 2** recipe with each material structure and varied annealing processes

Ohmic 3	Non-annealed	240/260 °C	280/300 °C	300/360 °C
A1258	Non-ohmic	Non-ohmic	Non-ohmic	0.23 ± 0.02
A1375	0.22 ± 0.02	1.13 ± 0.05	1.66 ± 0.07	0.36 ± 0.04
A1408	0.12 ± 0.02	0.78 ± 0.04	Non-ohmic	0.22 ± 0.02
A1409	0.14 ± 0.02	-0.15 (?)	0.42 ± 0.04	0.19 ± 0.02

Figure 6.1.2c - Ohmic contact resistances for **Ohmic 3** recipe with each material structure and varied annealing processes

Results with the Ohmic 1 recipe on the four material structures proved disappointing with non-ohmic or high resistance behaviour observed at lower annealing temperatures, and bad surface morphology at higher temperatures across all four materials. An image of the degradation in the Ohmic 1 morphology with higher temperatures is presented in **Figure 6.1.3**:

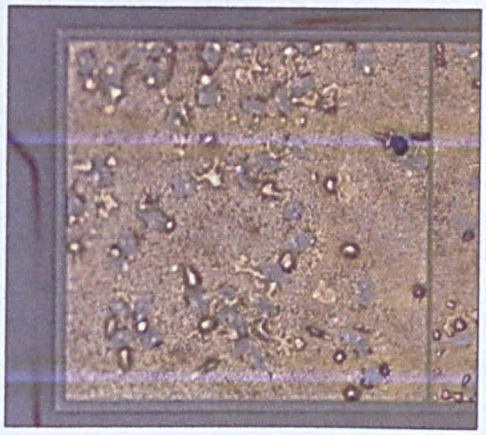


Figure 6.1.3 - Degradation in surface morphology of Ohmic 1 following annealing at 360°C

Figure 6.1.3 demonstrates the deterioration of the surface of the Ohmic 1 metallisation at higher temperatures as globules and patches appear to form, most likely due to the formation alloys from the ohmic metals and material.

At lower annealing temperatures, the Ohmic 2 metallisation appeared more promising with a low contact resistance of 0.16 Ω.mm occurring on A1408 at 260°C. This proved to be unstable however as repeated tests produced variable results. Higher temperatures also yielded a much higher resistance while non-annealed contacts remained non-ohmic.

Initially upon testing of the Ohmic 3 metallisation, the results began to look promising at higher annealing temperatures, while at the lower temperatures required for a self-aligned process the figures remained high if not erratic. However upon testing of the non-annealed Ohmic 3 contact with 1408 and 1409 materials, it was discovered that a low contact resistance of 0.12 - 0.14 Ω.mm could be attained. These figures were also found to be extremely repeatable and did not degrade with exposure to temperatures associated with resist baking.

The difference in contact resistance figures seen with each structure and with the Ohmic 3 metallisation can be explained by considering the magnitude of the Schottky barrier in each case. With the GaAs cap structures 1258 and 1375, only with increased doping in the cap with 1375 does the Schottky barrier become transparent enough to exhibit an ohmic response. Similarly with 1408 and 1409, the high doping concentration in the cap leads to a

thinner barrier in addition to the height of the barrier being reduced by using $In_{0.2}Ga_{0.8}As$ as opposed to GaAs, leading to a lower resistance than with the GaAs cap structures..

To ensure the ability to fabricate a self-aligned gate structure, a test sample incorporating an array of 120nm gate length T-gates was fabricated and the thin non-annealed Ohmic 3 metallisation deposited across it. A cross section SEM image of such a structure is given in **Figure 6.1.4**.

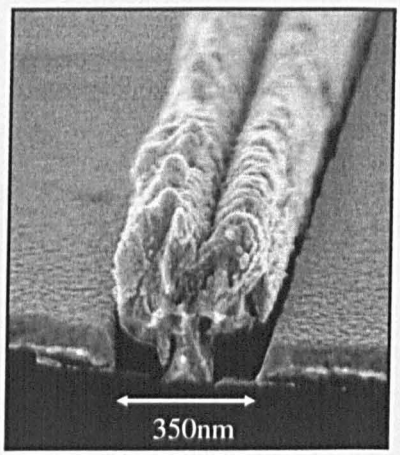


Figure 6.1.4 - SEM cross section of 120nm T-gate and thin Ohmic 3 metallisation

As demonstrated by the image presented in **Figure 6.1.4**, a uniform discontinuity exists along the width of the structure between the gate head and ohmic metallisation ensuring no shorting between the gate and potential source and drain contacts. This was further verified with electrical measurements across the array which demonstrated no shorting across a substantial amount of the test structures.

With the ability to form thin non-annealed ohmic contacts that exhibit a low enough contact resistance to be justifiably incorporated into a HEMT process, the main hurdle that obstructed the development of a self-aligned gate process appeared to have been surmounted. Effort was then focussed into the incorporation of the ohmic process into a self-aligned process flow, and in particular the development of a suitable gate recess etch process.

6.2 The Selective Gate Recess Process

Having developed a viable ohmic contact process dependent on using 1408 or 1409 material structures, dry etching of the cap for the gate recess etch could not be performed due to the lack of suitable low damage etch processes available for InGaAs. This led to an investigation into alternative wet etch processes that ideally would selectively etch InGaAs/GaAs over AlGaAs, i.e. the cap layer over the barrier layer in the material structure. A solution was found in the form of a pH balanced succinic acid based wet etch which is commonly used in gate recess processes in InP based HEMT fabrication [6.1]. Initial experimentation with the etch and various 1408 test samples incorporating standard 120nm T-gate resist profiles indicated that the etch was extremely selective, removing the InGaAs/GaAs cap and etching laterally once reaching the AlGaAs barrier layer. SEM cross section images verified that both the InGaAs and GaAs cap layers from the 1408 material were being removed, as the etch was stopped by the AlGaAs barrier layer. An SEM cross section image of a 120nm gate profile upon the 1408 material, and following an initial succinic acid etch test is presented in **Figure 6.2.1**:

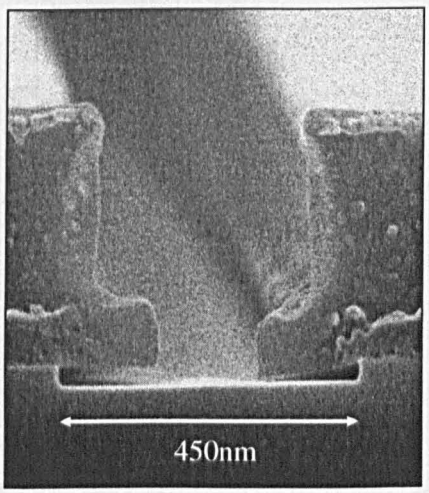


Figure 6.2.1 - High concentration succinic recess etch on 1408 with 120nm T-gate profile

The selectivity of the succinic acid etch with the 1408 structure is clearly demonstrated in **Figure 6.2.1**. The InGaAs and GaAs layers are completely removed under the gate resist profile, exposing the AlGaAs barrier layer before the etch moves laterally, continuing to etch

the cap. In this example a recess etch length of ~150nm is achieved with a 120nm gate profile. The magnitude of the etch has also resulted in the collapse of part of the resist profile. For this particular gate length a recess length of 25 - 30nm is desirable, and hence the process needed to adjusted to achieve this figure. Due to the aggressive nature of the etch, it would have been difficult to develop a repeatable process producing a desirable recess length simply by adjusting the duration of the etch alone, (again considering **Figure 6.2.1**, the structure was only etched for 30s). The etch solution was therefore diluted with water to half its original concentration. This reduced the etch rate significantly, and allowed the development of a repeatable etch resulting in a desirable recess length. A SEM image of a 120nm gate profile with a more desirable sized recess upon 1408 material is given in **Figure 6.2.2**:

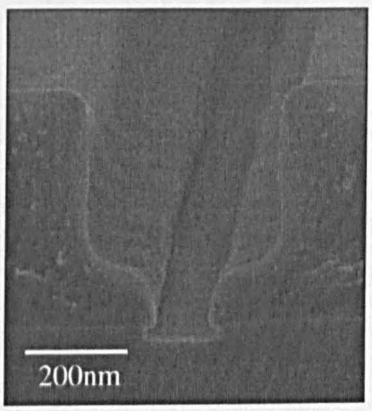


Figure 6.2.2 - Optimised succinic recess etch on 1408 with 120nm T-gate profile

The dimensions and selectivity of the diluted etch were then verified with AFM measurements by scanning the sample surface after removal of the gate resist. This confirmed the depth of the etch as that of the cap and also indicated a route mean square surface roughness of ~ 0.6nm within the etched area. A 3D AFM image of a scaled recess on the 1408 material is presented in **Figure 6.2.3**:

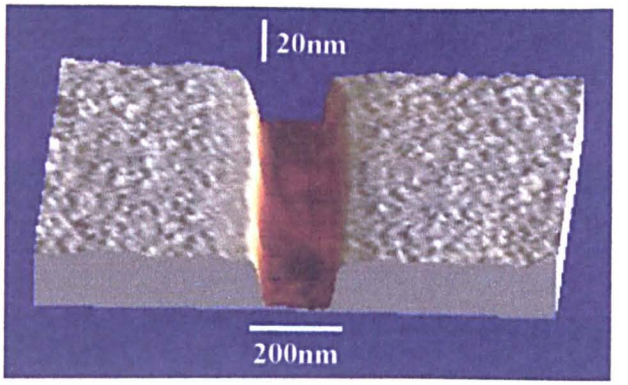


Figure 6.2.3 - 3D AFM image of optimised succinic recess etch on 1408

The repeatability of the etch was also verified across a range of similar etches and was found to be extremely robust and independent of temperature which only varied by a few degrees from room temperature with each etch.

Following the succinic acid recess and pre-metallisation, devices were treated with a hydrochloric acid based de-oxidation treatment to encourage better Schottky contact formation by removing any oxidised residue on the barrier layer.

6.3 Device Results

By adapting the standard GaAs pHEMT process by incorporating developed thin non-annealed ohmic and succinic recess processes, self-aligned 120nm gate length devices were fabricated using the 1408 material. The details of the 1408 self-aligned process flow are given in **Appendix AII**. Completed devices were characterised at DC and RF by processes discussed previously. The normalised DC output characteristics for a 25μm wide self-aligned device are presented in **Figure 6.3.1**:

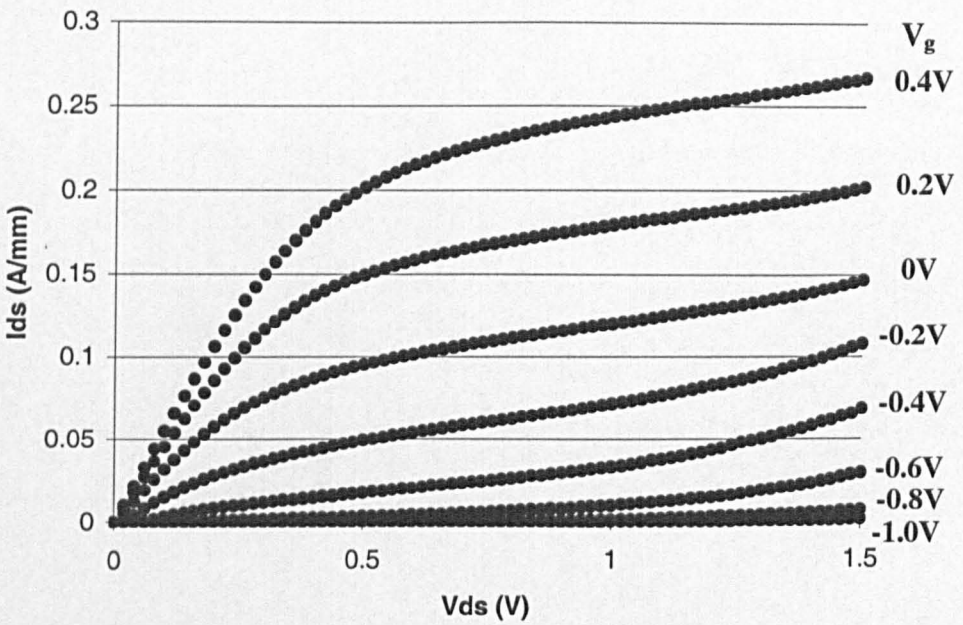


Figure 6.3.1 - Normalised output IV measurements from self-aligned 1408 device

Figure 6.3.1 displays the source-drain current through the device across a range of drain and gate voltages which indeed portrays the trend associated with typical HEMT output characteristics and were typical across the range of fabricated self-aligned devices. The drain voltage is swept from 0 to 1.5V (beyond which the initial effects of breakdown were observed), while the gate bias is stepped from +0.4V to –1V. Pinch off of the drain current is observed at –0.8V on the gate, i.e. when the current is effectively reduced to zero. The drain saturation current, I_{dss}, which is quoted to be the magnitude of the drain current within the saturated region with zero bias on the gate, is found to be 120mA/mm as normalised to the device width. By applying a positive gate voltage, the drain current was found to increase

considerably and indeed double the value for I_{dss} with only +0.4V on the gate, indicating perhaps too deep a recess etch resulting in significant depletion of the channel at zero gate bias.

In addition to extraction of the output device characteristics, the transfer and transconductance response of the devices were measured:

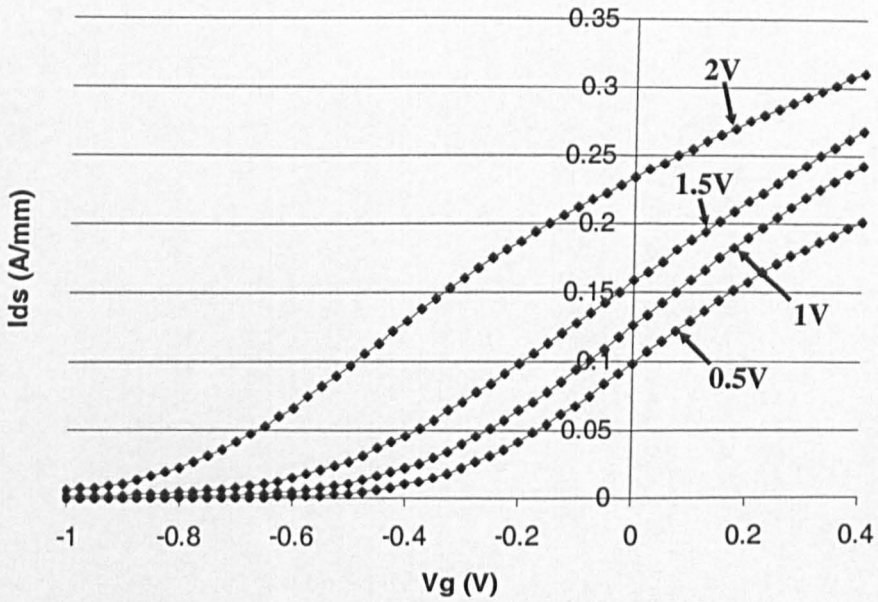


Figure 6.3.2 - Normalised transfer IV measurements from self-aligned 1408 device

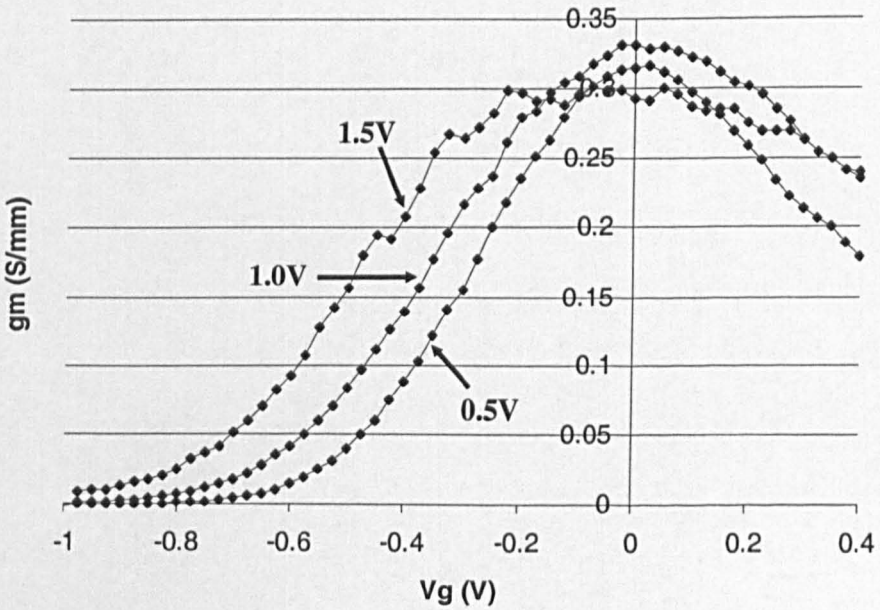


Figure 6.3.3 - Normalised transconductance IV measurements from self-aligned 1408 device

The transfer characteristics, or the variation in the drain current with varied gate voltage at a fixed source-drain voltage are shown for a typical self-aligned device in **Figure 6.3.2**. The source-drain voltage is varied in steps from 0.5 to 2V as indicated in the figure, while the gate voltage is swept from +0.4V to -1V. Pinch off of the drain current is again observed at ~ -0.8V V_{gs} at higher source-drain bias. The rate of change of the drain current with gate bias defines the device transconductance at a constant source-drain potential and hence the transconductance curves for each set source-drain bias were extracted and are presented in **Figure 6.3.3**. The peak transconductance of 340mS/mm occurs close to zero gate bias and at a source-drain bias of 1.5V.

Multi-bias S-parameters from varied width self-aligned devices were measured within the V-band range by methods described in **Section 4.3**. A model in the form of an equivalent HEMT circuit was then constructed with Microwave Office software that emulated the range of measured S-parameters under set bias conditions.

The S-parameters generated by this equivalent circuit can be compared with those extracted from the device under these bias conditions:

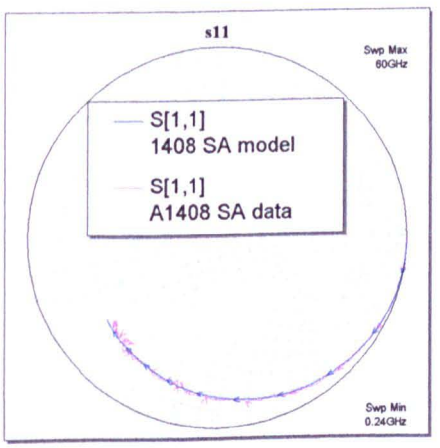


Figure 6.3.4a - S11 magnitude and phase (smith)

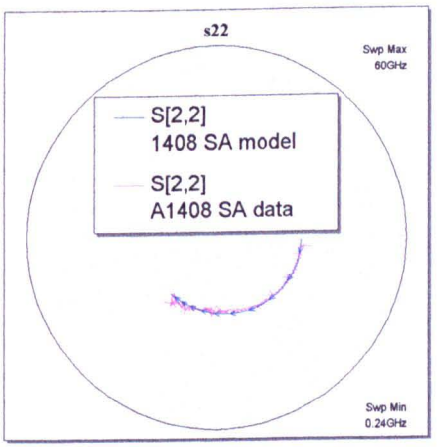


Figure 6.3.4b - S22 magnitude and phase (smith)

Figure 6.3.4 - 1408 measured (pink) and modelled (blue) S-parameters

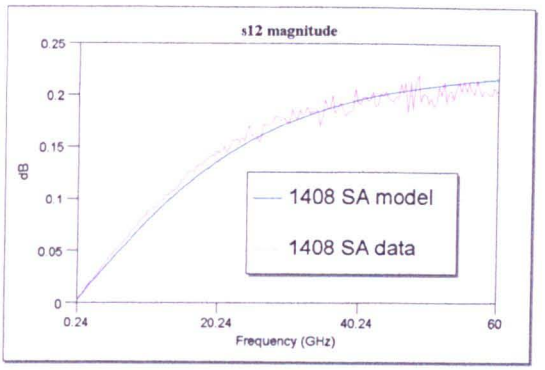


Figure 6.3.4c - S12 magnitude

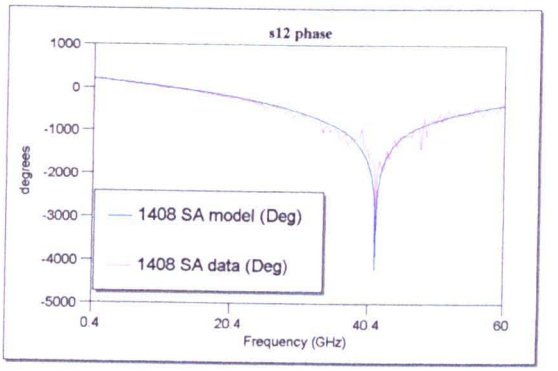


Figure 6.3.4d - S12 phase

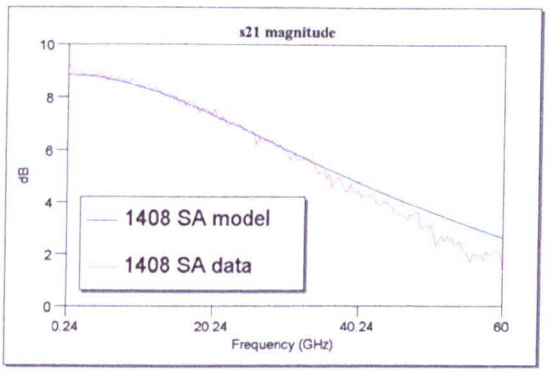


Figure 6.3.4e - S21 magnitude

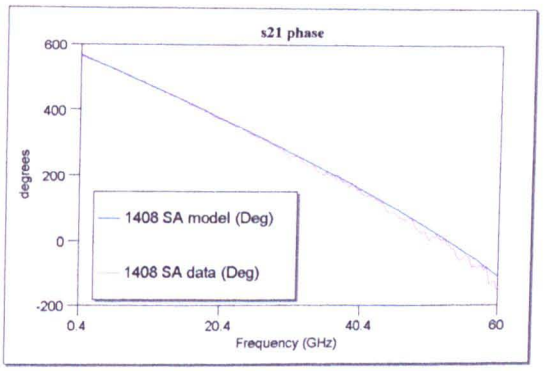


Figure 6.3.4f - S21 phase

Figure 6.3.4 - 1408 measured (pink) and modelled (blue) S-parameters

The S-parameters S11 and S22 are presented in Smith plots while the magnitude and phase of S21 and S12 are shown in separate linear plots. The experimental data (pink) is compared with that generated by the equivalent circuit (blue) for a 2x50μm self-aligned device at V_{ds} = 1.5V and V_g = -0.2V.

The good agreement between experimental data and the model is shown across all four parameters.

The de-embedded equivalent circuit element values used in this model are given in **Figure 6.3.5**:

Parameter	Value	Parameter	Value
C_{gs}	40.4 fF	R_i	0.1 Ω
C_{gd}	7.2 fF	C_{gsp}	1.14 fF
C_{ds}	14.3 fF	C_{gdp}	7.5 fF
g_m	44 mS	C_{dsp}	4.43 fF
R_{ds}	127.5 Ω	L_g	4.63 pH
R_d	1.04 Ω	L_s	0.1 pH
R_s	1.18 Ω	L_d	0.75 pH
R_g	2.08 Ω		

Figure 6.3.5 - Equivalent circuit parameters for a self-aligned 2 x 50μm device

Having identified an accurate equivalent circuit, the de-embedded device performance could then be simulated following the removal of the probe pad elements from the circuit.

The H21 and MAG response generated by the equivalent model is presented in **Figures 6.3.6** and **6.3.7** respectively.

