ROUTES TO COST EFFECTIVE RELISATION OF HIGH PERFORMANCE SUBMICRON GATE InGaAs/InAlAs/InP pHEMT

A thesis submitted to The University of Manchester for the degree of Doctor of Philosophy

In the faculty of Engineering and Physical Sciences

2013

Ka Wa Ian

School of Electrical and Electronic Engineering

Institute: School of EEE, the University of Manchester
Candidate: Ka Wa Ian
Degree: Doctor of Philosophy (PhD)
Title: ROUTES TO COST EFFECTIVE RELISATION OF HIGH PERFORMANCE SUBMICRON GATE InGaAs/InAlAs/InP pHEMT
Date: 19/09/2013

ABSTRACT

The Square Kilometre Array (SKA) is known to be the most powerful radio telescope of its type. In support of its high observational power, it is estimated that thousands of antenna unit equipped with millions of LNA (low noise amplifier) will be deployed over a large area (radius>3000km). The stringent requirements for high performance and low cost LNA design bring about many challenges in terms of material growth, device fabrication and low noise circuit designs.

For the past decade, the Manchester group has been wholeheartedly committed to the research and development of high performance, low cost Monolithic Microwave Integrated Circuit (MMIC) LNA with high breakdown (15V) and low noise characteristics (1.2dB to 1.5dB) for the SKA mid-frequency application (0.4GHz to 1.4GHz). The on-going optimisation of current design is hindered by the restriction of standard i-line 1μm gate lithography. The primary focus of this work is on the design and fabrication of new, submicron gate InGaAs/InAlAs/InP pHEMTs for high frequency applications and future SKA high frequency bands.

The study starts with the design and fabrication of InGaAs-InAlAs pHEMT sub-100nm gate structure using E-Beam lithography. To address the problems of short channel effect and parasitic components, devices with 128nm T-gate structure, and with optimised device geometries and enhanced material growth, having f_T of 162GHz and f_{max} of 183GHz are demonstrated, outlining the importance of device scaling for high speed operation. In addition, a gate-sinking technique using Pd/Ti/Au metallisation scheme was investigated to meet the requirement for single voltage supply in circuit design. Device with Pd-buried gate exhibits enhanced DC and RF characteristics and showed no degradation over 5 hours' annealing at 230˚C. The implementation of this highly thermal stable Pd Schottky gate is key to improving the device's long-term reliability at high-temperature operation.

To solve the problem of low productivity in E-Beam lithography, a simple, low cost, technique termed soft reflow was introduced by utilising the principle of solvent vaporisation in a closed chamber. It provides a hybrid solution for the fabrication of submicron device using low cost i-line lithography. The integration of this new soft reflow process with the Pd-gate sinking technique has enabled the large-scale fabrication of 250nm T-gate pHEMTs, with excellent f_T of 108GHz and a f_{max} of 119GHz and with device yields exceeding 80%. This novel soft reflow technique provides a high yield, fast throughput, solution for the fabrication of submicron gate pHEMT and other ultra-high frequency nanoscale devices.

DECLARATION

No portion of the work referred to in the dissertation has been submitted in support of an application for another degree or qualification of this or any other university or other institute of learning.

COPYRIGHT STATEMENTS

ACKNOWLEDGMENTS

I would like to express my thanks to a lot people in the M&N group, especially my PhD project supervisor, Prof. Mohamed Missous, who always provides support and guidance throughout my three and half years PhD studies.

A special thanks to our experimental officer Dr. James Sexton for his effort on the maintenance of our cleanroom facilities and professional advice on experimental works. I would also like to acknowledge my co-workers Dr. Sanae Boulay and Dr. Muammar Mohamad Isa for their kind instructions at the beginning of my PhD study and Dr. Michael Exarchos for his extensive knowledge of semiconductor Physics and support in my experimental work.

I would take this opportunity to show my appreciation to my co-worker Dr. Yan Lai, for her help on the pulse DC measurement and to Dr. Scott Lewis, from the School of Physics, for his great help on the E-beam lithography and imaging over the years.

I am greatly indebted to my sponsors EPSRC for providing funding under the Doctoral Training Account scheme.

Finally, this work would have been mission impossible without my parents' cares and supports every day during my studies, and the encouragement and understanding of my two younger brothers in Tokyo, Japan.