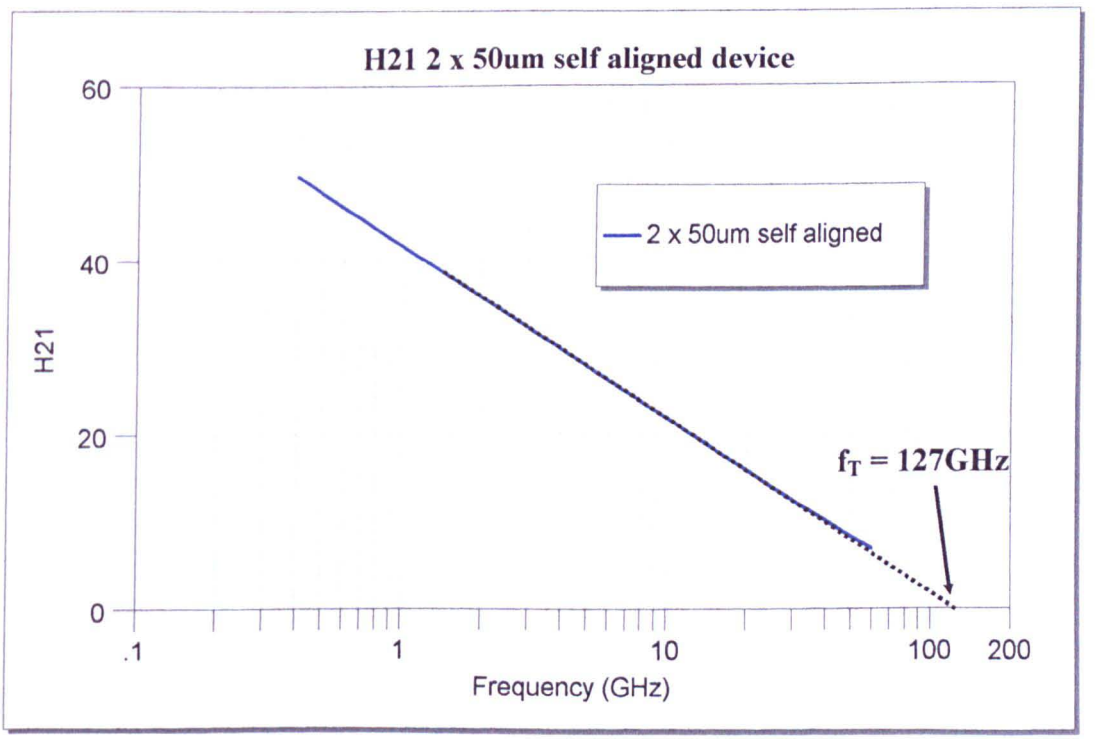


Figure 6.3.6 - H21 plot for 2 x 50μm self-aligned device

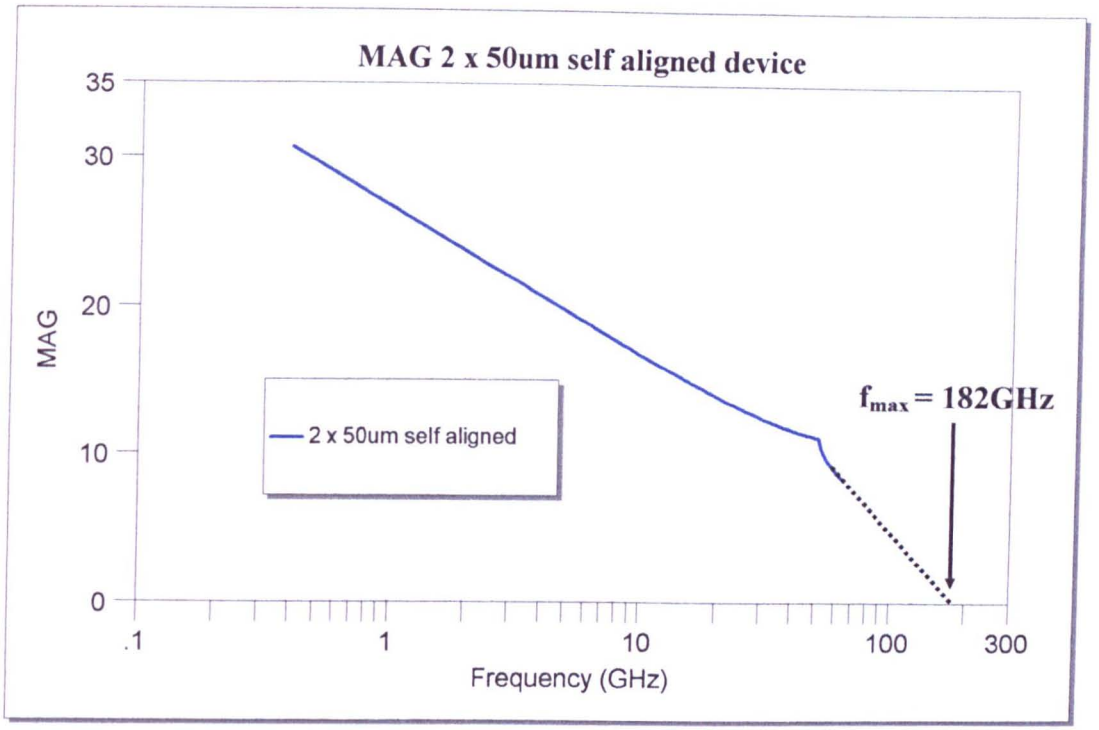


Figure 6.3.7 - MAG plot for 2 x 50μm self-aligned device

Figures 6.3.6 and **6.3.7** demonstrate figures of 127GHz and 182GHz for f_T and f_{max} respectively for a 2 x 50μm self-aligned device. It should be noted that the maximum f_T attained from this batch of self-aligned devices exceeded this value and was recorded at 135 GHz for a 2 x 100μm device. The explanation for the increase in this figure most likely lies with reduced influence of parasitic gate capacitance with wider devices. As discussed in **Section 3.5**, the gate feed metalisation makes contact to the active mesa to ensure continuity of the gate metal at the edge of the mesa. This additional effective gate capacitance will remain constant with varied device width and hence will reduce performance more significantly with narrower devices

Unfortunately the majority of operational self-aligned devices were of shorter width due to the probability of gate-ohmic contact shorting being greater with wider devices. The mechanical yield across the varied device widths is presented in **Figure 6.3.8**:

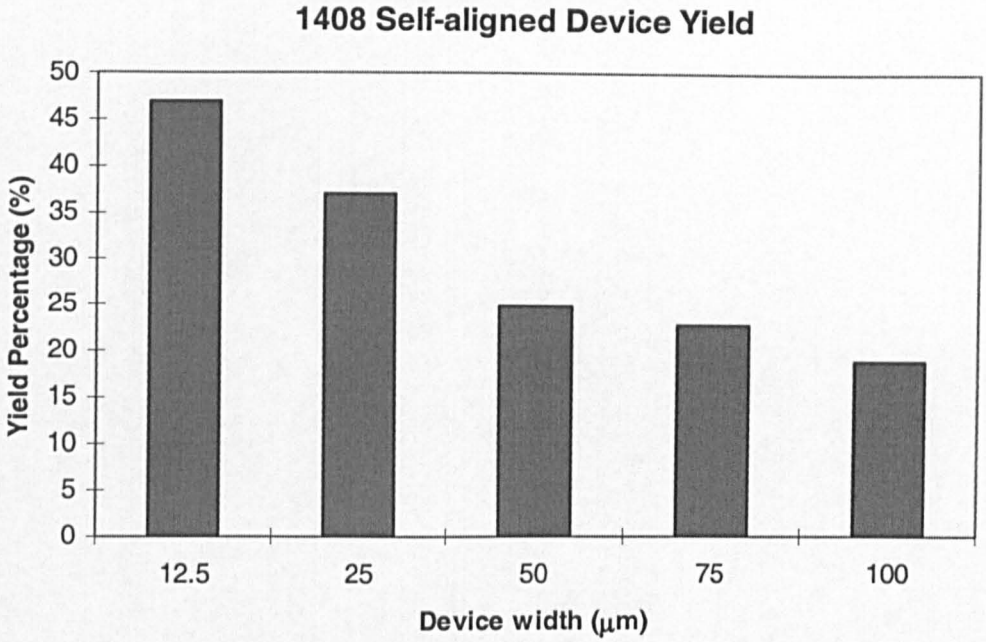


Figure 6.3.8 - 1408 device yield with device width

From the range of devices that were found to be operational, little non-scalable deviation was observed in their DC characteristics illustrating the stability of the self-aligned process.

6.4 Self-aligned and standard device comparison

To provide a greater understanding into the operation of the self-aligned devices, device results were compared with those extracted previously for standard GaAs pHEMT devices of 120nm gate length, fabricated using a standard pHEMT material structure similar to that of material 1258 (**Appendix BI**). The process flow used for the fabrication of these standard devices was similar to that described in **Section 3.5**.

The output characteristics for both standard (blue) and self-aligned (red) are presented in **Figure 6.4.1** for discussion.

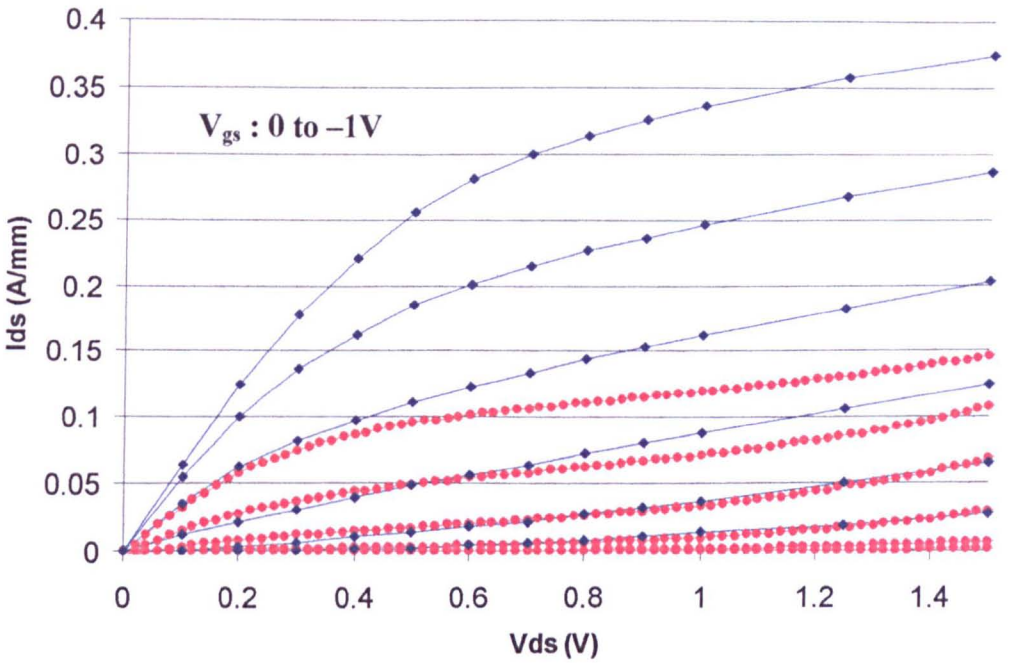


Figure 6.4.1 - Standard (blue) and 1408 self-aligned (red) device output data comparison

The example data presented in **Figure 6.4.1** clearly demonstrates a large decrease in the drive current with the self-aligned 1408 devices compared with the standard device data. In the saturation region for both devices only ~35% of the saturation current of that observed with the standard device is seen with the self-aligned device at zero gate bias. However the self-aligned device exhibits a smaller pinch-off voltage of ~ -0.8V V_{gs} compared to closer to -1V V_{gs} for the standard device. A maximum transconductance figure of 500mS/mm is extracted for the standard device compared with the peak value of 340mS/mm attained for

the self-aligned. Given that the self-aligned process should in effect reduce device parasitic resistances, this reduced DC performance compared to the standard device data was contrary to the expected performance enhancement.

Material characterisation of both the 1408 and standard pHEMT material structures were then compared through Van der Pauw (VDP) measurements:

	1408		Standard
	No etch, (Cap on)	Etched, (No cap)	Etched, (No cap)
Carrier Concentration, (cm^{-2})	1.2×10^{13} →	2.1×10^{12}	3.1×10^{12}
Sheet Resistance, ($\Omega/\square$)	180 →	500	370
Mobility, (cm^2 / Vs)	2850 →	5910	5800

Figure 6.4.2 - VDP 1408 and standard material comparison

VDP results yield values for the carrier concentration, sheet resistance and mobility through the material structure by methods described in **Section 4.1**. In **Figure 6.4.2**, these values are given for the 1408 material, with and without cap by measuring before and after recess etching. For the standard material only the etched, capless data is provided. By removing the cap of 1408, the carrier concentration is reduced leading to an increase in sheet resistivity. In contrast, the average mobility through the structure increases due to the removal of the ionised impurity scattering through the highly doped cap. Comparison between the capless data for the two materials indicates a significantly larger carrier concentration in the standard material, resulting in a lower sheet resistance. Values for the carrier mobility remain similar for the two suggesting that no damage was introduced through the succinic acid recess etch. Although this difference between the carrier concentrations for the two materials will directly affect the DC characteristics of the completed device, it is suggested that that magnitude of difference observed in the DC data cannot be explained by these differences in the material characterisation.

Considering again the comparison between the 1408 and standard device data given in **Figure 6.4.1**, the low field IV response which reflects the axial resistance through the device, i.e. the total resistance between source and drain contacts, is substantially higher with the self-aligned 1408 device indicating a much higher axial resistance than for the standard device. The magnitude of this resistance cannot be explained by the sheet resistance and the separation of the source and drain contacts defined in the self-aligned process. Instead the solution is found to lie with the characterisation of the ohmic contact resistance through the standard TLM method.

6.5 Analysis of the non-annealed ohmic process

As is discussed extensively in **Section 4.2**, to accurately extract a contact resistance figure for a HEMT type structure, a recessed TLM test structure must be employed. For standard HEMT type fabrication processes which involve annealing of the ohmic contacts to provide a low contact resistance, a standard TLM type structure with which the cap remains intact is typically used to attain values for the specific contact resistance. Often this method can provide sufficiently accurate figures for the contact resistance for the completed device. However for the non-annealed process, it becomes imperative to adopt the recessed TLM method if an accurate figure for the contact resistance is to be extracted.

By annealing the ohmic contact metals to encourage their diffusion into the material, the region below the contacts is effectively highly doped, minimising resultant potential barriers that arise throughout the heterostructure. However with a non-annealed ohmic process, this diffusion does not occur and therefore carrier transport vertically through the structure will depend solely on the band profile vertically through the structure. Again, as we are only concerned with electron transport, only the conduction band through the material structure is considered. To better understand the process of vertical conduction through the 1408 structure, a conduction band/carrier concentration plot as simulated using Greg Snider's 1D Poisson/Schroedinger solver for the active layers of the material is presented in **Figure 6.5.1.**

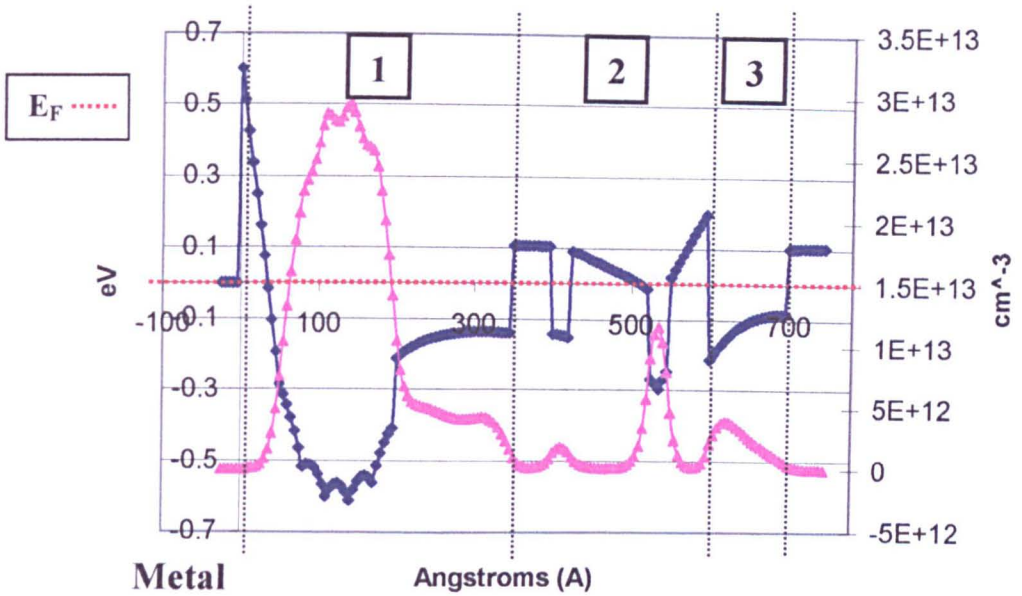


Figure 6.5.1 - 1408 Conduction band and carrier concentration profiles

Figure 6.5.1 demonstrates the conduction band trend (blue) and carrier concentration distribution (pink) through the cap, barrier/spacer and channel layers denoted by regions 1, 2 and 3 respectively. The ohmic metal is situated to the left of the left hand y axis and forms a Schottky barrier of 0.6eV in height from the Fermi level [6.2] defined at 0eV on the left hand y-axis, and ~ 3nm in depth into the highly doped cap layer. The right hand y-axis indicates the three-dimensional carrier concentration through the structure. For conduction to occur through the ohmic contact, electrons must travel between the ohmic metallisation and channel layer or vice versa, and hence must pass through the complex potential barrier distribution formed within the cap and barrier/spacer layers. The magnitude and shape of these barriers will therefore dictate the nature of transport across them and hence define the effective resistance through the contact. By considering the methods of current transport described in **Section 2.2**, the nature of transport through the 1408 structure is better understood. As an example, electrons moving from the ohmic metal to the channel under an applied external field are considered. Initially they must cross the large Schottky barrier formed at the metal/cap interface. Due to the large height and relatively narrow width of the barrier, it is speculated that tunnelling be the main method of transport in this region. Once in

the highly doped cap layer, a variety of smaller but thicker barriers arise as formed by the complex barrier and spacer layer structure before the channel layer. It is believed that due to the effective width of these barriers, the tunnelling current through them would be extremely low leading to thermionic emission accounting for the majority of the current flow across them. This process will of course be similar for electrons travelling in the opposite direction i.e. from the channel to the ohmic metallisation. However, due to the rectifying nature of the Schottky barrier as discussed again in **Section 2.2**, the effective impedance of the barrier should be less with the direction of current flow from semiconductor to metal than vice versa. The magnitude of this difference will of course depend on the nature of transport across the barrier as a reflection of the barrier dimensions, as thinner barriers relying on tunnelling transport will be less rectifying.

The accuracy of the standard TLM method in the extraction of a figure for the contact resistance will depend on the ratio of the resistance through the cap region of the structure to that through the channel via conduction across the barrier and spacer layers. In the instance of 1408, as is affirmed by the carrier distribution given in **Figure 6.5.1,** the large carrier concentration in the cap layer will promote a significant amount of current flow through the cap layer. This combined with the existence of the sufficiently large barriers across the barrier/spacer layers would result in the majority of the current to flow through the cap in the standard TLM structure, and hence the figure extracted for the contact resistance in this instance will merely be the effective Schottky contact resistance.

This process would further explain the inconsistency in the value attained for the resistance with the non-annealed ohmic contact process with the 1408 material structure using the standard TLM process compared with DC measurements of the actual devices. The initial value of 0.12 – 0.14Ω.mm when combined with VDP measurements and device dimensions could not account for the large axial resistance observed with the completed 1408 self-aligned devices. A realistic value for the non-annealed ohmic contact process was then

extracted from recessed TLM structure measurements upon the 1408 material using the process as described in **Section 4.2**. This did indeed demonstrate a substantial increase in the contact resistance from that produced by the standard TLM process and yielded a more realistic figure in the range of 0.45Ω.mm for the specific ohmic contact resistance using the non-annealed ohmic process.

6.6 Chapter Summary

In an attempt to develop a thin low temperature ohmic contact process for compatibility with a self-aligned gate GaAs pHEMT fabrication process flow, various ohmic metallisation recipes using four differing pHEMT type material structures were investigated across a range of annealing regimes. The quality of each strategy was examined using standard TLM type structures to yield figures for the specific contact resistances for each. From these figures it was concluded that using either 1408 or 1409 materials combined with a specific ohmic metallisation comprising of gold, germanium and nickel, the lowest resistance figures could be achieved without thermal annealing of the contacts. Having identified a suitable ohmic process, a selective succinic acid based recess etch was then developed which was compatible with the more exotic InGaAs/GaAs cap structure inherent to the 1408 and 1409 material structures. The etch was found to be extremely controllable as well as repeatable and hence was integrated into the developing self-aligned process. HEMT devices of 120nm gate length were then fabricated using the 1408 material structure and self-aligned gate process. Results from the completed devices indicated excellent transistor operation, however at DC the performance was found to be inferior to those from previous standard pHEMT devices of similar gate length. Investigation into the properties of the two material structures then suggested inconsistency in the measured contact resistances from the TLM test structures and the actual devices. Analysis of the accepted standard TLM method highlighted the inaccuracy of such a method for the extraction of an accurate contact resistance for the non-annealed process and hence was modified to produce an accurate resistance figure. This figure proved to be substantially larger than that initially extracted and corroborated the measured device DC data.

Beyond the DC device performance, the self-aligned devices performed admiralably at high frequency and indeed outperformed standard 120nm pHEMT results for both f_T and f_{max} figures.

Chapter Publications

[1] D. Moran, E. Boyd, H. McLelland, K. Elgaid. Y.Chen, D. S. Macintyre, S. Thoms, C. R. Stanley, and I. G. Thayne, *Novel technologies for the realisation of GaAs pHEMTs with 120 nm self-aligned and nano-imprinted T-gates*, Microelectronic Engineering 2003, Volume 67 - 68, p. 769 - 774.

References

[6.1] H. Fourre, F.D., A. Cappy, *Selective wet etching of lattice-matched InGaAs/InAlAs on InP and metamorphic InGaAs/InAlAs on GaAs using succinic acid/hydrogen peroxide solution.* Journal of Vacuum Science Technology, 1996. B14(5): p. 3400 - 3402.

[6.2] K. Kajiyama, Y. Mizushima, S. Sakata, *Schottky barrier height of n-InxGa1-xAs diodes.* Applied Physics Letters, 1973, 23(8): p. 458 - 459.

Chapter 7

Results II

Self-aligned Lattice-matched InP HEMT

Following from the success of the development and implementation of a self-aligned gate process for GaAs pHEMT, the task of developing a similar process for the realisation of self-aligned lattice matched InP HEMT devices was undertaken. By moving to the InP material system, the device performance will increase beyond that for GaAs pHEMT for a set gate length due to increased indium content within the channel, increasing carrier transport through the device [7.1]. However due to the large conduction band offsets that arise through the $In_{0.53}Ga_{0.47}As$ / $In_{0.52}Al_{0.48}As$ heterostructure, the ability to form an efficient low resistance non-annealed ohmic contact would at first appear more challenging than with the GaAs pHEMT system. This obstacle is tackled through the introduction of an additional layer of delta doping to the InP material structure which acts to reduce the magnitude of the potential barriers vertically through the active layers in addition to increasing the carrier concentration.

Using a process flow similar to that for the GaAs pHEMT self-aligned devices, devices of both 120nm and 70nm gate length were fabricated using the double delta doped InP material. In addition, devices with both self-aligned and standard ohmic contact separation were fabricated at both gate lengths for comparison and do indeed show a significant increase in performance at both DC and RF by moving from a standard to a self-aligned gate process.

7.1 Material design and ohmic optimisation

In depth characterisation of the 1408 self-aligned devices and in particular the extraction of an accurate ohmic resistance figure through recessed TLM measurements, highlighted the operation of the non-annealed ohmic process through the material structure. The existence of the distributed potential barrier system through the active layers as indicated by Poisson-Schroedinger modelling led to a large effective contact resistance of approximately three times that using a similar annealed ohmic process. For the self-aligned process to be an attractive alternative to standard device processing, the total parasitic resistances must be minimised to less than that for the standard process. This ohmic contact resistance must therefore be reduced to be comparable with standard annealed contact figures to be able to observe the reduction in the total parasitic resistances through the reduction in the separation of the source and drain contacts.

To integrate this self-aligned gate technology into a lattice matched InP system, the design of the material must be carefully considered to optimise the vertical carrier transport through the structure to produce a low resistance non-annealed ohmic contact. **Figure 7.1.1** shows the layer structure for a typical InP HEMT material system, scaled for 120nm gate length devices.

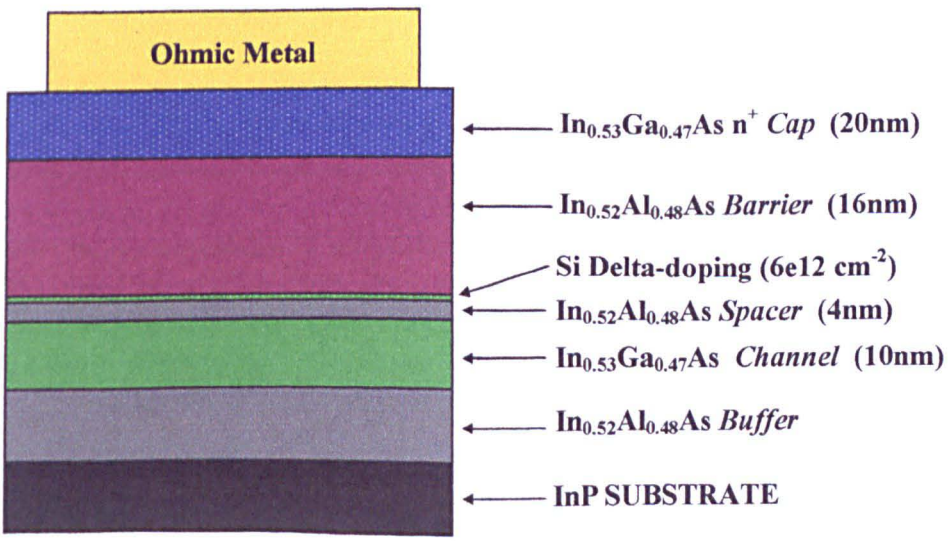


Figure 7.1.1 - 120nm scaled lattice matched InP HEMT structure

This standard structure is similar to the pHEMT structures presented in **Chapter 6** in that it comprises of a highly doped cap layer followed by a thick barrier layer containing a single layer of delta doping separated by a spacer layer from the channel. A thick buffer layer then separates these active layers from the InP substrate.

Figure 7.1.2 displays the conduction band profile (blue) and carrier concentration distribution (pink) through the active layers of the structure given in **Figure 7.1.1**:

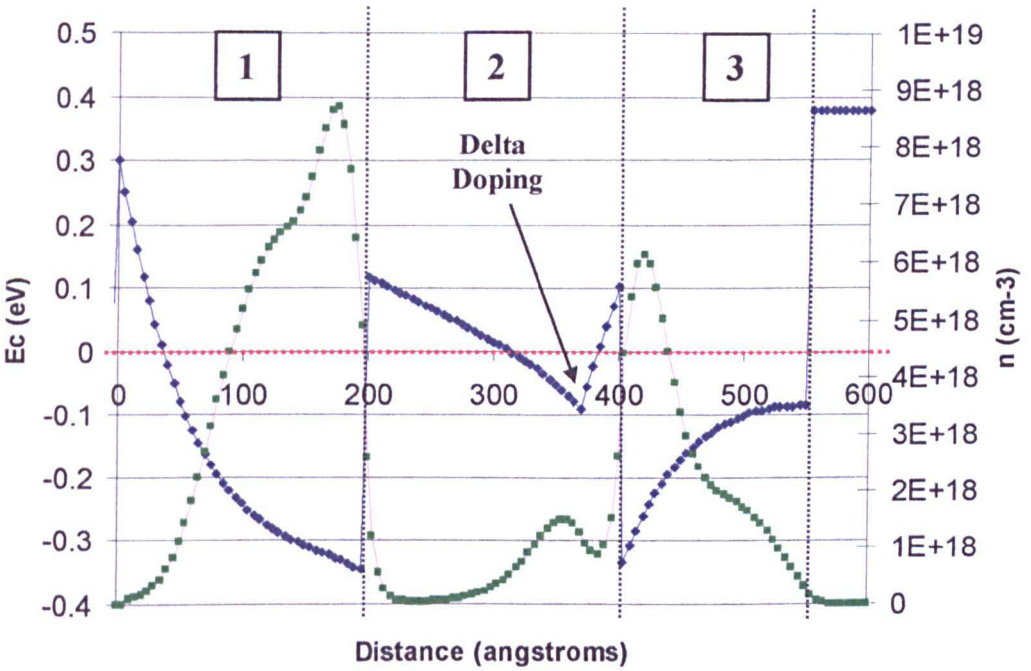


Figure 7.1.2 - InP conduction band and carrier concentration profiles

As before the structure is divided into cap, barrier/spacer and channel regions denoted by 1, 2 and 3 respectively. The Fermi level again is set as 0eV on the conduction band scale.

Similar to analysis of the transport through the 1408 structure, conduction of an electron between the ohmic metal and the channel layer is considered. As before a Schottky barrier is formed between the ohmic metal and cap layer, however due to the reduced band gap of InGaAs with increased indium concentration, the height of this barrier is significantly less than for the pHEMT system [7.2]. The width of this barrier will be dependent on the doping concentration within the cap and in this instance is about 3.5nm at the Fermi level. Due to the delta doping between the barrier and spacer layers, the conduction band edge is dragged to

well below the Fermi level within the effective area of doping. The potential barrier formed by these layers separating the cap and channel layers then consists of two triangular shaped barriers, both 0.1eV high with respect to the Fermi level. In this example, the long triangular barrier that is formed at the cap/barrier layer interface would most likely be the dominant element in the effective resistance between the ohmic metal and channel. In comparison with the 1408 structure, this material would arguably not demonstrate a smaller non-annealed contact resistance and hence needs to be modified to produce a lower contact resistance for compatibility with an efficient self-aligned process.

The effect of introducing a dopant material on the band structure through the material heterostructure is evident from the various conduction band simulations already presented. To ensure the majority of additional carriers introduced by the doping contribute to conduction, the doping material is chosen so that at the designated operating temperature, thermal excitation of carriers to the conduction band occurs. Therefore the larger the doping concentration the greater the distortion of the conduction band profile as within these regions the large carrier concentrations within the conduction band must be accounted for with respect to the Fermi level. This process can be exploited to minimise the potential barrier distribution through such a structure as that given in **Figure 7.1.2.** By introducing additional doping, the conduction band can be distorted further to reduce the magnitude of these potential barriers. This process is with limit however, as in addition to maximising conductance vertically through these layers, the recessed region of the device i.e. that below the gate must maintain a single channel of current flow through the channel as the 2DEG to maximise device performance. The Schottky barrier formed by the gate must also be sufficiently large to minimise gate leakage. A balance is then struck between the amount of additional doping that can be introduced whilst avoiding parallel conduction through the gate region. A solution is found in the addition of a single layer of delta doping within the barrier layer of the material, placed closer to the cap layer than the original delta doping in a bid to

reduce the large barrier formed between cap and channel layers. This process is illustrated in **Figure 7.1.3**:

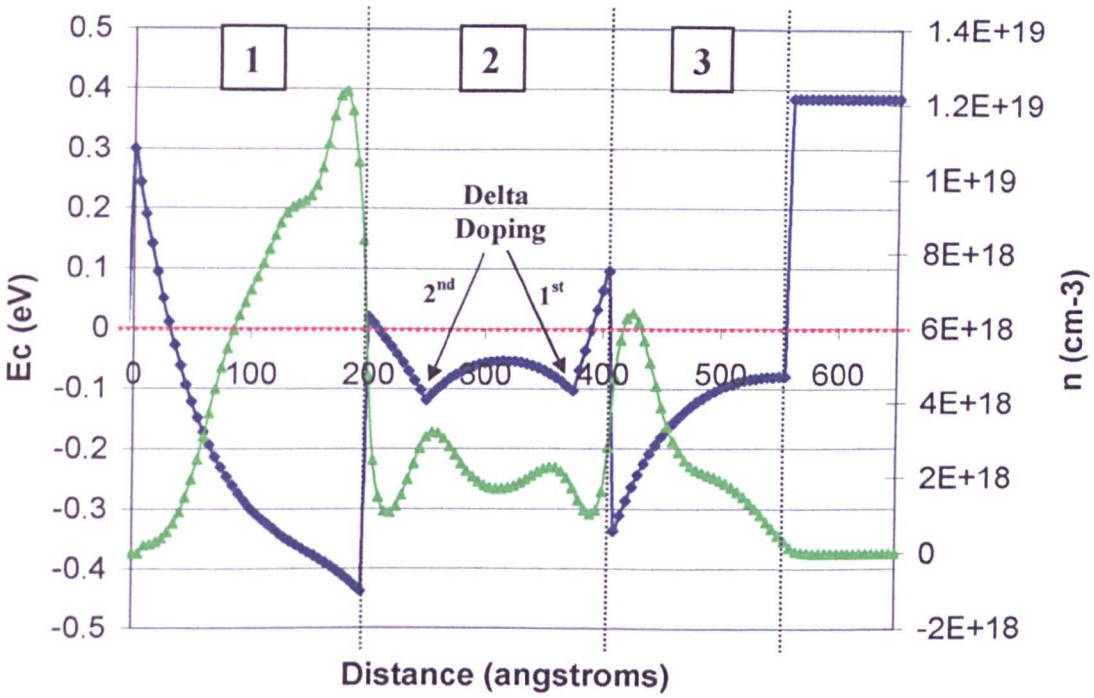


Figure 7.1.3 - Double delta doped InP conduction band (blue) and carrier concentration (green) profiles

The conduction band and carrier concentration profiles presented in **Figure 7.1.3** are those as simulated for a 120nm InP scaled material structure designed for use with a non-annealed ohmic process. The full material structure is given in **Appendix BV**. The benefits in the reduction of the potential barriers through the material with the additional doping are clearly illustrated in comparison with the single doped structure. In addition to the extra delta doping, the doping concentration within the cap layer has been increased to further minimise the Schottky barrier as well as implementing a slight reduction in the spacer layer from 4nm to 3nm.

Re-simulation of the structure with the cap removed to represent the region below the gate in the device is presented in **Figure 7.1.4**:

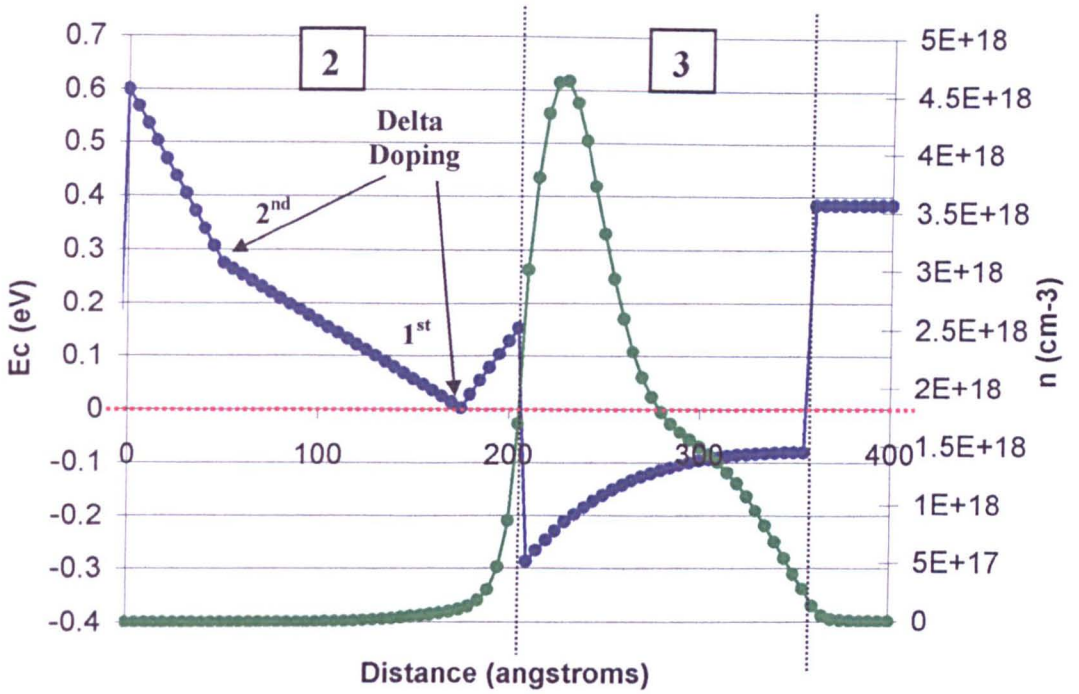


Figure 7.1.4 - Double delta doped InP (gate region)

Proper placement and choice of doping concentration of the second layer of delta doping ensures that the majority of the carrier concentration through the gate region of the device resides in the channel as shown in **Figure 7.1.4**. The magnitude of the Schottky barrier formed by the gate contact at a height of 0.6eV above the Fermi level [7.3] would also appear to be sufficiently large to minimise any current flow between the gate contact and device channel.

A figure for the contact resistance using the non-annealed process with the double delta doped material was then verified through recessed TLM measurements and yielded a value of 0.15 ± 0.02 Ω.mm which is comparable with figures attained with an annealed contact strategy.

By performing standard TLM measurements on the material structure, an equivalent resistance value for the transport across the Schottky barrier could be estimated. By minimising the amount of time between de-oxidation and metallisation of the ohmic level during fabrication in an attempt to minimise the build up of native oxide on the cap surface,

the effective Schottky barrier resistance could be kept to 0.05 ± 0.01 Ω.mm. These contact resistance values were found to be consistent with ohmic metallisation heights of 60nm and above and hence could be integrated into the self-aligned process without degradation of the quality of the contact.

Beyond the characterisation of the non-annealed ohmic process, VDP measurements provided further characterisation of the double delta doped InP material:

	Capped	**Cap removed**
Sheet Resistance (Ω/□)	125	260
Mobility (cm^2 / Vs)	4100	7100
Carrier Concentration (cm^{-2})	1.25×10^{13}	3.4×10^{12}

Figure 7.1.5 - VDP measurements of InP double delta doped material

The additional doping within the material is reflected in the VDP measurements presented in **Figure 7.1.5** by the large carrier concentration that remains after the cap has been removed, leading to a low sheet resistance of 260Ω/□. The mobility figure however is substantially lower than for typical single doped InP HEMT structures with which a capless mobility of between 9000 and 10,000 cm^2 / Vs is common. This reduced mobility is believed to be related firstly to the thin spacer layer of the structure of 3nm compared to a usual value of 4 to 5nm, but also to the carrier population distribution within the channel. As shown by **Figure 7.1.4**, the peak carrier concentration occurs close to the channel - spacer interface and hence will experience greater dopant coulomb scattering and lower low-field mobility.

7.2 120nm Self-aligned and Standard DC Characterisation

Using a process flow similar to that used for the fabrication of the 1408 self-aligned devices, both standard and self-aligned devices of 120nm gate length were fabricated from the double delta doped InP material. Both types of device were therefore identical except that the separation of the source and drain contacts of the standard devices were formed at distance of 1.6μm as is typical with the standard HEMT process. The self-aligned device ohmic contacts were formed at a separation of ~400nm as defined by the dimensions of the head of the gate. In addition, a 120nm Γ-gate process was devised from the 120nm T-gate process by altering the ebeam dose distribution through the T-gate profile. The aspect ratio of each Γ-gate was then tailored by altering the dimensions of the head of the gate whilst maintaining the 120nm foot profile. **Figure 7.2.1a** shows a 120nm Γ-gate resist profile with a 4:1 aspect ratio i.e. the ratio of the length of the drain side of the head of the gate to the source side. **Figure 7.2.1b** shows the resultant Γ-gate after metalisation and lift-off.

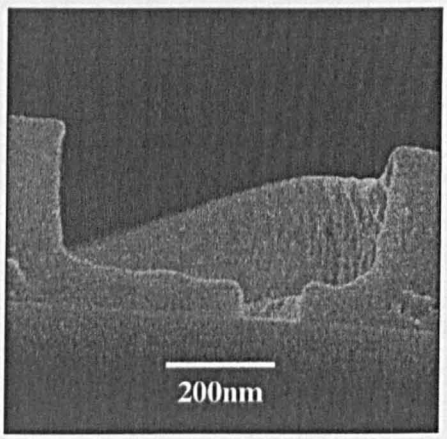


Figure 7.2.1a - 120nm Γ-gate resist profile

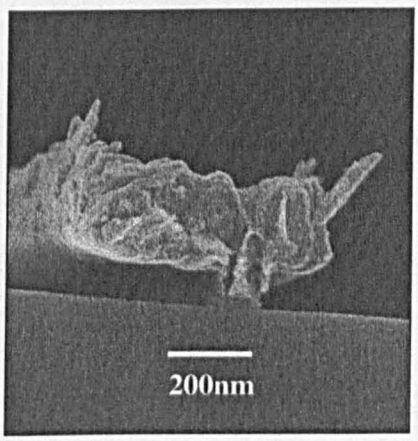


Figure 7.2.1b - Metallised 120nm Γ-gate

The process compatibility of these three types of device meant that standard T-gate (SD), self-aligned T-gate (SA) and self-aligned Γ-gate (SAG) devices could be fabricated upon the same sample for comparison.

In comparison with the GaAs pHEMT self-aligned process flow discussed in **Chapter 6**, the process was altered by the following for compatibility with the InP material system:

An orthophosphoric based wet etch was used to isolate active regions of mesa of the InP material. AFM measurements of the T-gate resist profile and in particular the height of the gate 'stalk' provided a method of determining a suitable ohmic metallisation height to minimise shorting between the gate and ohmics. The succinic acid based etch was modified for the gate recess process and tailored to produce the desired recess dimensions. This was followed by a short orthophosphoric clean prior to gate metallisation. With the introduction of the new metallisation unit (Plassys II), the Pd in the original 120nm gate recipe was replaced with Pt. An ohmic metallisation similar to that used for the GaAs pHEMT self-aligned process, but of 73nm in height was deposited to form the source and drain contacts.

The complete 120nm self-aligned gate InP fabrication process is described in **Appendix AIII**.

The resulting yield from the three types of device fabricated is given in **Figure 7.2.2**:

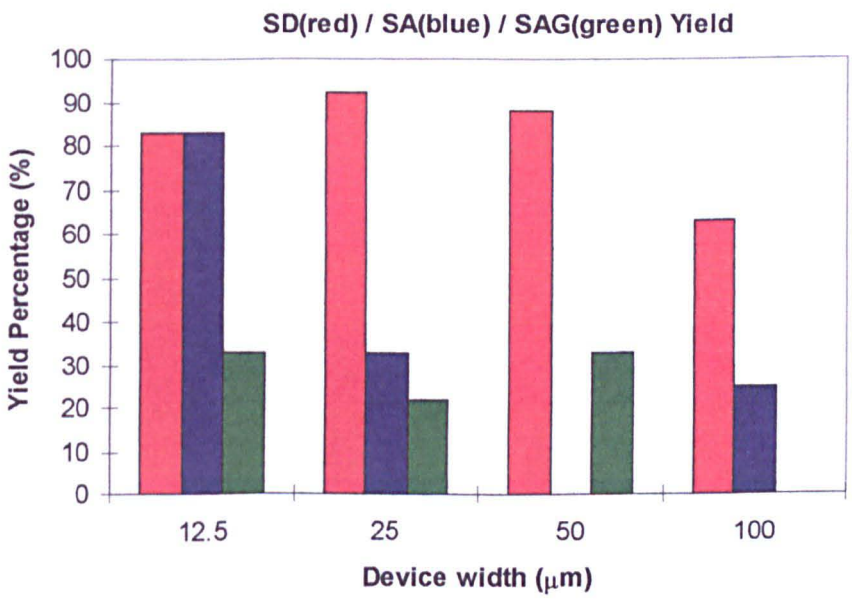


Figure 7.2.2 - Yield for 120nm SD, SA and SAG devices

Very high yield was attained with the SD devices due to the lack of potential shorting between the gate and ohmic contacts which is found to reduce the SA and SAG device yield, particularly for wider devices.

The various types of devices were then characterised at both DC and RF. Comparison between the SD and SA devices at DC indicated a large increase in performance through the reduction of parasitic resistances with the self-aligned system. Typical output characteristics for an SD and SA device are presented in **Figure 7.2.3**:

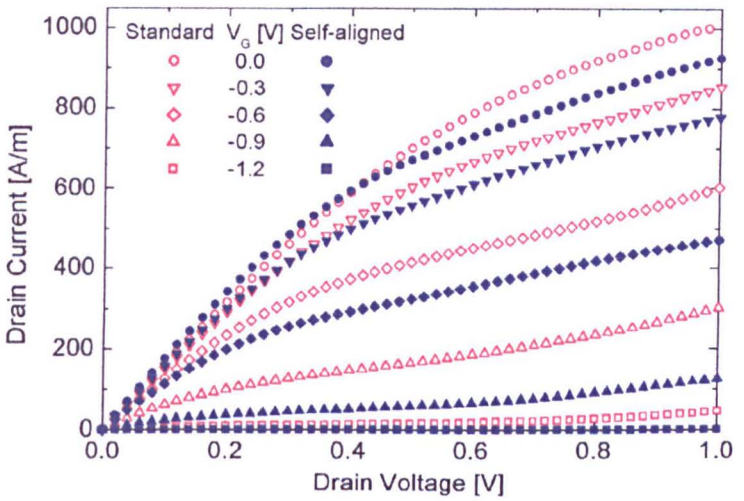


Figure 7.2.3 - SD (red) and SA (blue) 120nm device output characteristics

As before these results are normalised to the width of the device. When considering low source-drain bias measurements for both sets of devices, as expected the SA devices exhibit a lower axial resistance than with the SD devices demonstrated by their steeper IV response. This is further supported by the reduced pinch off voltage with the SA devices of in this instance –1V V_{gs} compared with –1.2V V_{gs} for the SD devices.

This reduced pinch off voltage is easily explained by the reduced voltage drop across the source resistance with the SA devices compared with the SD which results in a larger 'intrinsic' potential difference between the gate contact and channel for a similar gate-source voltage.

A substantial increase in the transconductance figures is also observed between the two types of device as demonstrated by the transfer characteristics for both presented in **Figures 7.2.4** (SD) and **7.2.5** (SA).

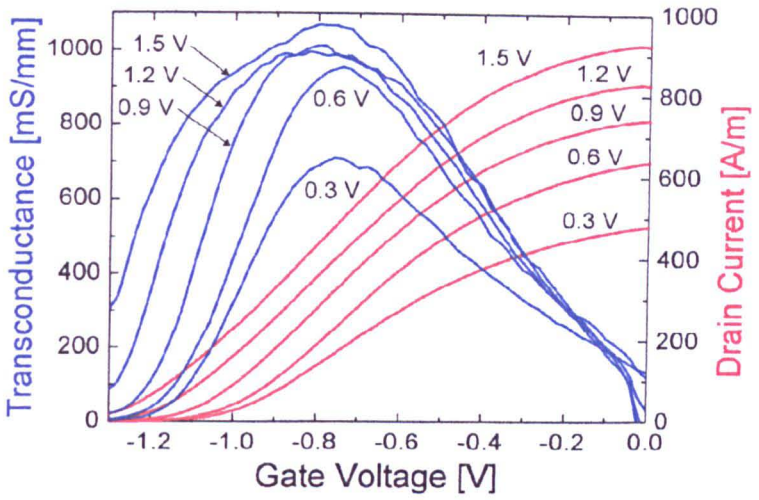


Figure 7.2.4 - SD 120nm transfer and transconductance characteristics

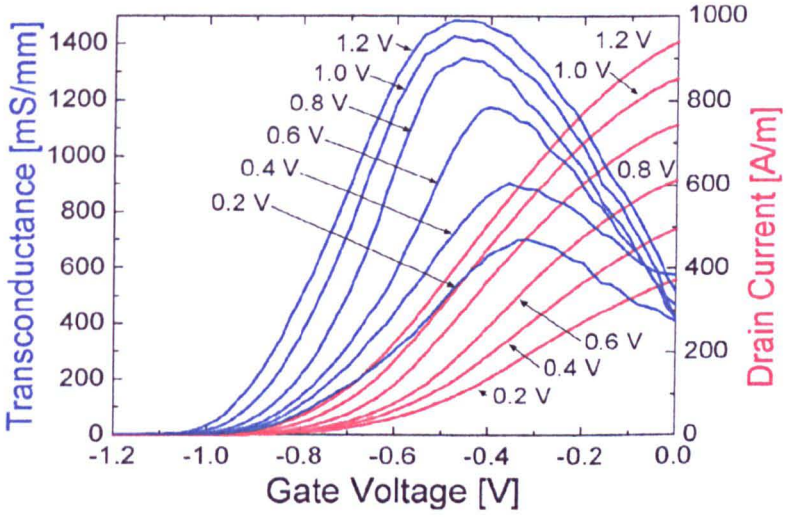


Figure 7.2.5 - SA 120nm transfer and transconductance characteristics

For the SD devices, the transconductance increases with source-drain bias and peaks at a figure of 1080mS/mm at a gate bias of –0.8V. Similarly with the SA devices the peak transconductance occurs with larger source-drain bias but peaks at a value of 1490mS/mm at a gate bias of –0.45V. This corresponds to an increase of 38% in the extrinsic transconductance between the SD and SA devices.

From the definition of the difference between intrinsic and extrinsic transconductance figures discussed in **Section 2.4**, the magnitude of difference between the parasitic source resistances of the SD and SA devices can be extracted from their extrinsic transconductance figures.

However this process assumes the intrinsic device remains similar for the two i.e. the intrinsic transconductance figures for the SD and SA devices are equal.

From **Eqn 2.4.6**, and taking $g_{mSD}^{*} = g_{mSA}^{*}$:

$$\frac{g_{mSD}}{1 - g_{mSD} R_{sSD}} = \frac{g_{mSA}}{1 - g_{mSA} R_{sSA}} \qquad \textbf{Eqn 7.2.1}$$

This can be re-written to express the difference in source resistances R_{sSD} and R_{sSA} as:

$$R_{sSD} - R_{sSA} = \frac{1}{g_{mSD}} - \frac{1}{g_{mSA}} \qquad \textbf{Eqn 7.2.2}$$

Taking the example peak transconductance values from **Figures 7.2.4** and **7.2.5** of g_{mSD} = 1080mS/mm and g_{mSA} = 1490mS/mm, this yields a reduction of $R_{sSD} - R_{sSA}$ = 0.25Ω.mm between the standard and self-aligned T-gate sources resistances.