DEDICATION

This thesis is dedicated

to my parents and my younger brothers.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

2DEG	2 Dimensional Electron Gas
AlGaAs	Aluminium Gallium Arsenide
ADC	Analogue-to-Digital Converter
AA	Aperture Array
BJT	Bipolar Junction Transistor
CMOS	Complementary Metal Oxide Semiconductor
DUV	Deep Ultraviolet
DI water	De-ionised water
D-mode	Depletion mode
DIBL	Drain-Induced Barrier Lowering
EEE	Electrical and Electronic Engineering
E-beam	Electron Beam
E-mode	Enhancement mode
FET	Field Effect Transistor
FOVs	Field of Views
FP	Field plate
FOM	Figures of Merits
GaAs	Gallium Arsenide
Ge	Germanium
G-S-G	Ground-Signal-Ground
HBT	Heterojunction Bipolar Transistor
HMET	High Electron Mobility Transistor
InAlAs	Indium Aluminium Arsenide
InGaAs	Indium Gallium Arsenide
InP	Indium Phosphide
IR	Infra-Red

IEMN	Institut d'Electronique, de Microélectronique et de Nanotechnologie
IC	Integrated Circuit
ITRS	International Technology Roadmap for Semiconductors
IPA	Isopropyl Alcohol
KSPs	Key Science Programs
LRRM	Line-Reflect-Reflect-Match
LNA	Low Noise Amplifier
MAG	Maximum Available Gain
MOCVD	Metal Organic Chemical Vapour Deposition
MOSFET	Metal Oxide Semiconductor Field Effect Transistor
MESFET	Metal Semiconductor Field Effect Transistor
M&N	Microelectronic & Nanostructures
Nf_{min}	Minimum Noise Figure
MODFET	Modulation Doped FET
MBE	Molecular Beam Epitaxy
MMIC	Monolithic Microwave Integrated Circuit
NMP	N-Methyl-2-pyrrolidone
NF	Noise Figure
nMOS	n-type channel MOS
PECVD	Plasma-Enhanced Chemical Vapor Deposition
PA	Power Amplifier
pHEMT	pseudomorphic High Electron Mobility Transistor
QW	Quantum Well
RA	Radio Astronomy
RF	Radio Frequency
RTD	Resonant Tunnelling Diodes
RT	Room Temperature

SDHT	Selectively Doped Heterojunction Transistor
SCE	Short Channel Effect
SNR	Signal to Noise Ratio
Si	Silicon
SiO_2	Silicon Dioxide
SiGe	Silicon Germanium
Si_3N_4	Silicon Nitride
SKA	Square Kilometre Array
SA	Succinic Acid
SoC	System on Chip
CF_4	Tetrafluoromethane
TE	Thermionic Emission
TFE	Thermionic Field Emission
TLM	Transfer Length Method
TEGFET	Two-Dimensional Electron Gas FET
UV	Ultraviolet
U	Unilateral power gain
UoM	University of Manchester
VNA	Vector Network Analyser
VLSI	Very Large Scale Integration

LIST OF PUBLICATIONS

1. **Ian, Ka Wa,** Michael Exarchos, and Mohamed Missous. "A novel low temperature soft reflow process for the fabrication of deep-submicron (< 0.35 μm) T-gate pseudomorphic high electron mobility transistor structures." Nanotechnology 24.5 (2013): 055202.

2. Packeer, F., Isa, M. M., Jubadi, W. M., **Ian, K. W.**, & Missous, M. "(2013). Fabrication and characterization of tensile In0.3Al0.7As barrier and compressive In0.7Ga0.3As channel pHEMTs having extremely low gate leakage for low-noise applications." Journal of Physics D: Applied Physics, 46(26), 264002.

3. **Ian, Ka Wa,** Michael Exarchos, and Mohamed Missous. "Fabrication of 250-nm T-gate $In_{52}Al_{48}As$-$In_{70}Ga_{30}As$ pHEMT using a novel solvent reflow technique at room temperature." Advanced Semiconductor Devices & Microsystems (ASDAM), 2012 Ninth International Conference on. IEEE, 2012.

4. **K. W. Ian**, M. Exarchos, M. Missous, "Novel submicron T-gate InGaAs-InAlAs pHEMT using a solvent reflow technique at room temperature", European Materials Research Society (EMRS) 2012 FALL MEETING, Warsaw University of Technology, Poland, 17-21 Sep. 2012.

5. **K. W. Ian**, M. Exarchos, M. Missous, "A simple and economical approach for the fabrication of submicron structures by means of bi-layer reflow technique using solvent vapour", U.K. Semiconductors 2012 Conference, University of Sheffield, 4-5 July 2012.

6. **K. W. Ian**, Y. Lai, M. Missous, "Thermal stability study of InGaAs-InAlAs pHEMT using Palladium gate metallisation", Postgraduate Poster Conference and Industrial Advisory Group Meeting, The University of Manchester, 30 Nov 2011.

7. **K. W. Ian**, Y. Lai, M. Missous, "Thermal stability study of InGaAs-InAlAs pHEMT using Palladium gate metallisation", Heterostructure Technology

(HeTech) 2011 European Workshop, Institute of Electronic, Microelectronic and Nanotechnology, Lille, France, 7-9 Nov. 2011.

8. **K. W. Ian**, Y. Lai, M. Missous, “Thermal stability study of InGaAs-InAlAs pHEMT using Palladium gate metallisation”, U.K. Semiconductors 2012 Conference and Exhibition, University of Sheffield, 6-7 July 2011.