This value can be compared with the difference in the low field axial resistances through both types of device:

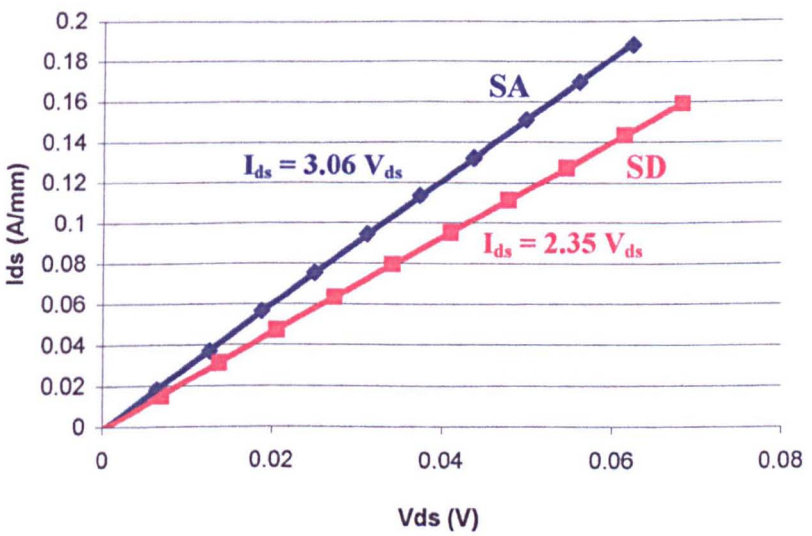


Figure 7.2.6 - Low-field SA (blue) and SD (red) output response at 0V V_{gs}

From **Figure 7.2.6**, under these low bias conditions and with zero gate bias, the IV response from source to drain should be ohmic and reflect the total axial resistance through the device. The difference in this resistance between the SD and SA devices should therefore correspond with the reduction in source and drain parasitic resistances using the self-aligned process.
In the example given in **Figure 7.2.6**, the difference in this low field axial resistance (LFAR) is found to be ~0.1 ± 0.01Ω.mm. Assuming that this additional resistance is divided equally between the source and drain parasitic regions, this corresponds in this example to a source resistance reduction of 0.05Ω.mm with the SA devices.

However from the calculations performed taking into account the extrinsic transconductances for both types of device, this figure should be closer to 0.25Ω.mm, five times as large as that extracted from the output characteristics.
Also, by considering the dimensions of each type of device and from the sheet resistance of the material extracted from VDP measurements, an estimate for the extra source resistance between SD and SA devices can be made. Given that the SD device source-drain contact separation is 1.5 ± 0.1μm and that the length of the T-gate head of the SA devices is 450 ± 20 nm, this corresponds to a reduction of ~ 525 ± 60nm of material conduction on either the source or drain side between the SD and SA devices. This of course assumes that the gate is aligned exactly mid source-drain gap for the SD devices. VDP measurements of the double delta doped InP material with the cap complete indicated a sheet resistivity of 130 Ω/ . This corresponds to a width normalised resistance figure of 0.068 ± 0.008Ω.mm across 525nm of the material.

Again this figure is much less than that predicted from the transconductance figures from the SA and SD devices of 0.25Ω.mm, and is in fact much closer to the LFAR figure extracted from the device output characteristics of 0.05Ω.mm. From these results it is concluded that although the SD and SA devices are identical with the exception of differing source - drain

contact separation, to justify the large increase in transconductance with the SA devices, the intrinsic transconductance of each must differ. This result is perhaps not so surprising when considering the different electric field distribution within the reduced geometry SA devices compared to that through the SD. It is believed that this increased field concentration leads to the process of *real space transfer* of electrons to surrounding layers as is discussed in **Section 2.6**. This idea is corroborated upon comparison of the output characteristics for SD and SA devices, where a distinct kink is observed with the SA device curve that is absent with the SD devices:

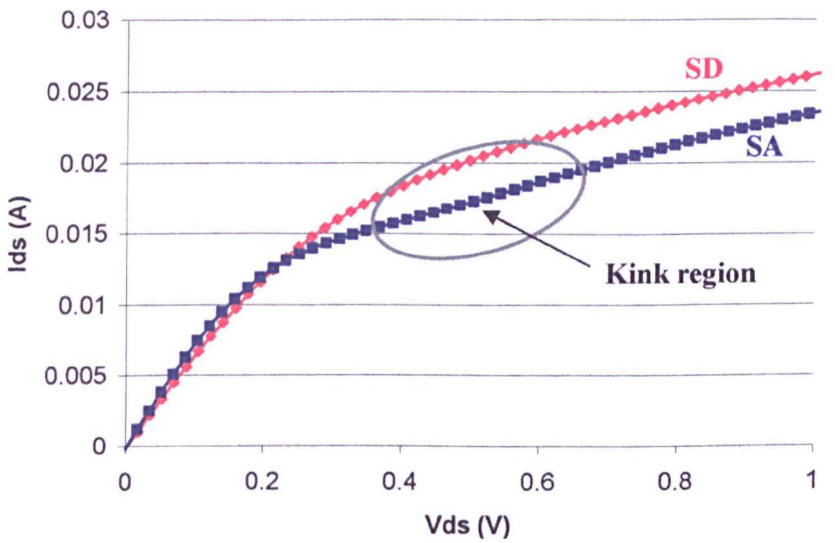


Figure 7.2.7 - SA (blue) and SD (red) output response at 0V V_{gs}

As shown in **Figure 7.2.7**, the knee voltage of the SA is lower than with the SD devices, indicating the onset of current saturation with the increased electric field through the device at a lower source-drain bias. Beyond the knee however, the current appears to begin to saturate before increasing more sharply again. This effect is better observed by plotting the differential resistance with respect to the source-drain voltage for each type of device, i.e. dV_{ds}/dI_{ds}:

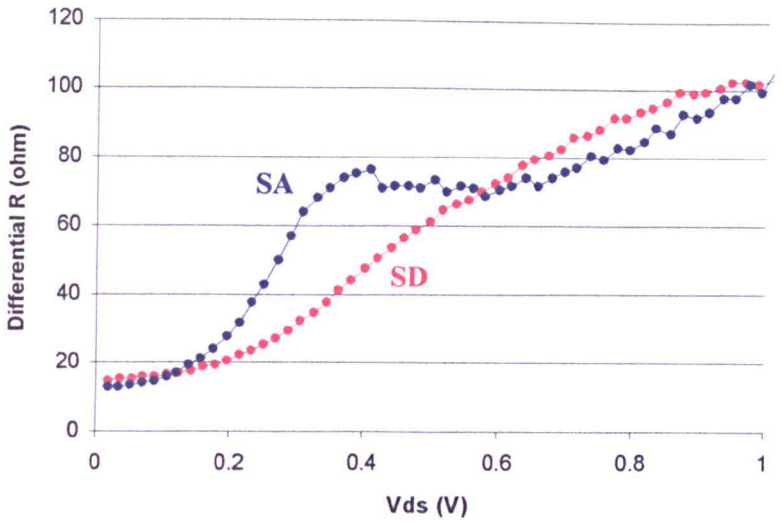


Figure 7.2.8 - Differential source-drain resistance for SD (red) and SA (blue) 120nm

The kink in the SA output characteristics becomes more evident by a peak followed by a slight decrease in the differential resistance for the SA device compared with smooth increase with the SD device as the current begins to saturate.

As discussed in **Section 2.6**, this change in differential resistance generally arises from the transfer of 'hot' electrons to surrounding layers. With the reduced source and drain contact separation of the SA devices, this process is exaggerated by the larger electric field through the SA device and would explain why this effect isn't observed in the SD device results.

Due to the nature of conduction through the device parasitic source and drain regions with the non-annealed ohmic process, the extent to which the self-aligned process benefits overall device performance is not as apparent when compared with the standard device results fabricated with the non-annealed process. In a typical HEMT architecture, current will flow vertically through the ohmic region beneath the source and drain contacts, and then horizontally through the device channel before reaching the intrinsic device region below the gate. The total parasitic source resistance for example will therefore equal the sum of the vertical resistance through the source contact plus that through the channel between the ohmic and intrinsic device regions. However, due to the optimisation of vertical transport between the cap and channels layers with the double delta doped material, conduction

between these layers will occur outside the ohmic region i.e. outwith the area directly below the ohmic contact. Parallel conduction through the cap and channel layers therefore acts to reduce the horizontal parasitic resistance between ohmic and intrinsic device regions. With an annealed ohmic process, this parallel conduction does not occur to the same extent due to the large impedance between the cap and channel layers outwith the ohmic region and hence conduction horizontally through the structure occurs predominantly through the channel.

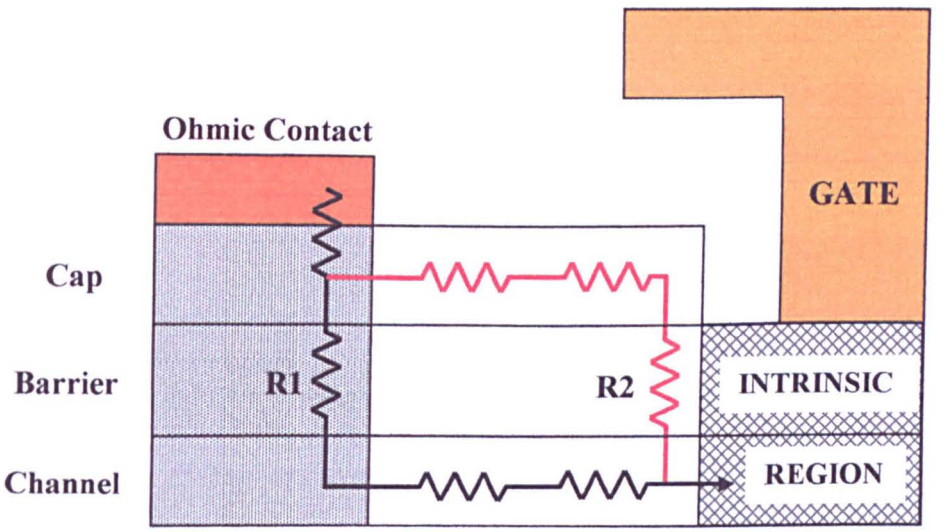


Figure 7.2.9 - Conduction through parasitic device region

In the example, the route by which current flows in a standard HEMT device from the ohmic contact to the intrinsic device region is illustrated in black by equivalent resistance elements. The annealing of the ohmic metals as discussed previously encourages low resistance conduction vertically through the area directly below the contact. With annealing therefore, the magnitude of the resistance R1 will typically be significantly less than that for R2 as the effective potential barrier resistance across the barrier layer will be large for a standard HEMT structure, forcing the majority of current to flow through the channel. However with R1 = R2, and with both of the same magnitude as R1 with the annealed process, parallel conduction will occur between the cap and channel layers depending on the sheet resistance of both layers in addition to the electric field distribution through the structure. With the non-annealed process and the double delta doped material, the low sheet resistance of the structure as measured with the cap on will therefore significantly reduce the total parasitic

resistance compared to that for an equivalent single delta doped and annealed ohmic structure.

As the self-aligned process acts to reduce this total parasitic resistance through a reduction in the separation of the ohmic and intrinsic device regions, the reduction in resistance seen using the non-annealed ohmic process is less than for a standard 'annealed ohmic' device due to a reduced sheet resistance and hence total horizontal resistance between ohmic and intrinsic regions.

For example, with a capped sheet resistance of 130Ω/ for the double delta doped material, total parasitic resistance figures of 0.24Ω.mm and 0.18Ω.mm can be calculated for SD and SA devices respectively taking the vertical contact resistance to be 0.15Ω.mm and from device dimensions. This can be compared with a standard HEMT structure with a capless sheet resistance of 350Ω/ (a standard figure) which would produce total parasitic resistance figures of 0.4Ω.mm for a SD device and 0.22Ω.mm for a SA device again using 0.15Ω.mm for the vertical contact resistance. With the higher effective sheet resistance resulting from conduction through the channel alone with a standard HEMT structure, the predicted reduction in the total parasitic resistance with a self-aligned process becomes more apparent than with the double delta doped material, (a reduction of almost 50% with the standard HEMT compared to only 25% with the double delta doped HEMT in this example).

Self-aligned Γ-gate (SAG) and T-gate (SA) device comparison

Due to the limited device yield from the number of devices fabricated, little statistical comparison could be made between SA and SAG devices. However comparison between this DC data indicated very little difference between SA and SAG device output and transport characteristics:

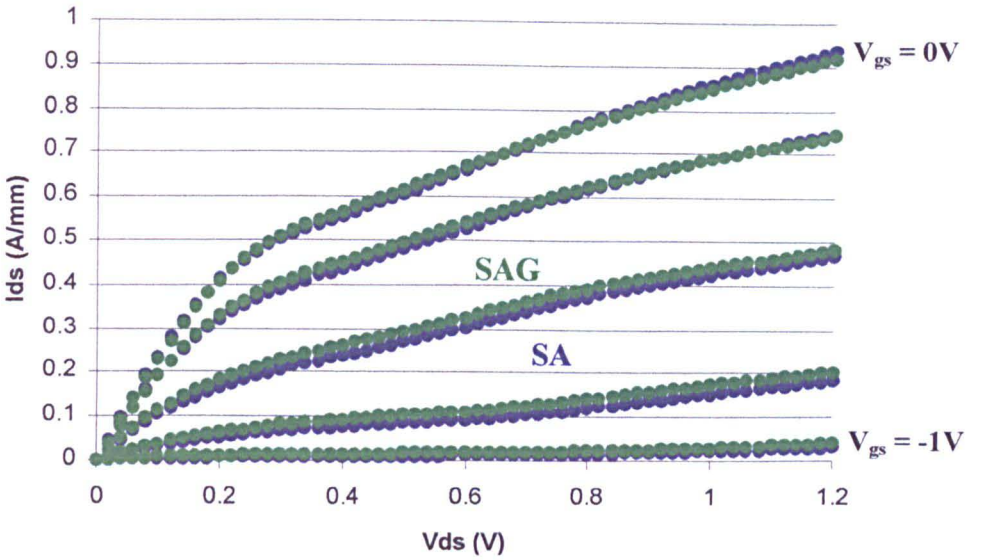


Figure 7.2.10 - 120nm SA (blue) and SAG (green) output characteristic comparison

The comparison between SA and SAG output characteristics presented in **Figure 7.2.10** clearly shows the close resemblance between the two types of device with both exhibiting similar pinch-off and saturation characteristics.

Similarly the peak transfer and transconductance characteristics for both the SA and SAG devices remained similar:

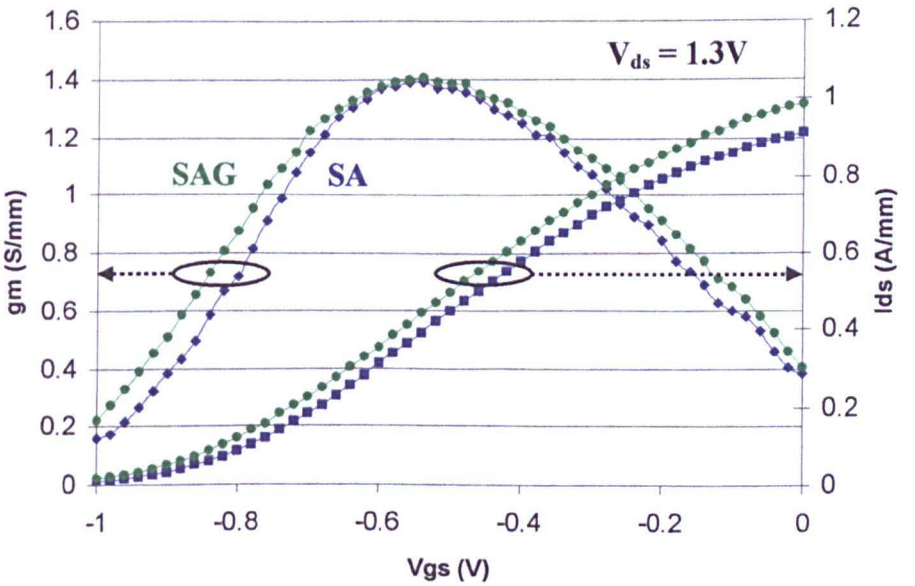


Figure 7.2.11 - 120nm SA (blue) and SAG (green) transfer / transconductance characteristics

Figure 7.2.11 demonstrates a comparable peak transconductance of ~1.4 S/mm for both types of device although the SAG exhibits a wider transconductance peak. Again both devices pinch-off at ~ -1V while the SAG device exhibits a slightly higher drain current with zero gate bias. This last result is unexpected as the axial resistance through the SAG should be larger than with the SA device, and hence exhibit a lower drain current at a similar source - drain bias. This is most likely due to statistical variation that is observed across the range of fabricated devices and is amplified by the scaling of the drain current.

The main objective when developing a Γ-gate process was to produce the performance of a self-aligned device whilst increasing the device breakdown voltage by extending the distance between the drain and gate contacts. The Γ-gate process used for the SAG devices was optimised for a 3:1 gate aspect ratio, which should increase the drain - gate contact distance to three times that of the self-aligned T-gate device. Again due to the limited number of operational SA and SAG devices, little analysis could be performed into the breakdown characteristics of each type of device without destroying the device. However from the few measurements performed it would appear that little improvement was introduced through the incorporation of a Γ-gate process as the breakdown voltage only increased by 0.2V from 1.4V to 1.6V with the SAG devices.

Monte Carlo simulation of SD device results

Working in collaboration with Dr Karol Kalna in the Device Modelling group, Monte Carlo (MC) simulation of the 120nm SD devices as calibrated from device results was performed. Details of the MC simulator are given elsewhere [7.4]. As this project is concerned with the design, fabrication and characterisation of real devices, the details of the simulation process by Dr Kalna are not provided. Instead the relevant results are summarised.

By including the external parasitic resistances extracted from recessed TLM measurements with the intrinsic MC simulations, the output characteristics for the SD device as generated by the simulator are compared with device measurements:

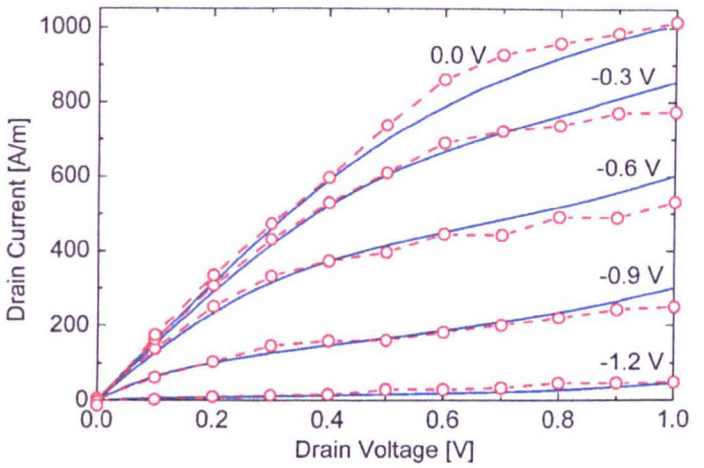


Figure 7.2.12 -120nm SD experimental (blue line) and MC (red circle) output characteristics

Figure 7.2.12 demonstrates the good agreement between experimental and simulated SD output results. The simulator was then used to model the carrier velocity characteristics through the gate region of the device.

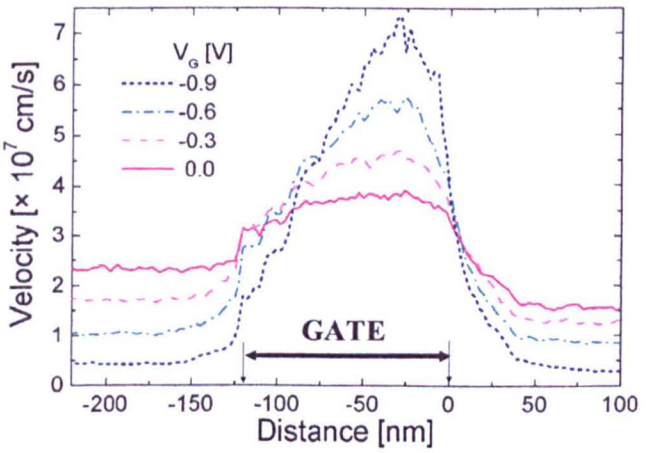


Figure 7.2.13 - 120nm SD carrier velocity characteristics for $V_{ds} = 1.2V$

The carrier velocity distribution for stepped gate bias at a constant source - drain bias of 1.2V is shown in **Figure 7.2.13**. As expected this steadily increases with more negative gate bias

and peaks towards the drain side of the gate due to the concentrated electric field in this region.

Having developed an accurate MC simulator for the 120nm SD devices, it was possible to re-simulate device characteristics after the removal of the additional layer of delta doping. This allowed comparison between simulated single and delta doped standard devices:

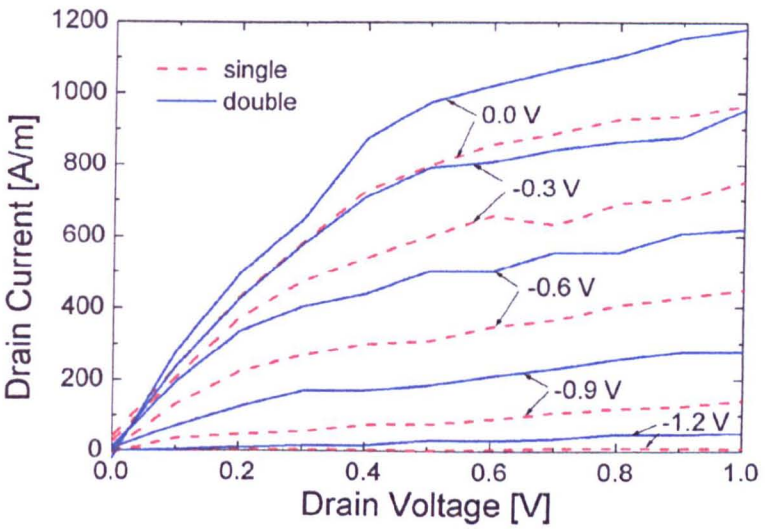


Figure 7.2.14 - 120nm SD single (red) and double (blue) delta doped output comparison

The benefits of the additional layer of delta doping in the standard device are clearly illustrated by the comparison presented in **Figure 7.2.14**. A significant increase of ~200A/m in I_{dss} is observed in addition to an increase in transconductance as a result of introducing the additional doping. This comparison of course also considers the external resistances for the two types of device to be equal. However as has been discussed, the additional layer of delta doping acts to reduce the access resistances through reducing the parallel resistances between ohmic and gate regions. The characteristics of the single doped device would therefore be improved beyond those predicted in **Figure 7.2.14** by the additional doping.

7.3 120nm Self-aligned and Standard RF Characterisation

Similar to the RF measurement procedure used for the characterisation of the 1408 self-aligned pHEMT devices, the InP double delta doped SD, SA and SAG devices were measured at V-band (40MHz to 60GHz) with varied source-drain and gate biases. Due to the similarity between the SA and SAG devices at DC and at RF, for the discussion of their RF performance they will be treated as one type of device. From extracted S-parameters an embedded equivalent circuit for each type of device was then constructed. The S-parameters generated by the embedded SA device model are compared with those extracted experimentally to verify the accuracy of the model:

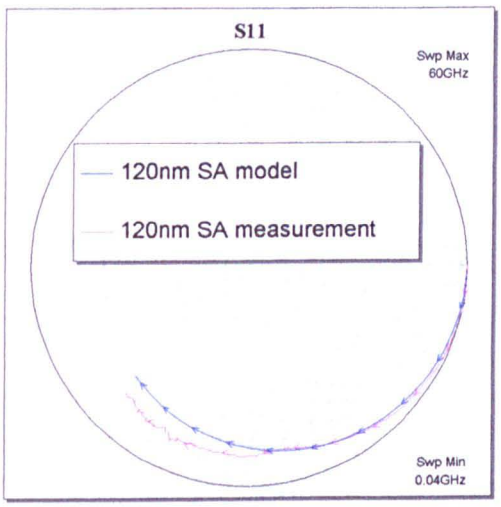


Figure 7.3.1a - S11 magnitude and phase (smith)

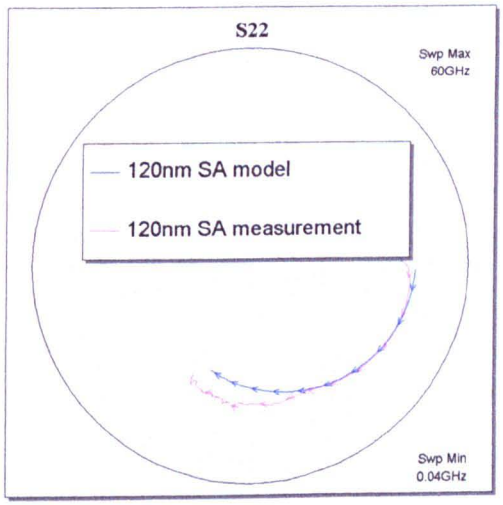


Figure 7.3.1b - S22 magnitude and phase (smith)

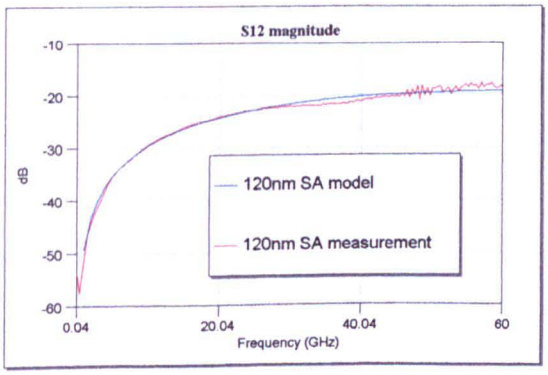


Figure 7.3.1c - S12 magnitude

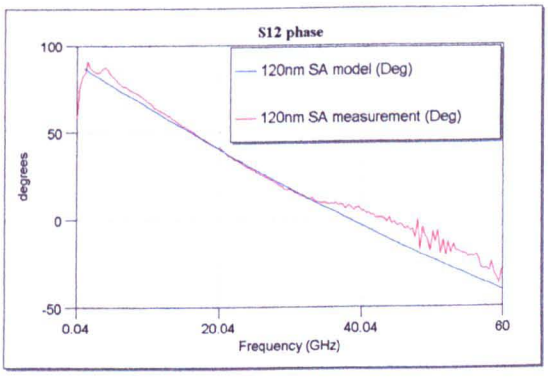


Figure 7.3.1d - S12 phase

Figure 7.3.1 - 120nm SA 2x12.5μm device measured (red) and model (blue) S-parameters

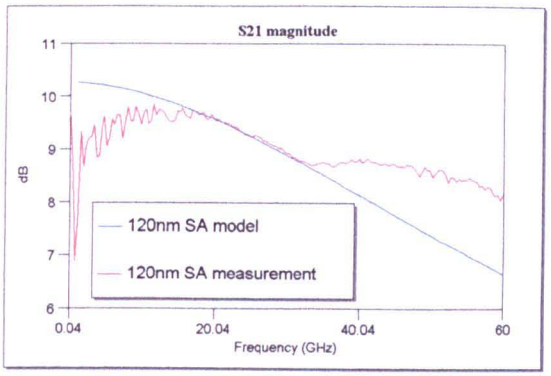


Figure 7.3.1e - S21 magnitude

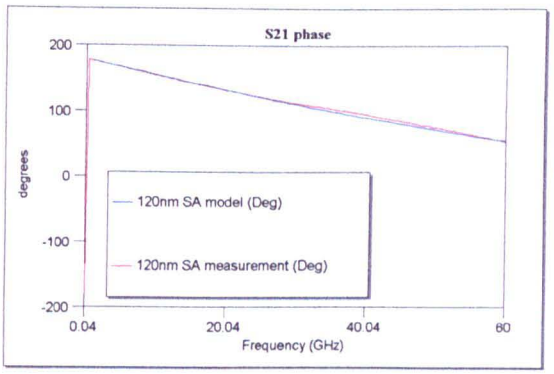


Figure 7.3.1f - S12 phase

Figure 7.3.1 - 120nm SA 2 x 12.5μm device measured (red) and model (blue) S-parameters

The measured S-parameters presented in **Figure 7.3.1** were taken from a 2 x 12.5μm device biased at 1.2V V_{ds} & -0.5V V_{gs}. Good agreement is found between the measured and modelled S-parameter phase, however discrepancies occur in the magnitude of S11, S22 and S21. These occur mainly as a kink in the normally smooth decay of the magnitude of each at a frequency of 33GHz in each case. It is believed that this effect is an artefact of the measurement process as for V-band measurements an RF switch internal to the system activates at this frequency. In addition this effect was observed across the range of bias conditions applied, as well as after separate calibration processes. Beyond stating that this results from the measurement procedure, and is not the real response of the devices beyond 33GHz, it is difficult to explain the cause of this process. However, by selecting sensible element values within the equivalent circuit, the response as shown in **Figure 7.3.1** which matches the S-parameters at frequencies from 15 - 33GHz, is generated. For example, values for the intrinsic device elements can be estimated to a certain degree from previously determined device models. Also, only a substantial reduction in the intrinsic gate capacitance elements C_{gs} or C_{gd} would account for magnitude of these parameters above 33GHz. The result of reducing these element values on the magnitude of S21 is shown in **Figure 7.3.2**:

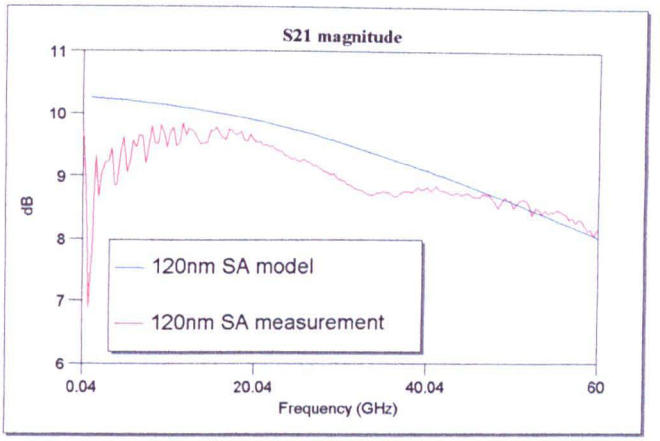


Figure 7.3.2 - S21 magnitude with reduced model C_{gs} & C_{gd}. Measured (red) / model (blue)

Not only does this reduction produce an unrealistically small gate capacitance for a 120nm gate length, but also destroys the fit between the model and measurements for lower frequencies that cannot be accounted for through adjustment of other elements. It is therefore argued that the model presented which fits the frequency range 15 - 33 GHz best describes the real device operation. Below this frequency range the magnitude of S21 and to a lesser degree S22 is lower than that predicted by the model. This artefact has been seen across many RF measurements of devices and although is not completely understood is believed to be related to a delay in the effective source-drain biasing of the device.

Using a similar argument, an equivalent circuit was constructed for the SD device which exhibited exactly the same deviation in S-parameter data from modelled and measured results observed with the SA device. The de-embedded circuit elements for each are presented in **Figure 7.3.3**:

Parameter	SD	SA	Parameter	SD	SA
C_{gs}	18 fF	18.6 fF	R_i	0.1 Ω	0.1Ω
C_{gd}	4.6 fF	4.3 fF	C_{gsp}	2.82 fF	3.02 fF
C_{ds}	9.4 fF	10.4 fF	C_{gdp}	1.94 fF	1.94 fF
g_m	32.9 mS	39.4 mS	C_{dsp}	2.1 fF	2 fF
R_{ds}	387 Ω	342 Ω	L_g	1 pH	1 pH
R_d	1.3 Ω	1.3 Ω	L_s	8.8 pH	8.8 pH
R_s	1.26 Ω	1.26 Ω	L_d	1 pH	1 pH
R_g	10.35 Ω	8.15 Ω			

Figure 7.3.3 - 120nm 2x12.5μm SD & SA equivalent circuit parameters

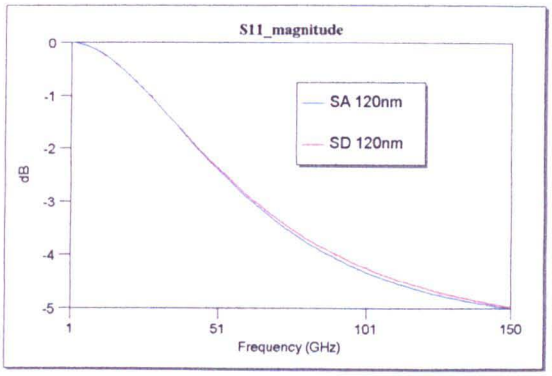


Figure 7.3.4a - SD(red) & SA(blue) S11 magnitude

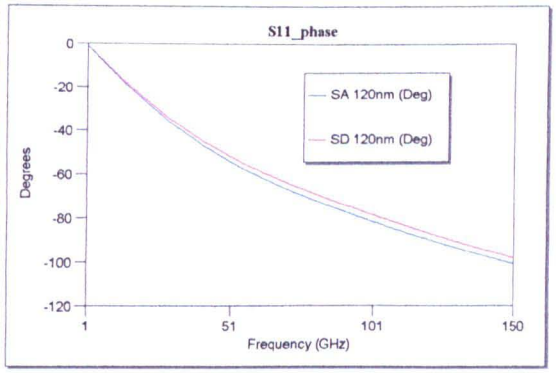


Figure 7.3.4b - SD(red) & SA(blue) S11 phase

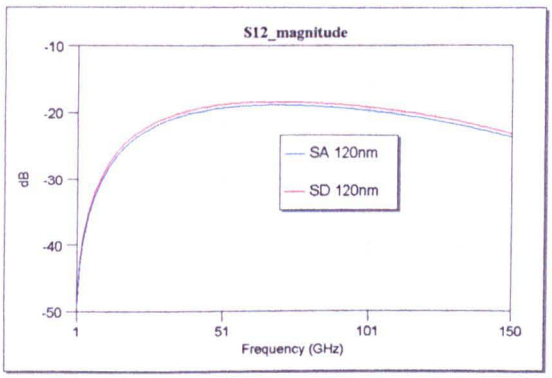


Figure 7.3.4c - SD(red) & SA(blue) S12 magnitude

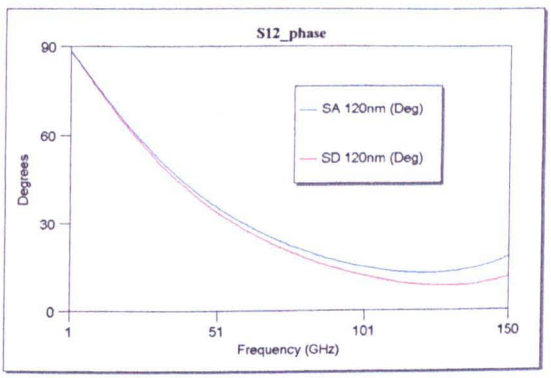


Figure 7.3.4d - SD(red) & SA(blue) S12 phase

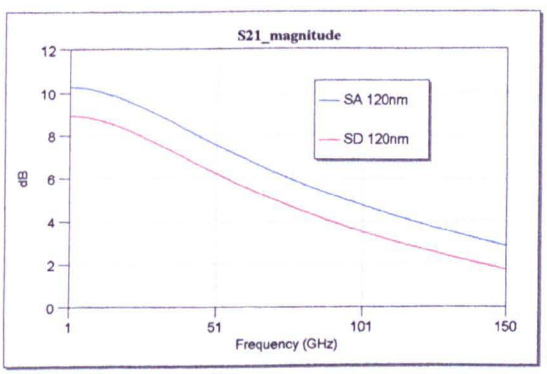


Figure 7.3.4e - SD(red) & SA(blue) S21 magnitude

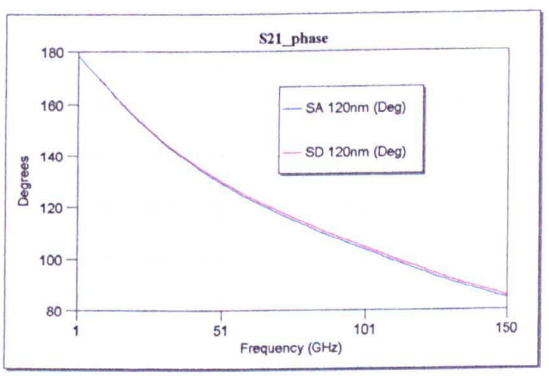


Figure 7.3.4f - SD(red) & SA(blue) S21 phase

Figure 7.3.4 - 120nm SD (red) and SA (blue) de-embedded S-parameters

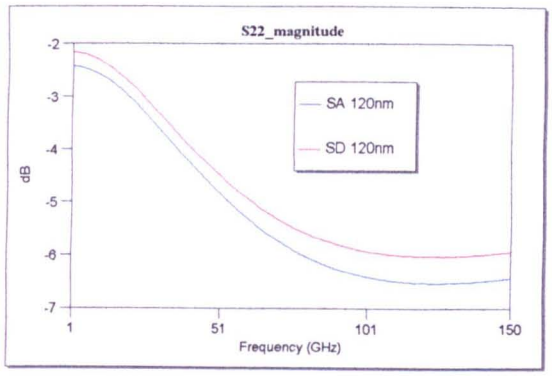


Figure 7.3.4g - SD(red) & SA(blue) S22 magnitude

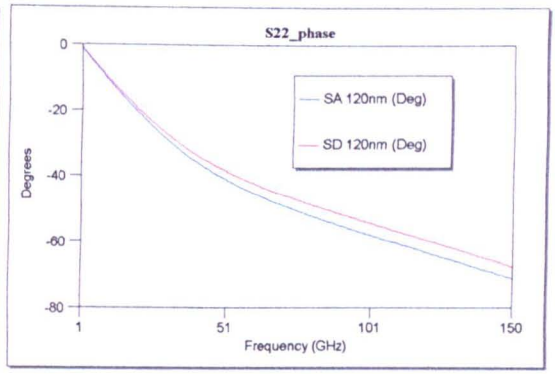


Figure 7.3.4h - SD(red) & SA(blue) S22 phase

Figure 7.3.4 - 120nm SD (red) and SA (blue) de-embedded S-parameters

Comparison of the S-parameters generated by the SD and SA de-embedded equivalent circuits given in **Figure 7.3.4** indicates a close resemblance across the range of magnitude and phase plots with the largest deviation occurring in the S21 and S22 magnitude plots. This difference is easily justified between the SD and SA devices by the difference in the transconductance and output conductance values used in the models that result from matching with measured S-parameters. Comparison of the de-embedded circuit elements also indicates similar intrinsic device values between the SD and SA device with a slight increase in C_{ds} as expected with the SA due to the increased proximity between the source and drain regions. A reduction in R_g is also observed with the SA device due to the additional metal added to the gate head as a side effect of the self-aligned process.

Generation of the H21 plot then provided a method of extracting a figure for f_T for each device:

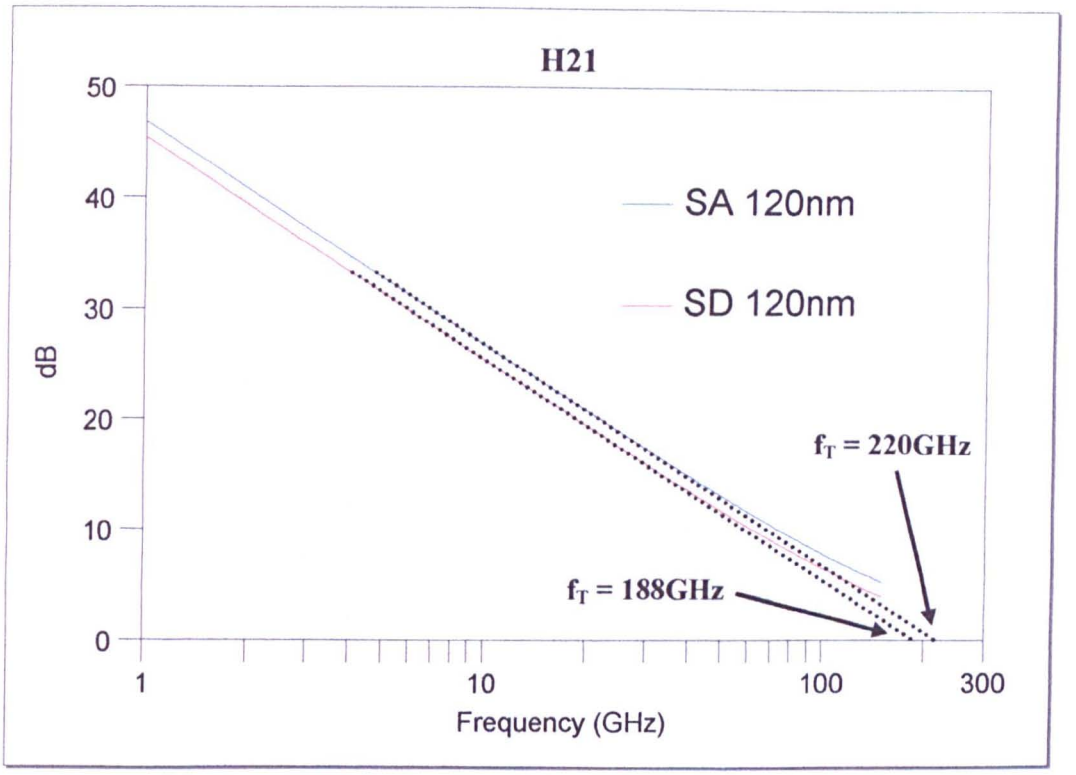


Figure 7.3.5 - H21 for 120nm 2x12.5μm SD (red) and SA (blue) devices

Again the benefits of the self-aligned structure are highlighted in these results through a substantial increase in f_T from 188GHz for the standard device to 220GHz for the self-aligned device. This corresponds to an increase of 17% and most likely results from the substantial increase in the transconductance observed between the SD and SA DC characteristics.

Similarly f_{max} figures for both types of device are extracted from MAG plots generated by their equivalent circuits:

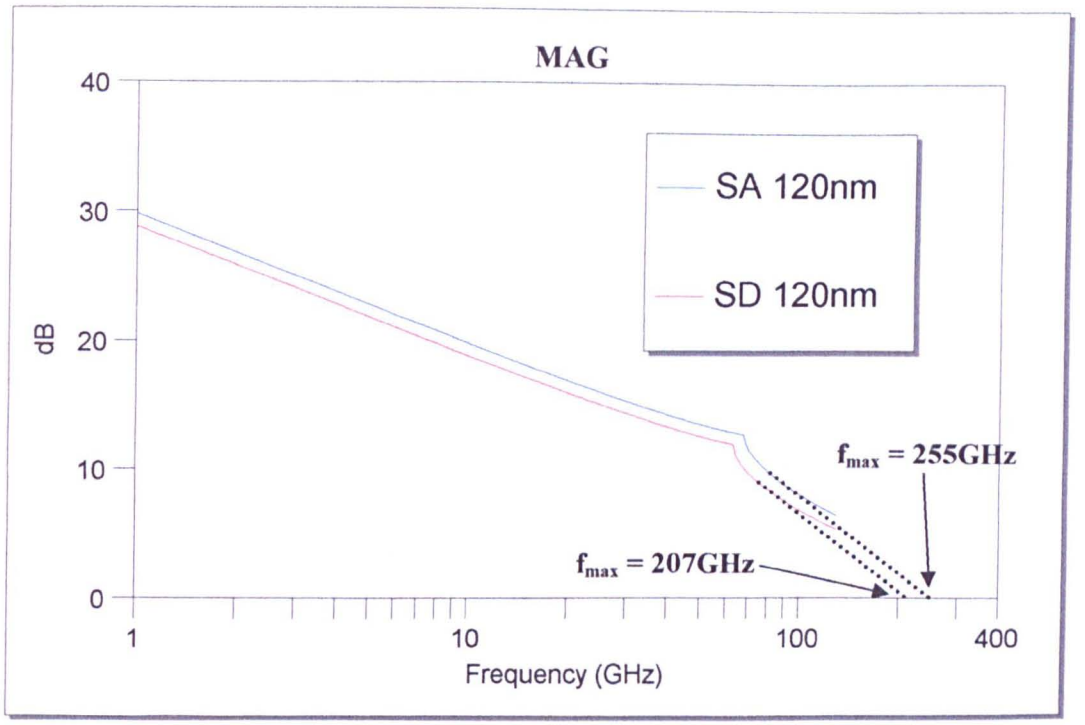


Figure 7.3.6 - MAG for 120nm 2x12.5μm SD (red) and SA (blue) devices

Through linear extraction from the MAG response for both types of device assuming a decay of 20dB/decade within the unconditionally stable region, figures of f_{max} = 207GHz and 255GHz for SD and SA devices are calculated respectively. This substantial increase in the maximum frequency between the two types device arguably results from a combination of a higher f_T figure and a reduced gate resistance with the SA device.

7.4 70nm Self-aligned and Standard Devices

To further investigate the potential increase in device performance at shorter gate lengths through a self-aligned methodology, self-aligned and standard devices of 70nm gate length were fabricated using the same InP double delta doped material as used with the 120nm self-aligned and standard devices. Due to the lack of variation between SA and SAG 120nm devices, only 70nm SD and SA devices were fabricated for comparison. The process used to fabricate these devices was identical to that used for the 120nm, with the exception of the lithography of the gate level which used a different resist and ebeam strategy for the production of the 70nm gate profile as discussed in **Section 3.5.4**. In addition, a double gate recess process was developed in an attempt to minimise the effective gate length of the 70nm gate as well as maximising drain current through the devices.

By developing a multi stage etch process comprising of succinic and orthophosphoric etch steps, a double recess profile can be achieved with the InP HEMT material structure:

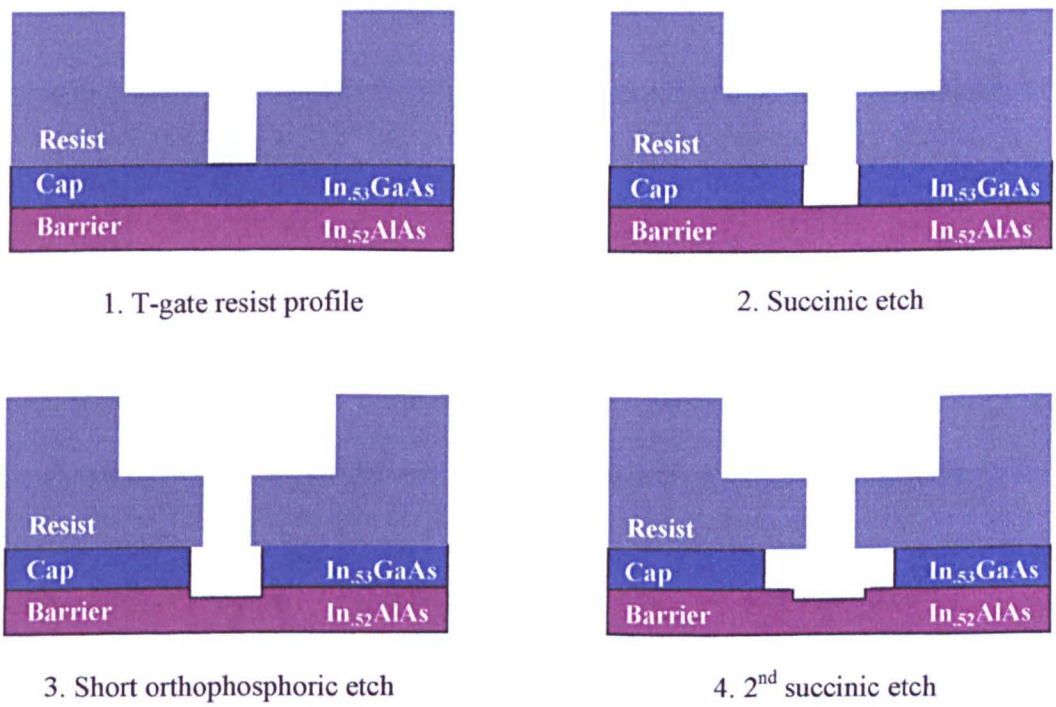


Figure 7.4.1 - 70nm double gate recess process

Following the 70nm T-gate resist profile definition, a short succinic acid etch is used to remove the cap layer below the gate foot. The selectivity of the etch ensures that only the cap is etched and by minimising the etch time little undercut occurs beneath the resist profile. A short non-selective orthophosphoric etch then acts to enlarge this initial etch profile by etching slightly into the barrier layer and laterally into the cap layer. A final succinic acid etch acts to selectively remove more of the cap layer, separating the smaller etch into the barrier layer from the cap. This process ensures a minimum recess length as defined by the etch into the barrier layer, whilst ensuring the deposited gate metal does not come into contact with the highly conducting cap layer.