9. Saguatti, D., Chini, A., Verzellesi, G., Isa, M. M., **Ian, K. W.**, and Missous, M. , “Improvement of Breakdown and DC-to-Pulse Dispersion Properties in Field-Plated InGaAs-InAlAs pHEMTs”, International Conference on Indium Phosphide & Related Materials (IPRM), 22 – 26 May, 2011 Berlin, Germany, pp. 84-86.

10. M. Mohamad Isa, D. Saguatti, G. Verzellesi, A. Chini, **K.W. Ian**, M. Missous, "Fabrication of novel high frequency and high breakdown InAlAs-InGaAs pHEMTs,", 8th International Conference on Advanced Semiconductor Devices & Microsystems (ASDAM), 2010, 25 – 27 Oct. 2010, pp. 41-44.

11. M. Mohamad Isa, D. Saguatti, G. Verzellesi, A. Chini, **K.W. Ian**, M. Missous, "Novel high frequency and high breakdown InAlAs-InGaAs pHEMTs for low noise and high power applications", UK Semiconductor 2010 Conference, The University of Sheffield, 7-8 July 2010.

12. D. Saguatti, M. Mohamad Isa, **K. W. Ian**, and M. Missous, "InAlAs-InGaAs pHEMT TCAD Simulation and Optimization”, UK Semiconductor 2010 Conference, The University of Sheffield, 7-8 July 2010.

13. Saguatti, D., Chini, A., Verzellesi, G., Isa, M. M., **Ian, K. W.**, and Missous, M., "TCAD optimization of field-plated InAlAs-InGaAs HEMTs", International Conference on Indium Phosphide & Related Materials (IPRM), 31 May – 4 June 2010. Takamatsu Symbol Tower, Kagawa, Japan, pp. 1-3

1 INTRODUCTION

1.1 Introduction: The objective of SKA

Since the "father of modern observational astronomy", Galileo Galilei, unveiled the tip of the Milky Way, the twinkling sky has been observed for hundreds of years [1]. Human being's curiosity has propelled rapid advances in astronomy for the past century. Man's footprints have been left on the moon and unmanned spacecraft and large-scale telescopes like the Hubble Space Telescope have extended human's reach to outer space [2].

Human being never ceases to explore the universe and resolve many enigmas such as the Big Bang, black holes and dark energy. However, uncovering these mysteries entails the assistance of cutting-edge technology of which the telescope is key. The observational power of common telescopes is determined by two important terms, angular resolution (θ) and sensitivity (S). θ represents the ability to distinguish between two neighbouring objects at a remote distance while S means the capability to detect an incoming weak signal from those objects, and these are characterized by Equation 1.1 and Equation 1.2:

$$\theta = \frac{1.22\lambda}{\mathrm{D}}$$ **Equation 1.1 [3,4]**

$$S = \frac{\mathrm{A_{eff}}}{\mathrm{T_{sys}}}$$ **Equation 1.2 [3,4]**

Where λ= wavelength of the incoming radiation, D = diameter of the lens aperture, A_{eff} = effective collecting area which is proportional to the magnitude of D and T_{sys} = receiver equivalent system temperature.

Figure 1.1 illustrates the relationship between spectrum wavelength and the use of different astronomical telescopes across those spectrums. There are two types of observations, namely ground and space-based observations, and their specific usages

depend on the spectrum absorption of the earth's atmosphere.

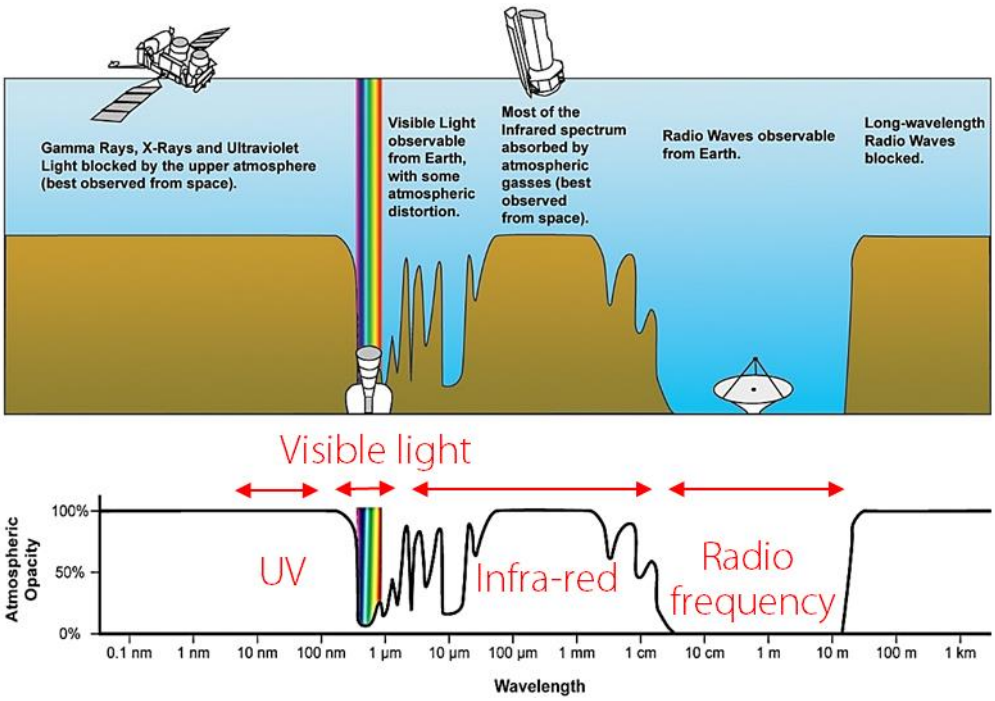


Figure 1.1 Relationship between spectrum wavelength and the different type of telescopes [5]

High frequency signals such as X-ray, Ultra-violet (UV) and also the Infra-Red (IR), are mostly blocked by the atmosphere, which suggests that space-based telescope is richer in features. Ground-based optical telescope, an economical option, is a very popular option in traditional astronomical observation. However, its observation power is physically confined by the incoming short wavelength, which can only be compensated by using a larger but very expensive lens aperture. The key to addressing this issue is to employ a telescope technology operating at radio frequency.

Radio astronomy is one of the most important fields in modern astronomy which involves searching and analysing the natural radio signals from the universe. Radio signals often travel hundreds or thousands of light years from celestial bodies. The detection of such weak signals relies on the sensitivity of the telescope, and this is largely determined by A_{eff}, a key parameter to assess radio telescope's performance. The Square Kilometre Array (SKA) radio telescope, whose effective collecting area, A_{eff}, is equivalent to one square kilometre, can tackle the aforementioned issues.

1.2 Overview of SKA

The SKA project focuses on the development of new generation of array-based radio telescope to address five most important topics in Key Science Programs (KSPs) in SKA parlance .Firstly, the strong field tests of gravity using pulsars, black holes; secondly, the origin and evolution of cosmic magnetism; thirdly, evolution of galaxies, cosmology, dark matter and dark energy; fourthly, probing the dark ages; fifthly, the cradle of life [6]. The goal of SKA is to build the most powerful radio telescope with the most appropriate performance in terms of searching speed, sensitivity and resolution in its class and the major specifications are as follows [6]-[7]:

1. An effective collection area of one square kilometre, which enables at least 50-fold better sensitivity and 100-fold higher resolution compared with legacy radio telescopes.

2. A wide range of operation frequency from 70MHz to 10GHz and will be extended to over 25GHz when the system is in full operation by the year 2025 and beyond.

3. Combinations of various antennas enable the SKA to operate in a wide range of field views (FOVs). For frequencies below 1GHz, it is able to provide 200 square degrees FOV which is approximately the area of five moons on the sky. This means for any higher frequencies, it is able to “view” more than 1 square degree. Such high coverage will enable it to search the sky at much higher speed than any other radio telescopes currently in use.

The operating frequency of the SKA radio telescope can be divided into 3 categories, SKA-low (70 to 300MHz), SKA-mid (300MHz to 10GHz) and SKA-high (10 to 25+ GHz). Each of these target different fields of astronomy and their implementation will be scheduled as a three-stage process as shown in Figure 1.2:

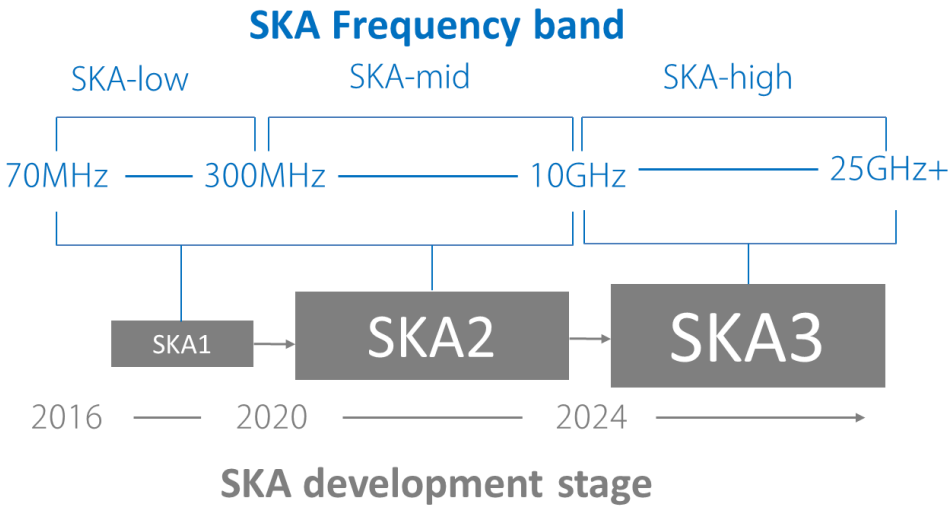


Figure 1.2 Relationship between the SKA schedule and the operating frequency [7]

The phase-one construction of SKA (SKA1), is to demonstrate its observation power in comparison with other current telescopes and the operating frequency will be limited to the SKA-low and SKA-mid frequency range for cost and technical consideration. As a subset of SKA2, the SKA1 development only provides 10% of capability and it will be extended to the full power at SKA2. Finally, the frequency band will be pushed beyond 25GHz at SKA3, which will hinge on the progress of SKA2 and the future development of relevant antenna and receiver technologies [7].