The results of this process are shown in **Figure 7.4.2**:

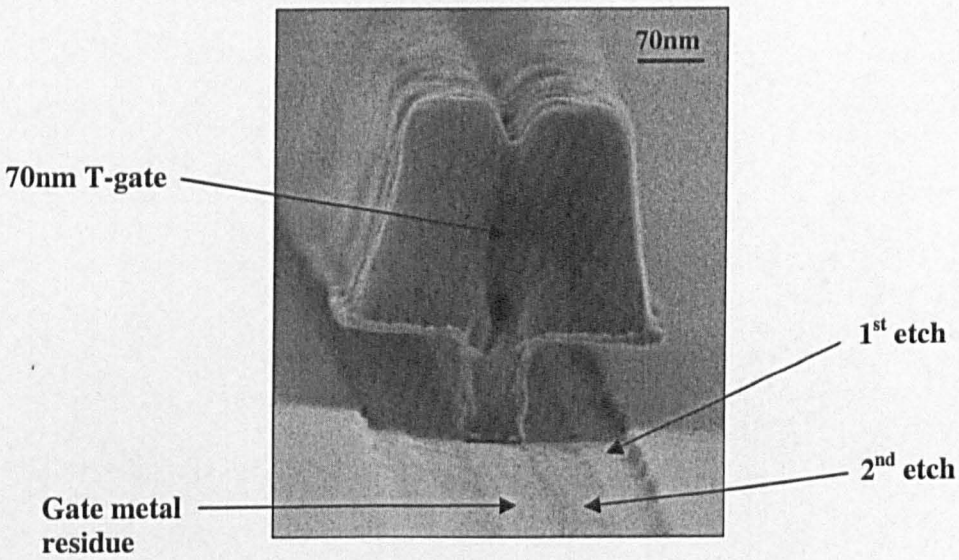


Figure 7.4.2 - 70nm T-gate with double recess

The SEM image in **Figure 7.4.2** clearly shows the double etch with which the cap is etched further laterally than the 2nd etch into the barrier layer. In this test the dimensions of the etch are exaggerated for clarity by overetching. From the figure the cap appears to have been etched ~140nm laterally from the gate stalk, while the 2nd etch appears to have gone ~70nm. In a real device, these dimensions are tailored to be more suitable and for the 70nm device are chosen to be in the range 100nm for the cap etch and 15 - 20nm for the Barrier etch.

These dimensions are verified through AFM scanning of the etch profile:

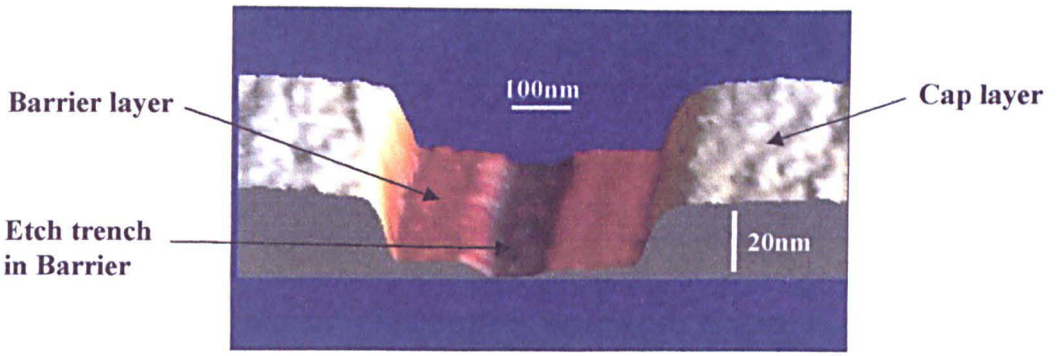


Figure 7.4.3 - AFM profile of optimised 70nm double recess

Figure 7.4.3 clearly shows the desired etch profile with a narrow 100nm long etch trench within the barrier layer, providing a 15nm recess with a 70nm gate length. The cap layer is also etched sufficiently far to avoid shorting between the cap and gate metal. The 2nd etch in this instance was measured using the AFM to be 5nm deep into the barrier layer compared with the initial 20nm etch to remove the cap layer. This corresponds to a double recess ratio of 1:4. The details of this etch process along with the complete 70nm SD/SA device fabrication process are given in **Appendix AIV**.

70nm Self-aligned (SA) and Standard Device (SD) characterisation

Upon D.C. measurement of the range of devices fabricated, the following figures for device yield were determined for SD and SA devices of various widths:

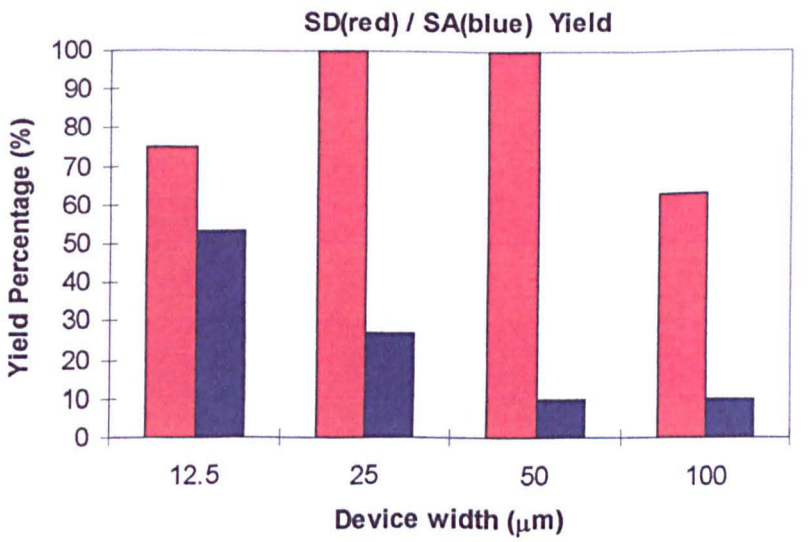


Figure 7.4.4 - 70nm SD (red) & SA (blue) device yield

Again excellent yield of the SD devices was attained with 100% of the 25μm and 50μm wide devices operational. These figures drop dramatically with the SA devices as shorting between the gate head and ohmic contacts renders almost all of the wider devices inoperable. Characterisation of the operational SD and SA devices was then performed, again to compare their DC output and transfer/transconductance characteristics.

Similar to that observed with the 120nm SD and SA devices, a slightly lower low field axial resistance (LFAR) is measured with the SA devices than for the SD. A more pronounced kink in the SA characteristics is also observed:

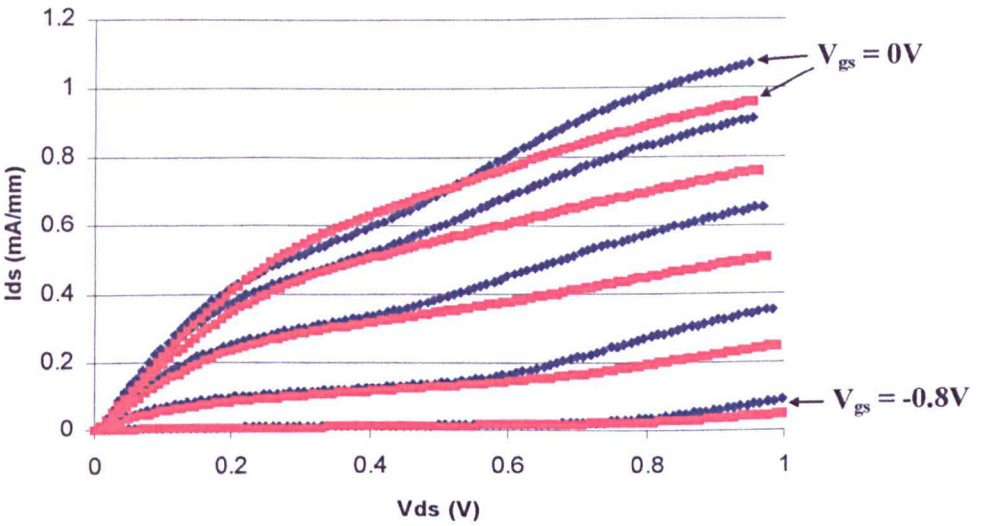


Figure 7.4.5 - 70nm SD (red) and SA (blue) output characteristics comparison

At low source - drain bias and with zero gate bias, the IV response of the SA device is slightly steeper than for the SD device, again reflecting the reduced axial resistance with the self-aligned process. Beyond this initial ohmic response however, a kink again is observed in the SA characteristics that appears to be absent with the SD devices. This process is similar to that seen with the 120nm SD and SA devices except that beyond the kink in the IV characteristics of the 70nm SA devices, the current begins to increase rapidly again and becomes larger than the saturation current for the 70nm SD device. The more pronounced kink in the 70nm SA characteristics that is seen with the 120nm SA devices supports the idea

that hot electron effects distort the ideal IV characteristics of the device under which transport of carriers occurs solely through the channel.

As with the 120nm SD and SA devices, the 70nm transfer and transconductance characteristics are also compared:

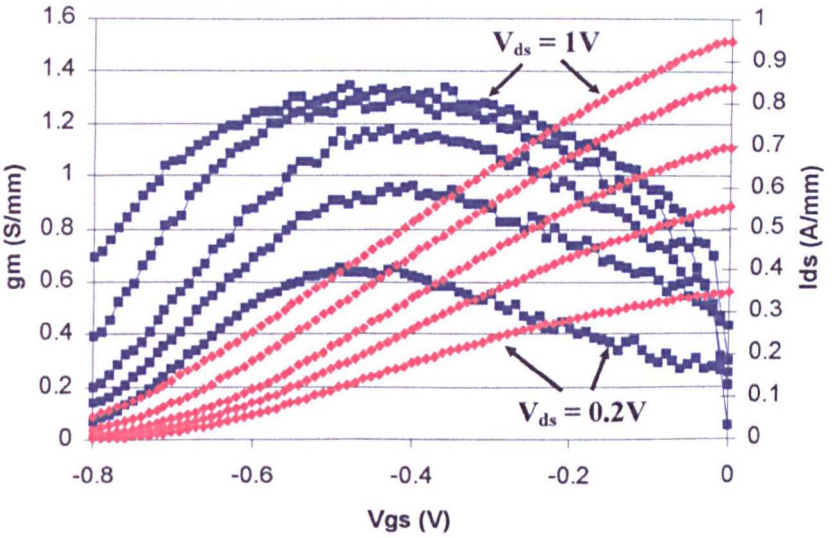


Figure 7.4.6 - 70nm SD transfer (red) and transconductance (blue) characteristics

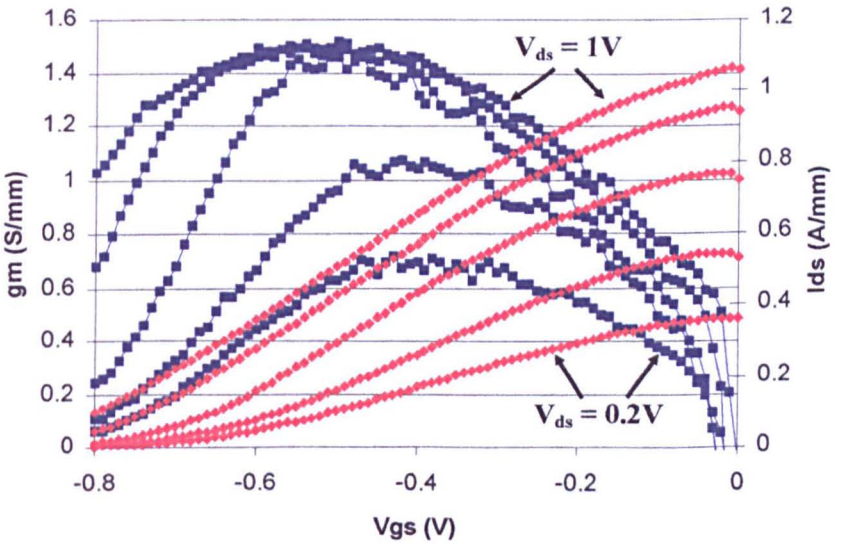


Figure 7.4.7 - 70nm SA transfer (red) and transconductance (blue) characteristics

Again an improvement in the DC performance is observed with the SA device over the SD, though not as pronounced with the 120nm devices. Also the peak transconductance of 1380mS/mm for the SD device occurs at a less negative gate bias than the figure of 1525mS/mm for the SA device. This is the opposite of that observed for the 120nm SD and

SA devices with which a more negative gate bias was required for the SD devices than for the SA, for them to reach their peak transconductance.

The difference in LFAR figures calculated from the 70nm SD and SA output characteristics at zero gate bias again indicates that the intrinsic operation of the two types of device differs. Using the process described by **Eqn 7.2.2**, to account for the increase in peak transconductance from 1380mS/mm to 1525mS/mm, the source resistance must vary by 0.067Ω.mm between the SD and SA device. However from the LFAR figures, only a difference of 0.03Ω.mm in the source resistance can be accounted for.

Comparison between the 120nm and 70nm devices indicates that SA devices at both gate lengths demonstrated similar DC characteristics, including transconductance figures in the range of 1500mS/mm. The kink effect observed in both device profiles however is more exaggerated with the 70nm SA devices. For the 70nm and 120nm SD devices, a significant increase of 28% in the transconductance is observed in the reduction of the SD device gate length from 120nm to 70nm. The saturation of the SA DC performance with respect to these gate lengths is open to speculation. However, given that devices of both gate lengths were fabricated from the same material system, the lack of scaling for the 70nm device could be hindering its potential performance, (as the material was designed for the 120nm node). Another factor of influence is the adoption of the double recess process used for both SD and SA 70nm devices. Although not seen with the SA devices, this process could be responsible for the significant performance enhancement observed between the 120nm and 70nm SD devices. Ideally to test this hypothesis, 120nm devices should be fabricated using an equivalent double recess process and 70nm using a single recess process to truly identify the influence of the double recess on the performance of the various devices. Unfortunately there was insufficient time to perform this additional research, however RF characterisation provides further insight into the benefits of the self-aligned system.

Measured multi-bias S-parameter data from both 70nm SD and SA devices indicated a similar trend in the magnitude of S11, S21 and S22 as for the 120nm devices in the form of a kink in their decay occurring at 33GHz.

Again a model was constructed which accurately emulated the S-parameter data below this frequency. The accuracy of this model is again argued from the inability to devise an accurate equivalent circuit to emulate the higher frequency response with sensible circuit values.

A comparison between the measured and modelled S-parameter data for a 2x12.5μm SA device is presented in **Figure 7.4.8**:

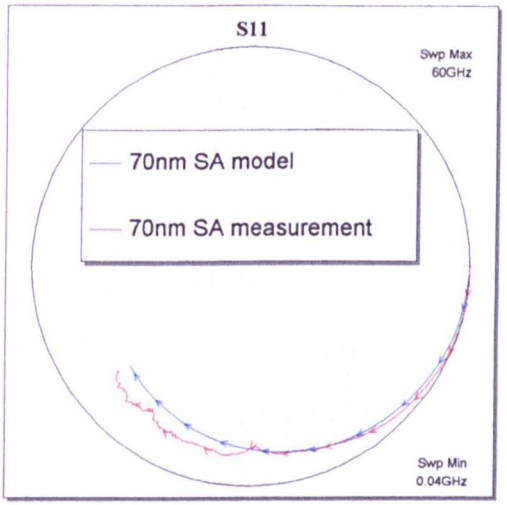


Figure 7.4.8a - S11 magnitude and phase (smith)

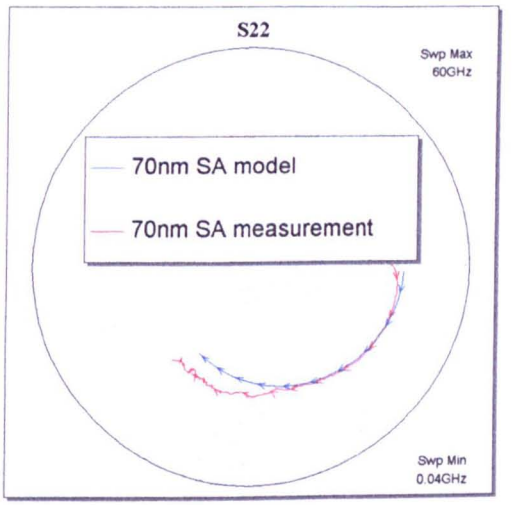


Figure 7.4.8b - S22 magnitude and phase (smith)

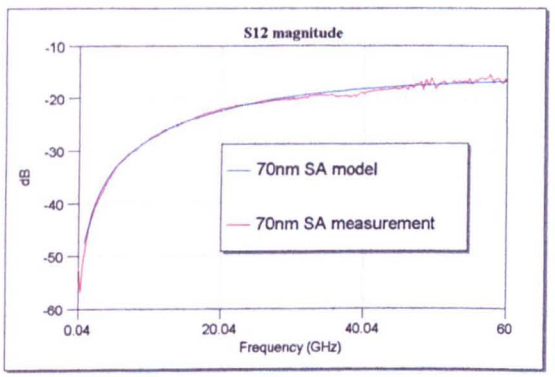


Figure 7.4.8c - S12 magnitude

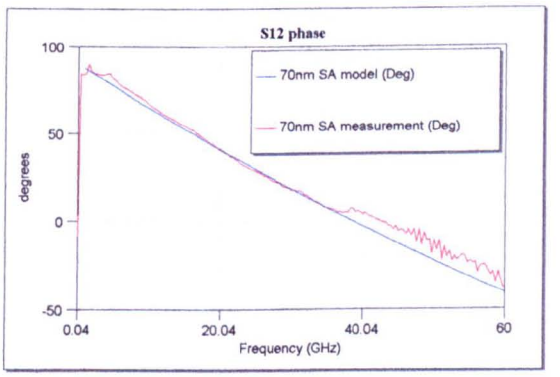


Figure 7.4.8d - S12 phase

Figure 7.4.8 - 70nm SA 2 x 12.5μm device measured (red) and model (blue) S-parameters

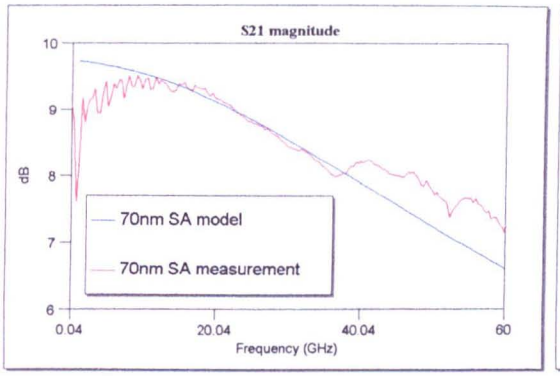


Figure 7.4.8e - S21 magnitude

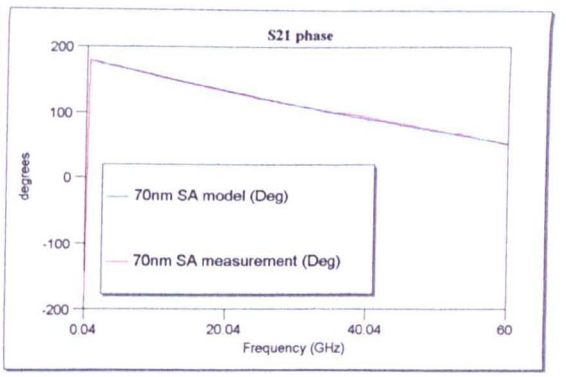


Figure 7.4.8f - S21 phase

Figure 7.4.8 - 70nm SA 2 x 12.5μm device measured (red) and model (blue) S-parameters

The S-parameters presented in **Figure 7.4.8** demonstrate the good agreement between the embedded model and measurements from the 2x12.5μm device biased at 1.1V V_{ds} & -0.5V V_{gs}, again within the frequency range 15 to 33GHz. A similar equivalent circuit was constructed for the SD device from measured data:

The element values for the de-embedded SD and SA models are presented in **Figure 7.4.9:**

Parameter	SD	SA	Parameter	SD	SA
C_{gs}	12.4 fF	9.5 fF	R_i	0.1 Ω	0.1 Ω
C_{gd}	5.1 fF	4.6 fF	C_{gsp}	3 fF	5 fF
C_{ds}	9.1 fF	7.9fF	C_{gdp}	1 fF	3.3 fF
g_m	34.8 mS	38.3 mS	C_{dsp}	2fF	2 fF
R_{ds}	279 Ω	267 Ω	L_g	15.3 pH	15.3 pH
R_d	1.3 Ω	1.3 Ω	L_s	8.8 pH	8.8 pH
R_s	1.26 Ω	1.26 Ω	L_d	1 pH	1 pH
R_g	3.2 Ω	1.9 Ω			

Figure 7.4.9 - 70nm 2x12.5μm SD & SA equivalent circuit parameters

The S-parameters as generated by each of these de-embedded models are then compared:

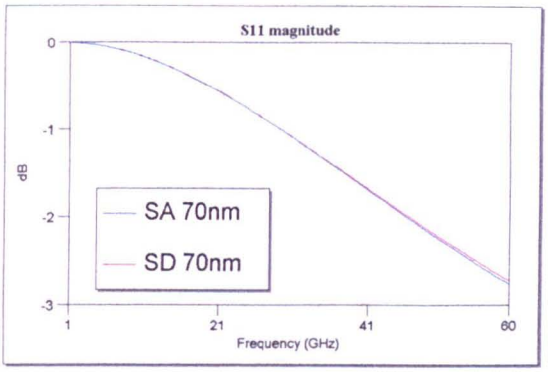


Figure 7.4.10a - SD(red) & SA(blue) S11 magnitude

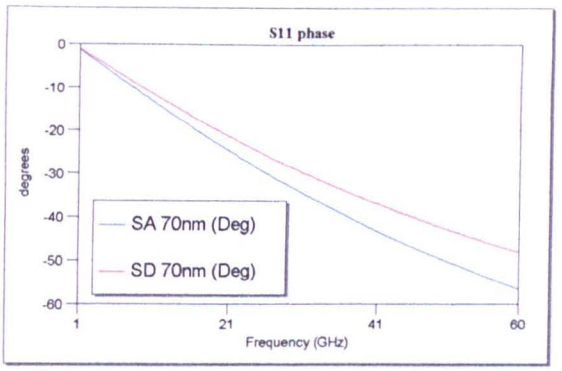


Figure 7.4.10b - SD(red) & SA(blue) S11 phase

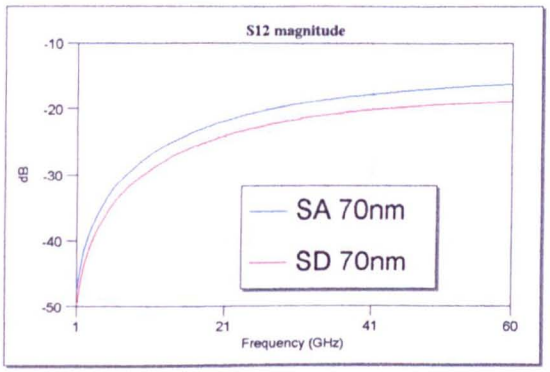


Figure 7.4.10c - SD(red) & SA(blue) S12 magnitude

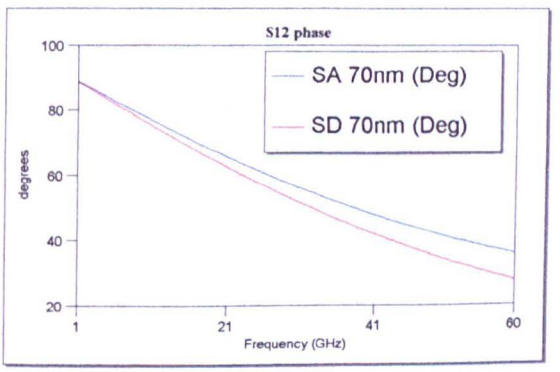


Figure 7.4.10d - SD(red) & SA(blue) S12 phase

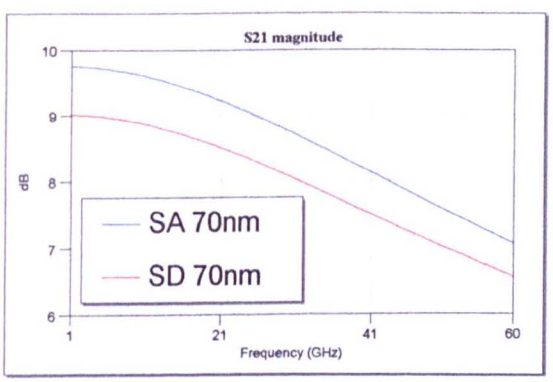


Figure 7.4.10e - SD(red) & SA(blue) S21 magnitude

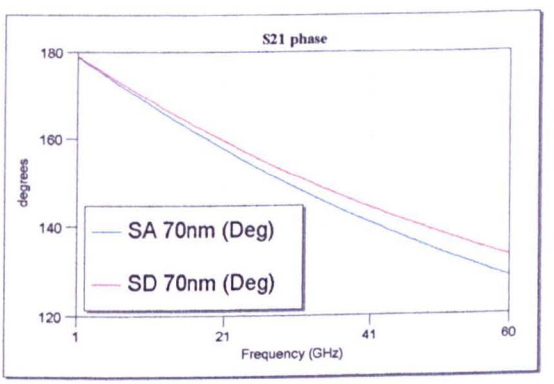


Figure 7.4.10f - SD(red) & SA(blue) S21 phase

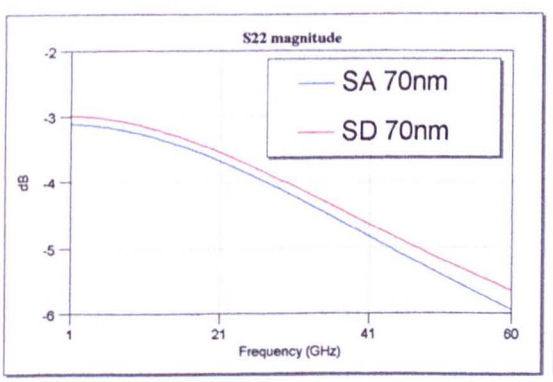


Figure 7.4.10g - SD(red) & SA(blue) S22 magnitude

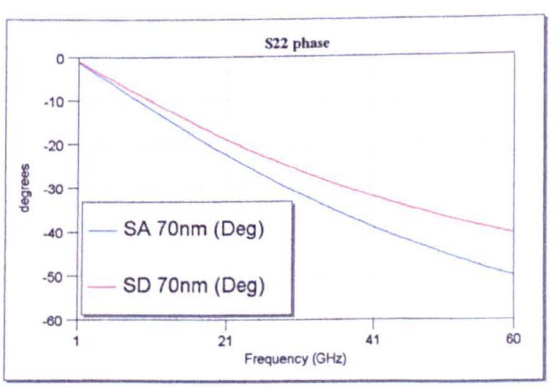


Figure 7.4.10h - SD(red) & SA(blue) S22 phase

Again the smaller output conductance and transconductance of the SD device compared with the SA is demonstrated by the difference in the magnitude of S22 and S21 respectively. At this geometry an increase is seen in the parasitic capacitances C_{gdp} and C_{gsp} with the SA device. This is to be expected as the capacitance effects through the air between the gate and source/drain contacts will be increased due to the closer proximity of the ohmic contacts and gate. This difference in parasitic capacitance values leads to a deviation in the phase across the four S-parameters for the two types of device, in addition to increasing the magnitude of S12 for the SA device.