The wide frequency band operation of SKA undoubtedly poses a great challenge to the sensors' engineering. There are many reference design proposed at the pre-construction stage of SKA1 but the final decision will be made based on consideration of cost, performance and maturity of particular sensor technology at the construction stage. The European Square Kilometre Array Design Studies (SKADS) has proposed a feasible solution for the SKA system and sensors' implementation. In this proposal, SKA-low and SKA-mid operating frequencies are further divided into three sub-sets, where three antenna designs are proposed to meet the specific performance requirement [8]. Figure 1.3 presents three unique sensor technologies and their correlation to three different operating frequencies.

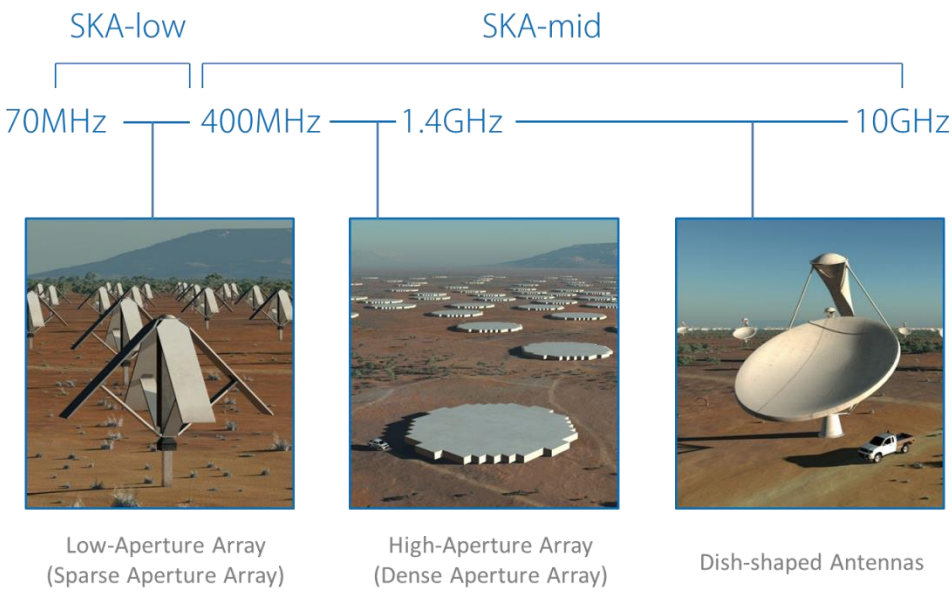


Figure 1.3 A schematic of the proposed antenna designs and their operating frequencies [8]-[9]

The Aperture Array (AA) is a key feature of the SKA, which consists of many small receptor units. The deployment of AA facilitates the construction of very large observation area at low cost. However, its usage would be restricted to frequencies below 1.4GHz due to its poor efficiency at higher frequencies, where traditional dish antenna technology becomes a better option [8].

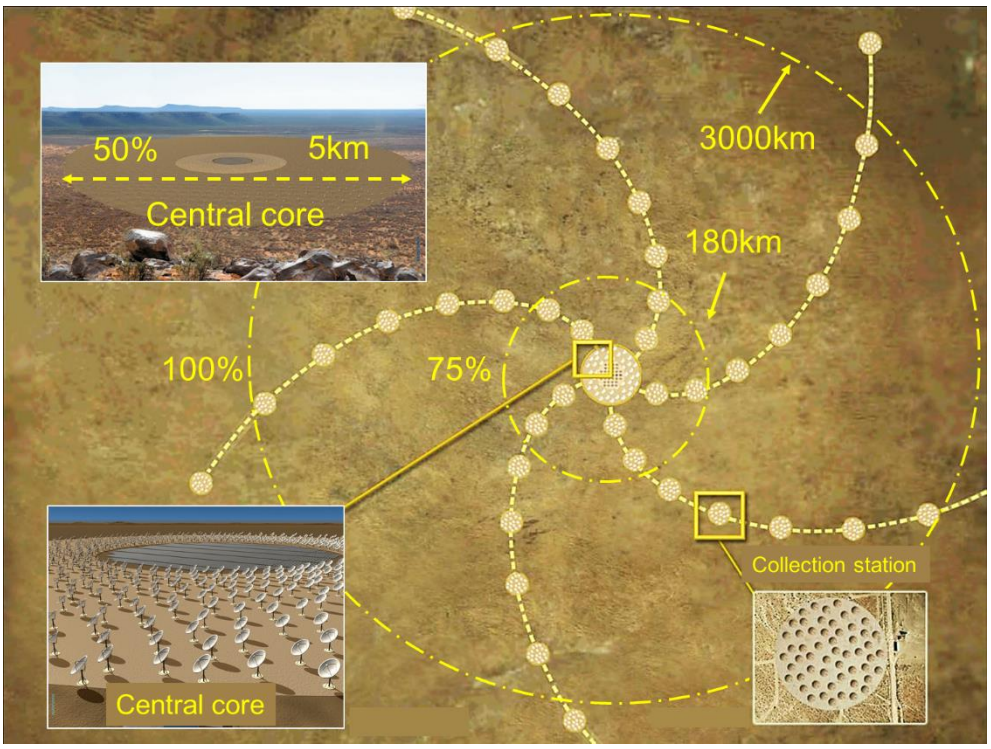


Figure 1.4 SKA layout and the antenna configuration [8]-[10]

The SKA thus, consists of a huge amount of various antennae covering a very large area. The final system configuration is currently under discussion and it is very likely to be changed before the completion of SKA1. For illustration, a system layout is shown in Figure 1.4 according to the information given in [8]-[10]. The system is made up of three different regions, the central core region, mid region and five spiral arms, covering a 3000km diameter area.

The 5km central core region covers approximately 50% of the overall collecting area by employing a huge number of dish antennas and low and high frequency aperture arrays. From the central region extending to 3000km diameter, a number of collecting stations will be built along five spiral arms. For the first 180km, the collection station mainly consists of aperture arrays and a small number of dishes antennae to generate 75% coverage while the remaining 25% will be covered by stations only equipped with dishes antennas for higher frequency detection.

The key feature of SKA is to detect the weak incoming radio signals by employing a very large collection area, which consists of millions of small receptor units [8]. As the first stage signal processing component, the Monolithic Microwave Integrated Circuit Low Noise Amplifier (MMIC LNA) plays an important role in trimming cost, and guaranteeing performance and reliability of the front end system. In other words, a MMIC LNA is a highly efficient and economical signal amplification circuitry which consists of passive (capacitor, inductor and resistor) and active (transistor) components integrated on one single chip. Pseudomorphic High Electron Mobility Transistor (pHEMT) is often utilised as the active part of LNA design thanks to its optimal mobility and excellent noise performance. The development of the pHEMT and hence the passive components will be pivotal to the successful LNA implementation. In the next section, a comprehensive review will cover all the previous works conducted by the Manchester group on the pHEMT and hence the MMIC LNA development for the SKA.

1.3 Previous LNA studies for the SKA

Currently, the SKA is one of the most powerful and complex radio telescope in terms of technical designs, development and construction stages. Since the birth of the SKA concept in 1991, there have been over 50 institutes in 20 different countries across the world working on different aspects of SKA development [9]. The University of Manchester, as the head office of the SKA project, has been involved in many pathfinder studies for the key building blocks such as antenna, Analogue-to-Digital Converter (ADC) and MMIC LNA, etc [11]. The School of Electrical and Electronic Engineering (EEE) at Manchester have been pioneering the development of ultra-high speed ADC and MMIC LNA over the past 10 years. The relationship between the SKA timeline and our researches is presented in Figure 1.5.

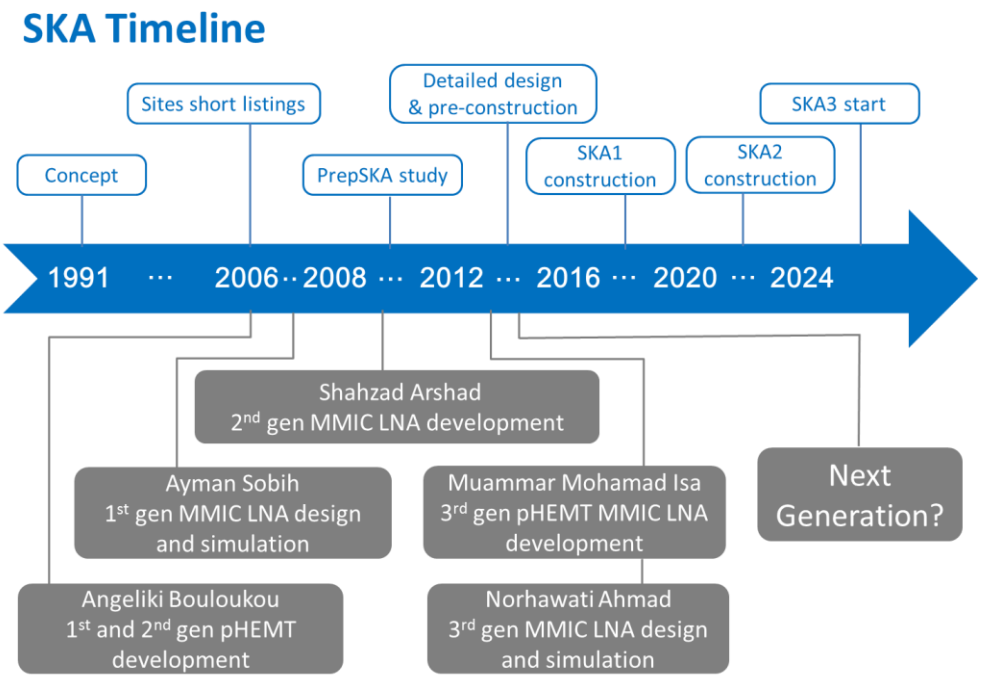


Figure 1.5 The relationship between SKA timeline and the Manchester LNA MMIC work