A decrease in the intrinsic gate capacitances C_{gs} and C_{gd} is also observed between the SD and SA devices. This result is more difficult to explain but could arguably be related to the difference in the dimensions of the depletion region formed in each type of device, accentuated by the double recess process.

Generation of the H21 response then allowed the calculation of the f_T figures for each type of device:

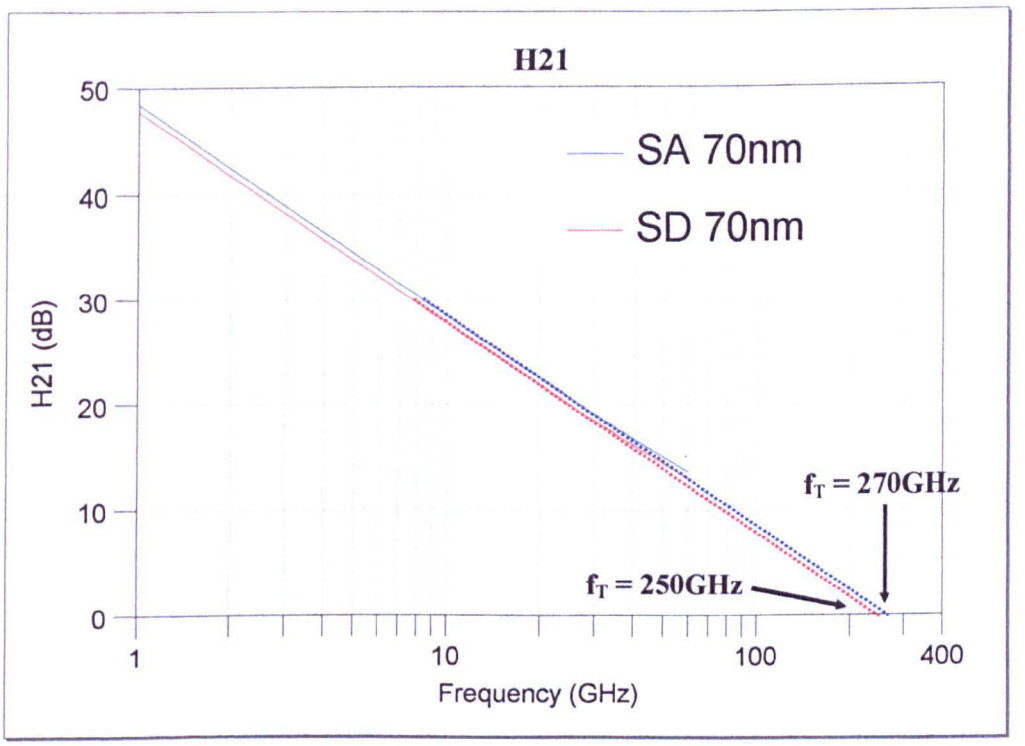


Figure 7.4.11 - De-embedded H21 plot for 70nm SA (blue) and SD (red) devices

And similarly the MAG response for the calculation of f_{max} figures:

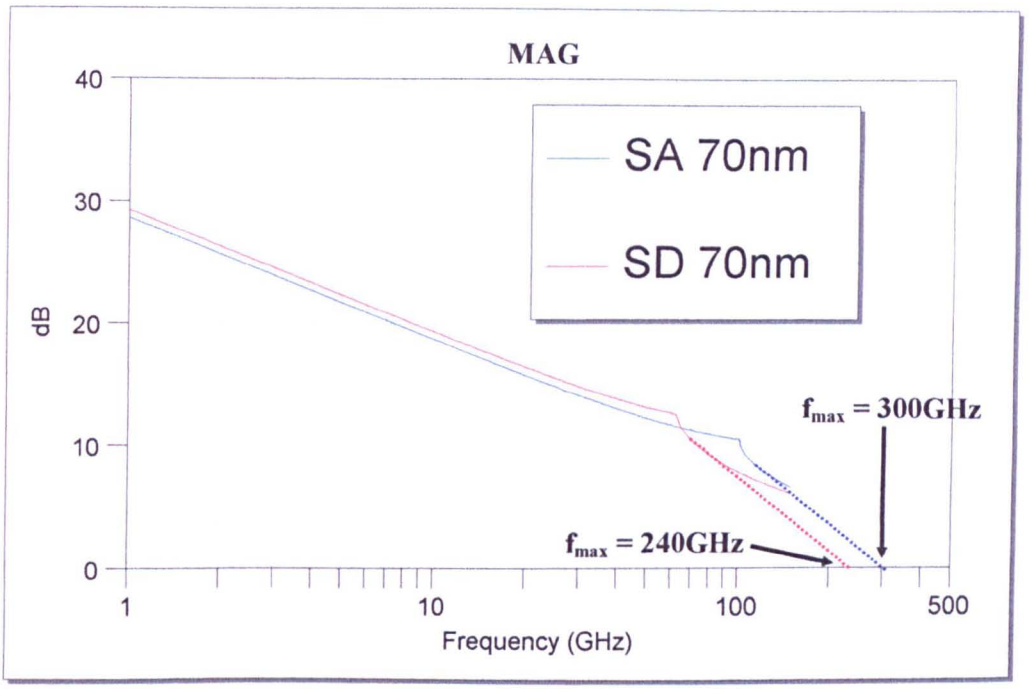


Figure 7.4.12 - De-embedded MAG plot for 70nm SA (blue) and SD (red) devices

Figures of f_T and f_{max} as extrapolated from H21 and MAG plots indicate that as with the 120nm devices, an increase in performance is observed when moving to a self-aligned process. From the examples shown in **Figures 7.4.11** and **7.4.12** this relates to an increase of 8% in f_T and a large increase of 25% in f_{max} between the SD and SA devices. This performance improvement can be compared with that seen between the 120nm SD and SA devices at DC and RF:

	% increase between 120nm SD and SA	% increase between 70nm SD and SA
DC transconductance	38%	11%
Cut-off frequency f_T	17%	8%
Maximum frequency f_{max}	23%	25%

Figure 7.4.13 - Performance increase between SD and SA devices (120nm and 70nm)

In general the largest performance increase seen between SD and SA devices is at the 120nm node with a substantial increase observed in the transconductance and cut-off/maximum frequency figures. Although improvement is seen at the 70nm node, with the exception of the maximum frequency, the performance increase does not match than seen with the 120nm devices. This result might at first appear to contradict the theory that the benefits of the self-aligned process should become more apparent at shorter gate lengths. However, other factors have to be considered beyond the reduction of the parasitic resistances in these two gate length systems. Firstly, the material used for devices of both gate lengths was designed and scaled for 120nm device operation, in which case the gate-channel separation might be too large for optimum gate control with the 70nm devices. Secondly, by introducing a double recess process for the 70nm devices an additional variable is introduced between the 120nm and 70nm devices, the exact benefits of which are not known. As mentioned this process could prove more beneficial to the SD device structure than for the SA, reducing the performance increase seen between the two. However without testing of both single and double recess processes at each gate length this effect cannot be verified. Beyond these issues, there also exists the influence of the parasitic capacitances between the gate and ohmic contacts which become more apparent in the 70nm SA device. As the gate length of the SA device is continuously reduced, reducing the intrinsic gate capacitance, the magnitude of these parasitics will remain relatively constant and hence this lack of parasitic scaling would begin to dominate performance at shorter gate lengths.

For example, if the parasitic capacitance values for C_{gdp} and C_{gsp} within the 70nm SA de-embedded model are replaced with those from the 70nm SD model, the figure for f_T for the 70nm SA device is increased to 330GHz. f_{max} remains similar after the reduction of these parasitic capacitances. This example highlights the importance of the magnitude of these parasitics as devices are continuously scaled to reduced gate lengths, as even a few extra femtofarads will severely degrade device performance.

With these arguments, the incorporation of a self-aligned gate process for ultra-small gate length devices might seem futile as the benefits in device performance from reduced parasitic resistances are undermined by increased parasitics capacitances. Potentially however, a compromise could be reached with which these resistances are reduced from that of the standard HEMT structure, whilst minimising the parasitic capacitances between the gate and ohmic contacts. A potential advanced self-aligned T-gate process that addresses this issue is presented in **Figure 7.4.14**:

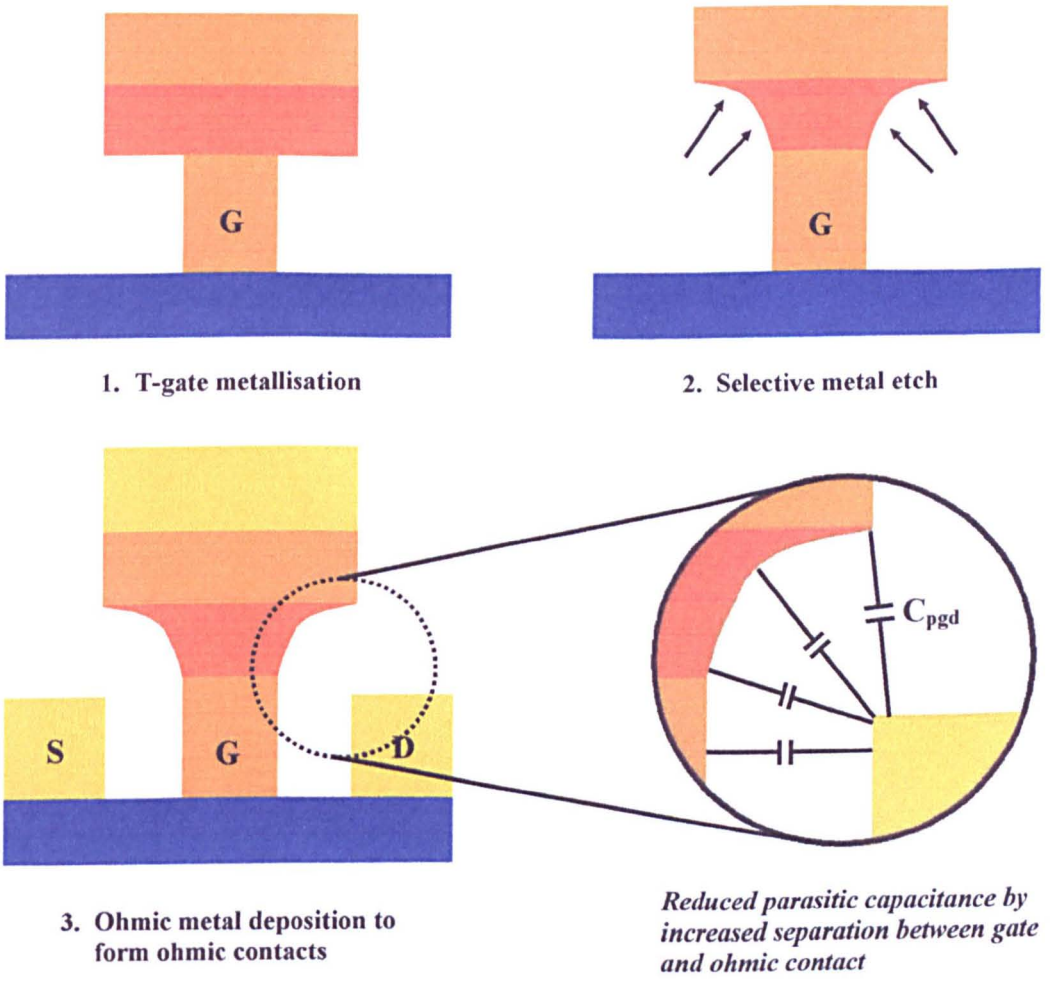


Figure 7.4.14 - Etched self-aligned T-gate process

From **Figure 7.4.14**, a specific gate metallisation strategy is chosen to allow the selective etching of the middle gate metal layer and hence reduce the separation between subsequent ohmic metallisation and gate head, thereby reducing associated capacitances. This process would also improve the poor mechanical yield associated with the self-aligned process by

reducing the likelihood of shorting between gate head and ohmic metallisation. Although this process has not as yet been put into practice, if suitable metallisation and etch processes could be adopted, the yield and performance of the standard self-aligned process described in this work could potentially be improved dramatically. Unfortunately time was not available to pursue this concept further.

7.5 Chapter Summary

Following the development of a self-aligned gate process for the fabrication of 120nm GaAs pHEMT devices, an improved material structure based upon a lattice matched InP architecture was designed. The incorporation of an additional layer of delta doping ensured the maximisation of conduction vertically through the structure for use with a non-annealed process and hence provided a reduced ohmic contact resistance to that for the GaAs pHEMT process. The development of a suitable PMMA/Co-poly based 120nm Γ-gate process together with the original 120nm T-gate process allowed the fabrication of standard (SD), self-aligned T-gate (SA) and self-aligned Γ-gate (SAG) devices for comparison. The benefits of the self-aligned process were apparent, even at the 120nm node through increased performance at both DC and RF. Discrepancies between the external resistance figures predicted from transconductance and IV characteristic comparison for the SD and SA devices suggested the intrinsic operation of the two differed. This idea was supported by a kink effect observed in the SA output characteristics indicating the possible existence of hot electron effects within the SA device.

Using a Γ-gate aspect ratio of 3:1, little difference was observed between the SA and SAG devices beyond a slight increase in the breakdown voltage. This was difficult to confirm however from the few measurements taken from the limited number of operational SA and SAG devices. The benefits of the presence of the second layer of delta doping alone were then highlighted through Monte Carlo simulation of the SD devices. Comparison between equivalent single and double delta doped device Monte Carlo simulations indicated a substantial increase in drive current and transconductance with the double delta doped device, resulting from increased carrier concentration and reduced device access resistances over the single doped device.

Using the same double delta doped material structure, 70nm SD and SA devices were then fabricated using an identical process to that for the 120nm apart from the incorporation of a double recess process at the gate level. Again improvement in device performance was

observed with the self-aligned system, although not as pronounced as with the 120nm devices. This was believed to be due to lack of appropriate scaling of the material for the 70nm node, but also due to the effect of the parasitic gate - ohmic contact capacitances which are exaggerated by the self-aligned architecture and impact device performance more severely at shorter gate lengths.

Although this result suggests the redundancy of the self-aligned device for ultra-short gate lengths, (potentially 50nm and below), it is suggested that tailoring of the self-aligned process to minimise these capacitances whilst maintaining reduced access resistances could still be beneficial to realisation of such ultra-short gate length devices.

Chapter Publications

[1] D. A .J. Moran, K. Kalna, E. Boyd, F. McEwan, H. McLelland, L. L. Zhuang, C. R.Stanley, A. Asenov, I. Thayne, *Self-aligned 0.12μm T-gate $In_{53}Ga_{47}As/In_{52}Al_{48}As$ HEMT technology utilising a non-annealed ohmic contact strategy*, in ESSDERC 2003 p. 315 – 318.

[2] D. A. J. Moran, E. Boyd, K. Elgaid, F. McEwan, H. McLelland, C.R. Stanley, I.G. Thayne, *Self-aligned T-gate InP HEMT realisation through double delta doping and a non-annealed ohmic process*, In print Microelectronic Engineering, 2004

References

[7.1] P. M. Smith, K. Nichols, W. Kong, L. MtPleasant, D. Pritchards, R. Lender, J Fisher, R. Actis, D. Dugas, D. Meharry, A. W. Swanson. *Advances in InP HEMT technology for high frequency applications.* in *Indium Phosphide and Related Materials*, 2001. p. 9 - 14

[7.2] K. Kajiyama, Y. Mizushima, S. Sakata, *Schottky barrier height of n-InxGa1-xAs diodes.* Applied Physics Letters, 1973, 23(8): p. 458 - 459.

[7.3] E. Skuras, G. Pennelli, A. R. Long, C. R. Stanley, *Molecular-beam epitaxy growth of InGaAs--InAlAs high electron mobility transistors with enhanced electron densities and measurement of InAlAs surface potential.* Journal Vacuum Science and Technology, 2001. B19(4): p. 1524 - 1528.

[7.4] K. Kalna, S. Roy, A. Asenov, K. Elgaid, I. Thayne, *Scaling of pseudomorphic high electron mobility transistors to decanano dimensions.* Solid-State Electronics, 2002. 46: p. 631 - 638.

Chapter 8

Conclusions

Throughout this work the role of the parasitic resistances inherent to HEMT design, in the performance of short gate length III-V HEMT technology, has been explored. The motivation to reduce these parasitic elements is initially brought to light through inspection of the theory of operation of such short gate length devices. This highlights the necessity to scale the physical elements of the device, both intrinsic and extrinsic to ensure the maximum performance at a particular gate length.

Although the importance of this process becomes apparent for short gate length devices, little research has been focussed into the reduction of these parasitic resistances and in particular the development of self-aligned gate technologies. Instead, the main routes adopted for the maximisation of HEMT performance have been focussed upon the lithographical challenges associated with reducing the gate length of such devices, combined with using more exotic material structures. In addition, although substantial work into the development of self-aligned gate processes for III-V HEMT has been carried out by Nguyen *et al*, no direct comparison between self-aligned and standard device technology have been reported [8.1 – 8.3].

With this in mind, self-aligned gate technologies compatible with short gate length GaAs pHEMT and lattice matched InP HEMT have been developed with the aim of providing insight into the benefits of the self-aligned gate process upon device performance. This was initially approached by investigating the challenges associated with modifying a standard HEMT process flow for use with a GaAs pHEMT self-aligned process. In particular the ability to produce an efficient ohmic contact without degradation to the gate by thermal annealing was investigated through experimentation with varied ohmic metallisation and annealing strategies. This, combined with a modified succinic acid etch used for the gate recess, allowed the fabrication of 120nm GaAs pHEMT self-aligned devices. Inspection of the performance of these devices highlighted the inefficient operation of the non-annealed ohmic contacts in comparison with similar, non-self-aligned devices.

Analysis of the operation of the self-aligned GaAs pHEMT devices and specifically the non-annealed ohmic process led to the design of a lattice-matched InP HEMT material structure optimised for use with a more efficient non-annealed ohmic and hence self-aligned process. The InP material scaled for 120nm gate length operation was then used to realise both standard and self-aligned devices at 120nm and 70nm nodes. The incorporation of a self-aligned gate process was found to improve device performance substantially at both device gate lengths as verified by DC and RF measurements. However the magnitude of improvement in performance observed between standard and self-aligned devices was found to be significantly larger with the 120nm devices. Although this result is contrary to the premise that a reduction in the parasitic resistances should be more beneficial at shorter gate lengths, it has been argued that a combination of inadequate scaling of the material for use at the 70nm node combined with large parasitic gate capacitances associated with the self-aligned structure restrict the potential of the 70nm self-aligned device.

The dominating effect of these parasitic gate capacitances within ultra-short gate length devices would counteract the benefits of reducing the parasitic resistances and hence render the self-aligned process pointless. However, the increase in device performance as a direct result of reducing the parasitic resistances has been demonstrated throughout this project. It has also been discussed that most of the improvement in device performance, in comparison with a standard device process, results from reduced access resistances using the non-annealed ohmic process. Therefore at reduced gate lengths, where the additional parasitic gate capacitances associated with a self-aligned process will degrade performance, the non-annealed process used with a standard device geometry provides a method of reducing access resistances without concern for increased capacitance parasitics. In addition, the ability to define the gate level prior to the ohmic contacts is a desirable feature for sub100nm gate features, as this eliminates potential uniformity issues with the gate profile due to resist thickness fluctuations that result from spinning of the gate resist between the ohmic contacts. This process is readily achieved with the non-annealed ohmic technology presented throughout this work.

With these issues in mind, it is concluded that for optimum device performance and maximum yield for sub100nm HEMT devices, the following process should be used:

- Use of non-annealed ohmic contacts (double delta-doped)
  - Minimises access resistances leading to increased performance
  - Allows gate level definition prior to ohmics
  - Minimises process thermal budget
- A 'standard' device structure as opposed to self-aligned
  - Minimises parasitic gate capacitances, crucial at shorter gate lengths
  - Increases device breakdown voltage
- Double gate recess process
  - Reduces short channel effects
  - Improves control and pinch-off voltage

Potential Future Work

Based on the research reported in this dissertation, potential further research into the development of such short gate length self-aligned gate III-V HEMT technology should address these issues:

- Further identification of the influence of parasitic gate capacitances at shorter gate lengths through the realisation of self-aligned devices at the 50nm node and below.
- Minimisation of parasitic capacitance elements through an "advanced" self-aligned process in which the gate profile is tailored by selective metal etching.

References

[8.1] U. K. Mishra, A.S. Brown, L.M. Jelloian, M. Thompson, L.D. Nguyen, S.E. Rosenbaum, *Novel high performance self-aligned 0.15 micron long T-gate AlInAs-GaInAs HEMTs.* in *Electron Devices Meeting, 1989. Technical Digest., International.* 1989.

[8.2] L. D. Nguyen, L.M. Jelloian, M. Thompson, M. Lui, *Fabrication of a 80 nm self-aligned T-gate AlInAs/GaInAs HEMT.* in *Electron Devices Meeting, 1990. Technical Digest., International.* 1990.

[8.3] L. D. Nguyen, A.S. Brown; M. A. Thompson, L. M. Jelloian, *50-nm self-aligned-gate pseudomorphic AlInAs/GaInAs high electron mobility transistors.* Electron Devices, IEEE Transactions on, 1992. 39(9): p. 2007-2014.