The evolutionary cycle of the MMIC LNA are mainly divided into two parts: the development of the single pHEMT active component and the whole MMIC LNA. Dr. Angeliki Bouloukou, the developer of the first (1st gen) and second (2nd gen) generation pHEMT, spent a lot of effort on the optimisation of 1μm gate Indium

Phosphide (InP) based pHEMT for SKA application [12]. The advantage of InP-based pHEMT became obvious in comparison with the conventional Gallium Arsenide-based (GaAs) pHEMT with regard to noise performance. The introduction of highly selective gate recessing etching and non-doped cap layer on the strained Indium Gallium Arsenide-Indium Aluminium Arsenide (InGaAs-InAlAs) pHEMT improved the gate leakage and also the breakdown voltage (-18V). Concomitantly, Dr. Ayman Sobih concentrated on the design and simulation of the 1st gen MMIC LNA by employing extensive linear and non-linear modelling on the fabricated pHEMT and the relevant passive components for the frequency range 0.3 to 2GHz. However, the noise performance was not quite satisfactory at that time due to the absence of multiple gate fingers and the insufficient optimisation of the passive modelling [13].

Subsequently, the 2nd gen pHEMT addressed the aforementioned issues by employing thicker gate metallisation and multi-gate fingers (4 to 6) design. This was then integrated with Dr. Shahzad Arshad's physical and empirical device modelling for the design and simulation of the 2nd gen MMIC LNA. Three different designs, including single-ended, fully differential, and differential to single ended MMIC LNAs, were demonstrated by using the Manchester in-house fabricated 2nd gen pHEMT and this outperformed the MIC design with commercial GaAs product in terms of power consumption [14].

Dr. Muammar Mohamad Isa and Dr. Norhawati Ahmad were working together for the development of 3rd gen product [15]-[16]. Considerable efforts were invested in optimising the InP-based pHEMT. A better solution was developed for the gate recess etching step to create a high quality Schottky gate contact which incorporated a wide band gap barrier layer to suppress the gate leakage to the lowest level possible. In addition, the introduction of a new Field-Plate (FP) feature improved the breakdown voltage even further (-21V). The modelling data of this excellent device was used for the design and simulation of the 3rd gen MMIC LNA. The full MMIC LNA was fabricated and characterised for the first time. However, the measured noise

performance was about 1.5dB which is 0.1 to 0.2dB higher than the modelling data partly because of the poor passive modelling.

1.4 Objective

In past decades, researchers in the Manchester group have been wholeheartedly dedicated to improving performance of InP-based MMIC LNA. Many aspects, including band gap engineering, the optimisation of fabrication process, the implementation of new physical feature, the improvement of modelling and simulation, have been intensively investigated in previous studies. The performance of the 1μm pHEMT had reached a plateau that no significant improvement could be made without employing a smaller gate length; while at the same time keep the cost effectiveness of i-line lithography employed to date.

The development of next-generation submicron gate InGaAs-InAlAs pHEMT is a key objective in this study for the SKA3 implementation or other higher frequencies such as X-band applications in the near future. The major achievement of this work is the novel soft-reflow process using conventional 1μm i-Line photography and solvent vapour technique for the fabrication of high performance quarter-micron gate (0.25μm) device. This hybrid solution solves conventional problems such as speed of production, equipment and photo mask cost in traditional Electron Beam (E-beam) and Deep Ultra-Violet (DUV) lithography.

1.5 Thesis outline

Chapter 2 begins with a literature review of the basic concepts of III-V compound semiconductors. A historical background which details the development of different III-V FET devices in relation to the advancement of material engineering and the more demanding RF applications is presented. From a comparison between the different FET structures (MESFET, HEMT and pHEMT) and material systems (InGaAs-AlGaAs and InGaAs-InAlAs), the advantage of InP-based pHEMT for low noise/high speed applications will be outlined.

Chapter 3 describes the characteristics and performance of E-beam fabricated submicron gate InGaAs-InAlAs pHEMT in collaboration with other institutes. The importance of gate shrinkage for low noise and high speed operations and the problem of fine lithography are presented. The associated issues with submicron gate structure are explored from our first attempt at fabricating submicron (120nm planar gate) pHEMT. The results and feedbacks obtained were instrumental for the optimisation of device geometry and material growth in the fabrication of high performance (128nm T-gate) pHEMT.