Appendix A

Fabrication Processes

AI – Standard 120nm GaAs pHEMT process flow

1. Markers

Clean substrate – 5min ultrasonic acetone + 5min ultrasonic IPA

Spin resist – 12% 2010 5k rpm 60s. 1hr 180°C bake. 4% 2041 5k rpm 60s. 2hr 180°C bake

Exposure – Dose 310, Res 0.15, Spot 300

Develop – 1:1 MIBK:IPA 30s @ 23 °C, IPA rinse

Ash – low power O_2 60s

Metallise – 30s 4:1 H_20:HCl de-ox, H_20 rinse, Plassys 20nm Ti, 130nm Au

Lift-off – 2hr warm acetone, pipette clean, IPA rinse

2. Ohmic Contacts

Clean substrate – 5min ultrasonic acetone + 5min ultrasonic IPA

Spin resist – 12% 2010 5k rpm 60s. 1hr 180°C bake. 4% 2041 5k rpm 60s. 2hr 180°C bake

Exposure – Dose 225, Res 0.15, Spot 160

Develop – 1:1 MIBK:IPA 30s @ 23 °C, IPA rinse

Ash – low power O_2 60s

Metallise – 30s 4:1 H_20:HCl de-ox, H_20 rinse, Plassys Mikeohm metallisation

Lift-off – 2hr warm acetone, pipette clean, IPA rinse

Anneal – 20s 300°C - 60s 360°C

3. Mesa Isolation

Clean substrate – 5min ultrasonic acetone + 5min ultrasonic IPA

Spin resist – 12% 2010 5k rpm 60s. 1hr 180°C bake. 4% 2041 5k rpm 60s. 2hr 180°C bake

Exposure – Dose 225, Res 0.15, Spot 300

Develop – 1:1 MIBK:IPA 30s @ 23 °C, IPA rinse

Ash – low power O_2 60s

Isolate – 200:1:1 H_20:H_2O_2:NH_4, H_2O rinse (depth characterisation or electrical depending on ohmic/isolation level order)

Resist removal – 2hr warm acetone, pipette clean, IPA rinse

4. Gates

Clean substrate – 5min ultrasonic acetone + 5min ultrasonic IPA

Spin resist – 4% 2041 5k rpm 60s, 1hr 180°C bake. 9% Coploy 5k rpm 60s, 1hr 180°C bake. 2.5% 2010 5k rpm 60s, overnight 180°C bake

Exposure – Dose 118, Res 0.02, Spot 40 (CFA)

Develop – 2.5:1 MIBK:IPA 30s @ 23 °C, IPA rinse

Recess – Dry etch $SiCl_4/SiF_4/O_2$

Metallise – 10s 10:1 H_20:HF, H_20 rinse Plassys 15nm Ti 15nm Pd 160nm Au

Lift-off – 2hr warm acetone, pipette clean, IPA rinse

5. Bondpads

Clean substrate – 30min warm acetone, IPA rinse

Spin resist – 15% 2010 5k rpm 60s, 1hr 120°C bake. 4% 2041 5k rpm 60s. 2hr 120°C bake

Exposure - Dose 330, Res 0.15, Spot 300

Develop – 1:1 MIBK:IPA 30s @ 23 °C, IPA rinse

Ash – low power O_2 60s

Metallise – 30s 4:1 H_20:HCl de-ox, H_20 rinse Plassys 50nm NiCr 1200nm Au

Lift-off – 2hr warm acetone, pipette clean, IPA rinse

AII – Self-aligned gate 120nm GaAs pHEMT process flow (1408)

1. Markers

Clean substrate – 5min ultrasonic acetone + 5min ultrasonic IPA

Spin resist – 12% 2010 5k rpm 60s. 1hr 180°C bake. 4% 2041 5k rpm 60s. 2hr 180°C bake

Exposure – Dose 310, Res 0.15, Spot 300

Develop – 1:1 MIBK:IPA 30s @ 23 °C, IPA rinse

Ash – low power O_2 60s

Metallise – 30s 4:1 H_20:HCl de-ox, H_20 rinse, Plassys 20nm Ti, 130nm Au

Lift-off – 2hr warm acetone, pipette clean, IPA rinse

2. Mesa Isolation

Clean substrate – 5min ultrasonic acetone + 5min ultrasonic IPA

Spin resist – 12% 2010 5k rpm 60s. 1hr 180°C bake. 4% 2041 5k rpm 60s. 2hr 180°C bake

Exposure – Dose 225, Res 0.15, Spot 300

Develop – 1:1 MIBK:IPA 30s @ 23 °C, IPA rinse

Ash – low power O_2 60s

Isolate – 200:1:1 H_20:H_2O_2:NH_4, H_2O rinse (depth characterisation)

Resist removal – 2hr warm acetone, pipette clean, IPA rinse

3. Gates

Clean substrate – 5min ultrasonic acetone + 5min ultrasonic IPA

Spin resist – 4% 2041 5k rpm 60s, 1hr 180°C bake. 9% Coploy 5k rpm 60s, 1hr 180°C bake. 2.5% 2010 5k rpm 60s, overnight 180°C bake

Exposure – Dose 118, Res 0.02, Spot 40 (CFA)

Develop – 2.5:1 MIBK:IPA 30s @ 23 °C, IPA rinse

Recess – 30s 4:1 H_20:HCl de-ox, 25s Succinic acid (pHEMT), H_2O rinse

Metallise – 10s 10:1 H_20:HF, H_20 rinse, Plassys 15nm Ti 15nm Pd 160nm Au

Lift-off – 2hr warm acetone, pipette clean, IPA rinse

4. Ohmic Contacts

Clean substrate – 5min ultrasonic acetone + 5min ultrasonic IPA

Spin resist – 12% 2010 5k rpm 60s. 1hr 180°C bake. 4% 2041 5k rpm 60s. 2hr 180°C bake

Exposure – Dose 225, Res 0.15, Spot 160

Develop – 1:1 MIBK:IPA 30s @ 23 °C, IPA rinse

Ash – low power O_2 30s

Metallise – 30s 4:1 H_20:HCl de-ox, H_20 rinse, Plassys thin Mikeohm metallisation

Lift-off – 2hr warm acetone, pipette clean, IPA rinse

5. Bondpads

Clean substrate – 30min warm acetone, IPA rinse

Spin resist – 15% 2010 5k rpm 60s, 1hr 120°C bake. 4% 2041 5k rpm 60s. 2hr 120°C bake

Exposure - Dose 330, Res 0.15, Spot 300

Develop – 1:1 MIBK:IPA 30s @ 23 °C, IPA rinse

Ash – low power O_2 60s

Metallise – 30s 4:1 H_20:HCl de-ox, H_20 rinse, Plassys 50nm NiCr 1200nm Au

Lift-off – 2hr warm acetone, pipette clean, IPA rinse

AIII – Self-aligned gate 120nm InP process flow

1. Markers

Clean substrate – 30min warm acetone, IPA rinse

Spin resist – 12% 2010 5k rpm 60s. 1hr 180°C bake. 4% 2041 5k rpm 60s. 2hr 180°C bake

Exposure – Dose 310, Res 0.15, Spot 300

Develop – 1:1 MIBK:IPA 30s @ 23 °C, IPA rinse

Ash – low power O_2 60s

Metallise – 30s 4:1 H_20:HCl de-ox, H_20 rinse, Plassys 20nm Ti, 130nm Au

Lift-off – 2hr warm acetone, pipette clean, IPA rinse

2. Mesa Isolation

Clean substrate – 30min warm acetone, IPA rinse

Spin resist – 12% 2010 5k rpm 60s. 1hr 180°C bake. 4% 2041 5k rpm 60s. 2hr 180°C bake

Exposure – Dose 225, Res 0.15, Spot 300

Develop – 1:1 MIBK:IPA 30s @ 23 °C, IPA rinse

Ash – low power O_2 60s

Isolate – 100:1:1 H_20:H_2O_2:Orthophosphoric, H_2O rinse (depth characterisation)

Resist removal – 2hr warm acetone, pipette clean, IPA rinse

3. Gates

Clean substrate – 30min warm acetone, IPA rinse

Spin resist – 4% 2041 5k rpm 60s, 1hr 180°C bake. 9% Coploy 5k rpm 60s, 1hr 180°C bake. 2.5% 2010, 5k rpm 60s, overnight 180°C bake

Exposure – Dose 118, Res 0.02, Spot 40 (CFA)

Develop – 2.5:1 MIBK:IPA 30s @ 23 °C, IPA rinse

Ash – low power O_2 30s

Recess – 30s 4:1 H_20:HCl de-ox, 55s Succinic acid (InP), H_2O rinse

Metallise – 10s 300:1:1 H_20:H_2O_2:Orthophosphoric, H_20 rinse, Plassys 15nm Ti 15nm Pt 200nm Au

Lift-off – 2hr warm acetone, pipette clean, IPA rinse

4. Ohmic Contacts

Clean substrate – 30min warm acetone, IPA rinse

Spin resist – 12% 2010 5k rpm 60s. 1hr 120°C bake. 4% 2041 5k rpm 60s. 1hr 120°C bake

Exposure – Dose 225, Res 0.15, Spot 160

Develop – 1:1 MIBK:IPA 30s @ 23 °C, IPA rinse

Ash – low power O_2 60s

Metallise – 30s 4:1 H_20:HCl de-ox, H_20 rinse, Plassys thin Mikeohm metallisation

Lift-off – 2hr warm acetone, pipette clean, IPA rinse

5. Bondpads

Clean substrate – 30min warm acetone, IPA rinse

Spin resist – 15% 2010 5k rpm 60s, 1hr 120°C bake. 4% 2041 5k rpm 60s. 1hr 120°C bake

Exposure - Dose 330, Res 0.15, Spot 300

Develop – 1:1 MIBK:IPA 30s @ 23 °C, IPA rinse

Ash – low power O_2 60s

Metallise – 30s 4:1 H_20:HCl de-ox, H_20 rinse, Plassys 50nm NiCr 1200nm Au

Lift-off – 2hr warm acetone, pipette clean, IPA rinse

AIV – Self-aligned gate 70nm InP process flow

1. Markers

Clean substrate – 30min warm acetone, IPA rinse

Spin resist – 12% 2010 5k rpm 60s. 1hr 180°C bake. 4% 2041 5k rpm 60s. 2hr 180°C bake

Exposure – Dose 310, Res 0.15, Spot 300

Develop – 1:1 MIBK:IPA 30s @ 23 °C, IPA rinse

Ash – low power O_2 60s

Metallise – 30s 4:1 H_20:HCl de-ox, H_20 rinse, Plassys 20nm Ti, 130nm Au

Lift-off – 2hr warm acetone, pipette clean, IPA rinse

2. Mesa Isolation

Clean substrate – 30min warm acetone, IPA rinse

Spin resist – 12% 2010 5k rpm 60s. 1hr 180°C bake. 4% 2041 5k rpm 60s. 2hr 180°C bake

Exposure – Dose 225, Res 0.15, Spot 300

Develop – 1:1 MIBK:IPA 30s @ 23 °C, IPA rinse

Ash – low power O_2 60s

Isolate – 100:1:1 H_20:H_2O_2:Orthophosphoric, H_2O rinse (depth characterisation)

Resist removal – 2hr warm acetone, pipette clean, IPA rinse

3. Gates

Clean substrate – 30min warm acetone, IPA rinse

Spin resist – 2.5% 2041 2k rpm 60s, 1hr 180°C bake. 1:4 LOR 5k rpm 60s, 15min 180°C bake. 80% UVIII 5k rpm 60s, hotplate 120°C 60s

Exposure – Dose 85 (head) / 2800 – 3000 (foot), Res 0.01, Spot 12 (CFA)

Develop – 90s 120°C hotplate. MFCD26 60s, H_2O rinse. O-xylene 60s 23°C, H_2O rinse.

Ash – low power O_2 30s

Recess – 30s 4:1 H_20:HCl de-ox. 15s succinic acid, 5s ortho, 4s succinic. H_2O rinse

Metallise – 3s 300:1:1 H_20:H_2O_2:Orthophosphoric, H_2O rinse, Plassys 15nm Ti 15nm Pt 200nm Au

Lift-off – 2hr warm acetone, pipette clean, IPA rinse

4. Ohmic Contacts

Clean substrate – 30min warm acetone, IPA rinse

Spin resist – 12% 2010 5k rpm 60s. 1hr 120°C bake. 4% 2041 5k rpm 60s. 1hr 120°C bake

Exposure – Dose 225, Res 0.15, Spot 160

Develop – 1:1 MIBK:IPA 30s @ 23 °C, IPA rinse

Ash – low power O_2 60s

Metallise – 30s 4:1 H_20:HCl de-ox, H_20 rinse, Plassys thin Mikeohm metallisation

Lift-off – 2hr warm acetone, pipette clean, IPA rinse

5. Bondpads

Clean substrate – 30min warm acetone, IPA rinse

Spin resist – 15% 2010 5k rpm 60s, 1hr 120°C bake. 4% 2041 5k rpm 60s. 1hr 120°C bake

Exposure - Dose 330, Res 0.15, Spot 300

Develop – 1:1 MIBK:IPA 30s @ 23 °C, IPA rinse

Ash – low power O_2 60s

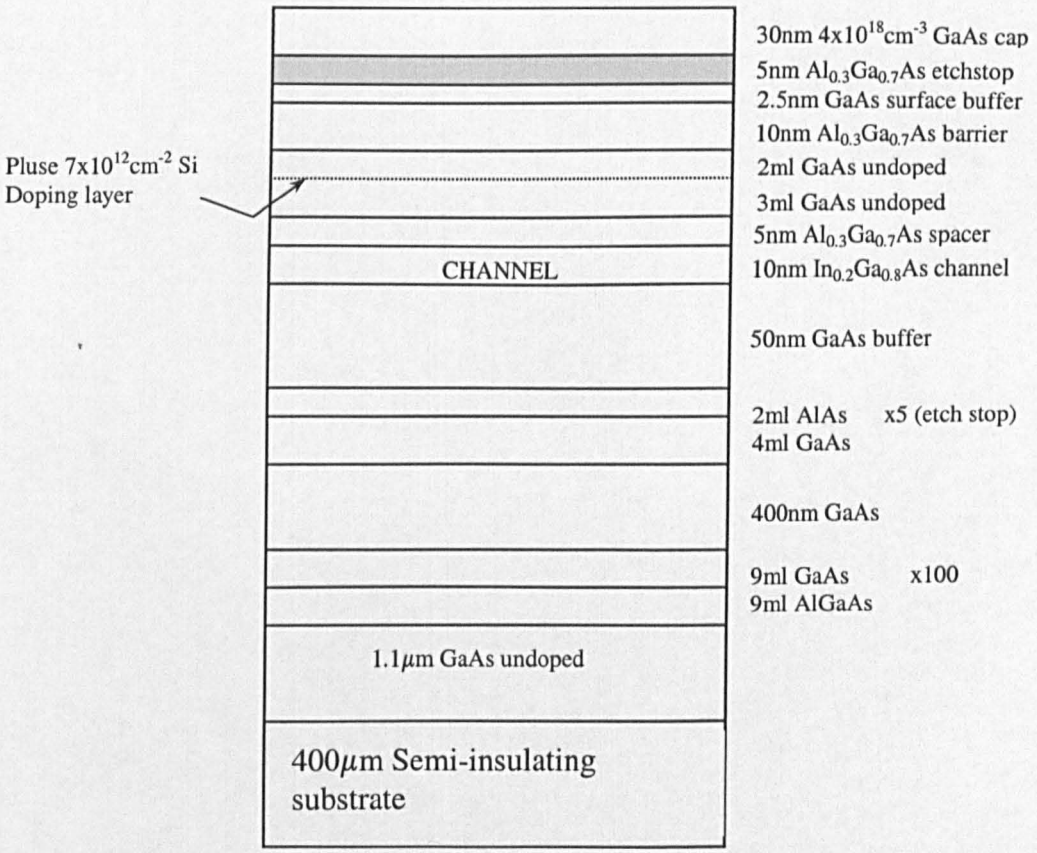
Metallise – 30s 4:1 H_20:HCl de-ox, H_2O rinse, Plassys 50nm NiCr 1200nm Au

Lift-off – 2hr warm acetone, pipette clean, IPA rinse

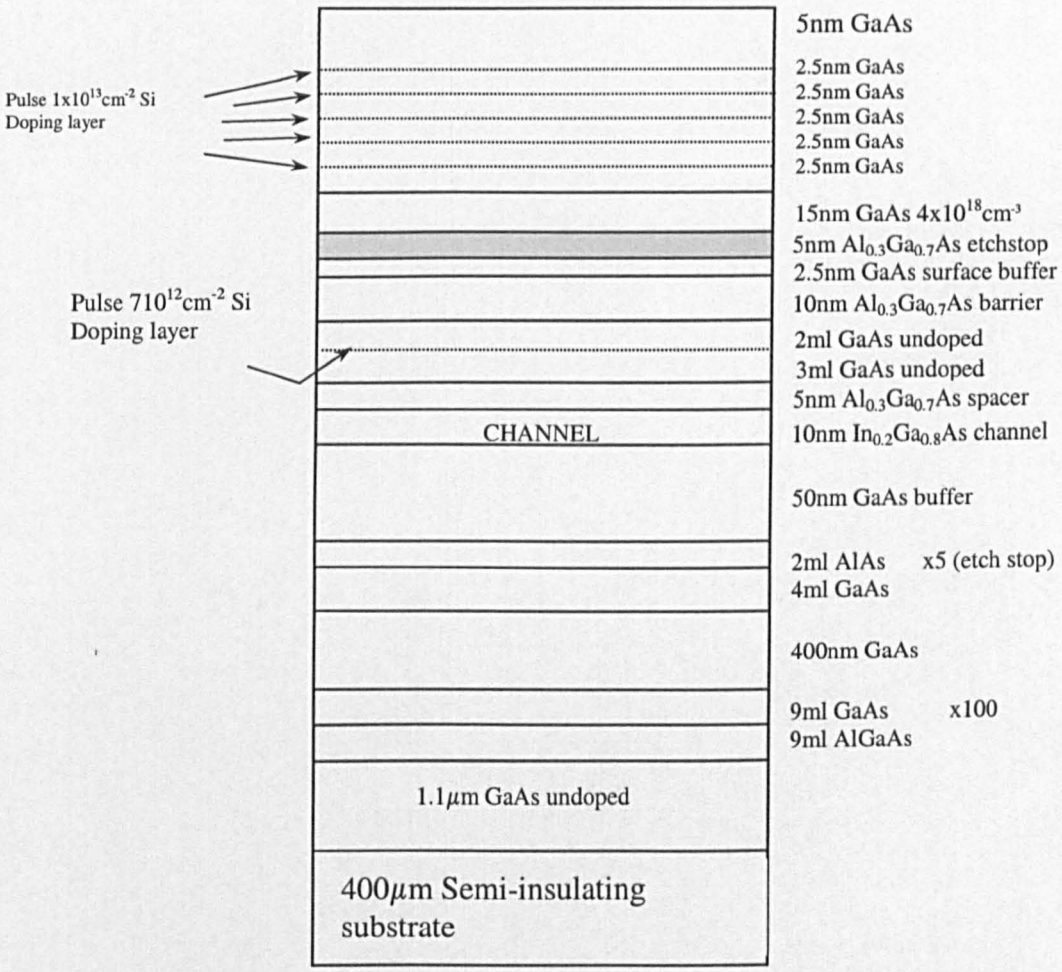
Appendix B

Material Layer Structures

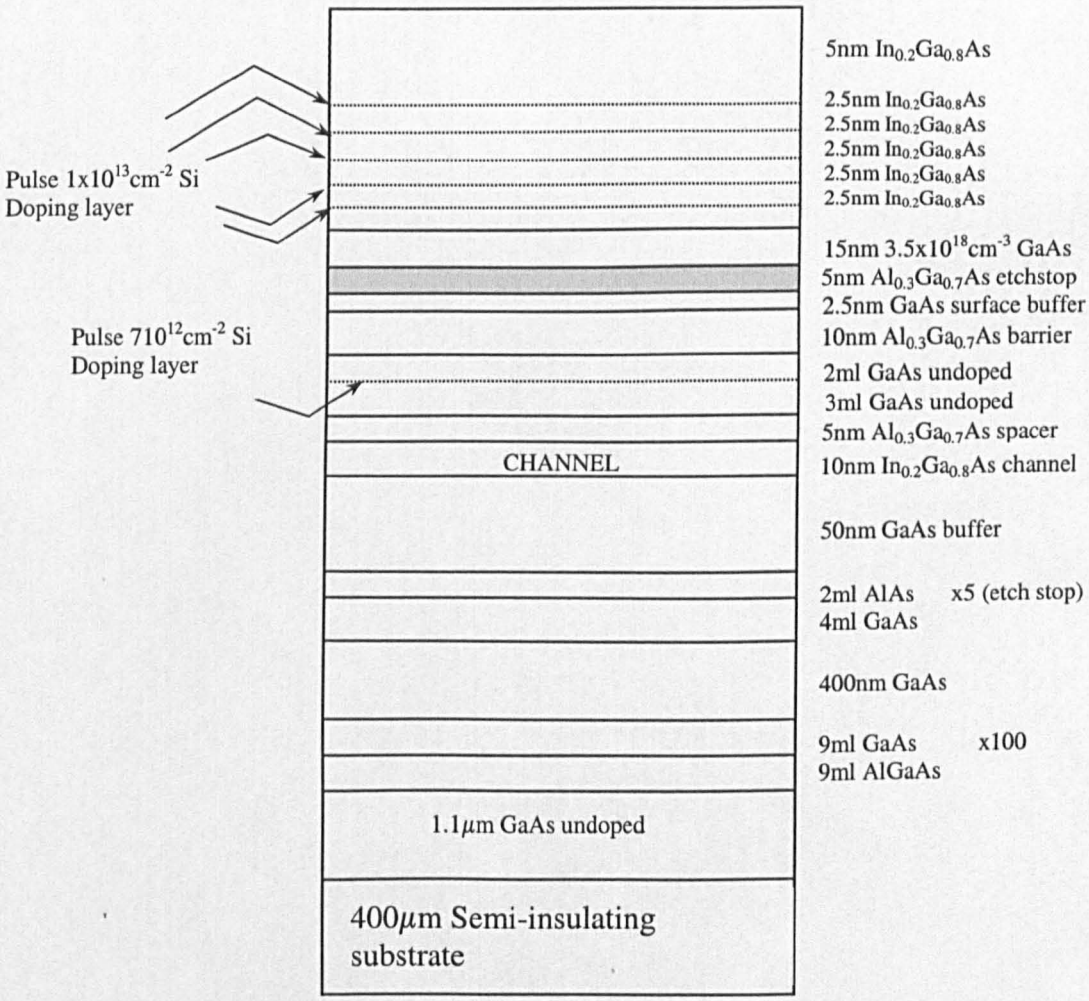
BI 1258 GaAs pHEMT



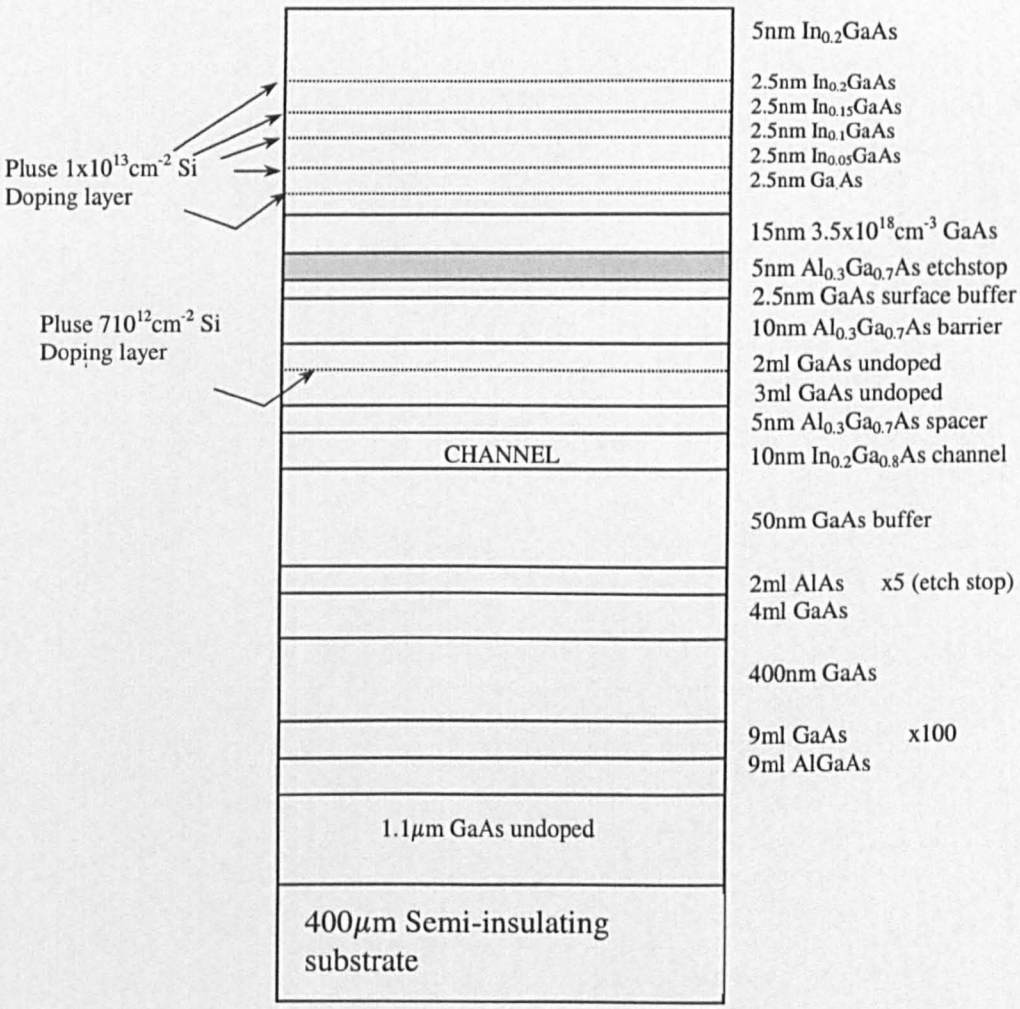
BII 1375 GaAs pHEMT



BIII 1408 GaAs pHEMT



BIV 1409 GaAs pHEMT



BV **Double delta doped InP (1744)**

Layer	Material	Doping	Thickness (nm)
1.	$In_{.53}Ga_{.47}As$ *cap*	1e19 cm^{-3} Si	20
2.	$In_{.52}Al_{.48}As$ *barrier*	-	5
3.	Delta Doping	4e12 cm^{-2} Si	-
4.	$In_{.52}Al_{.48}As$ *barrier*	-	12
5.	Delta Doping	6e12 cm^{-2} Si	-
6.	$In_{.52}Al_{.48}As$ *spacer*	-	3
7.	$In_{.53}Ga_{.47}As$ *channel*	-	15
8.	$In_{.52}Al_{.48}As$ *buffer*	-	400
9.	SI InP Substrate	-	-

GLASGOW UNIVERSITY LIBRARY