Chapter 4 presents the heat treatment study on InGaAs-InAlAs pHEMT with Pd/Ti/Au gate metallisation scheme. The behaviours of Pd-buried gate are studied with respect to the metal thickness; annealing temperature and time are reported for the first time on this material system. The device thermal stability for high temperature operation is also investigated by performing a long term annealing study.

Chapter 5 provides a detailed description on the development of a novel soft reflow technique, which provides an economical and efficient solution for the fabrication of submicron T-gate pHEMT. The principle and characteristics of soft reflow technique will be presented. The integration of this soft reflow process, together with the feedbacks from **Chapter 3** and the Pd-gate sinking technique from **Chapter 4**, are employed in the fabrication of high performance 250nm T-gate pHEMT.

Chapter 6 draws conclusion from this work, emphasising the significance of the project for the next generation LNA design and also its potential for other high speed applications. In addition, further works are proposed for future project development based on this thesis.

2 LITERATURE REVIEW

2.1 Introduction: A historical review of III-V compound semiconductor devices

For the past 60 years, semiconductor electronics have been one of the most extraordinary industries of the 20th century. Semiconductor chips are becoming more powerful, smaller and more economical and energy efficient. From bulky personal computers to portable tablet computer, from space satellite communication to convenient mobile broadband, semiconductor technology is indispensable in our daily life. Such achievement is mainly attributed to the continuous reduction of device size and the advancement of material engineering. In this chapter, a panoramic view of semiconductor evolution, in particular radio frequency (RF) applications, with respect to the development of material and device structure is presented. The significance of III-V compound semiconductors will be highlighted for its advantages and application in this field.

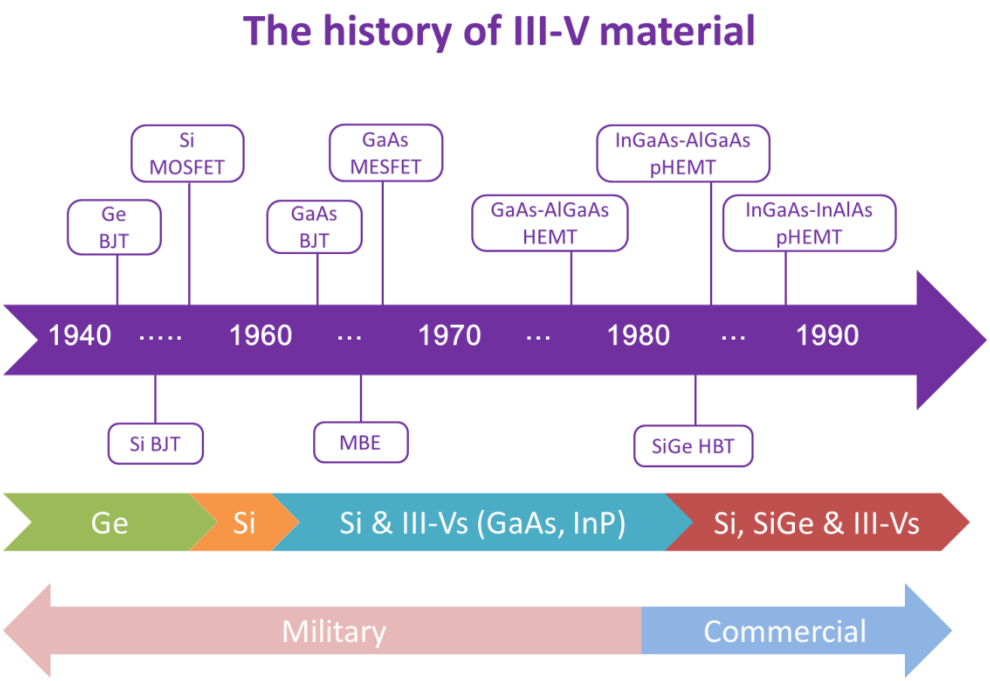


Figure 2.1 Major achievement of semiconductor devices for radio frequency applications

Figure 2.1 illustrates the major achievement of semiconductor devices in terms of structure, material and application. In 1947, the first semiconductor transistor device called the point contact transistor was invented by William Shockley at Bell Laboratories and it was based on a Germanium (Ge) Bipolar Junction Transistor (BJT) structure [17]. Ge BJTs developed rapidly and became dominant in the semiconductor market until the early 1960s. Silicon (Si), the basis of modern semiconductor electronics, was chosen as a replacement for Ge thanks to its better thermal stability and, more importantly, its low cost of manufacturing [18]. Another important property of Si lies in the formation of uniform native oxide, the Silicon Dioxide (SiO_2), which heralded the invention of the Metal Oxide Semiconductor Field Effect Transistor (MOSFET). This is the fundamental component of digital electronics such as microprocessor and flash memory. From 1960s, Si replaced Ge to lead the development of the semiconductor industry in the form of the Si BJT and Si MOSFET.

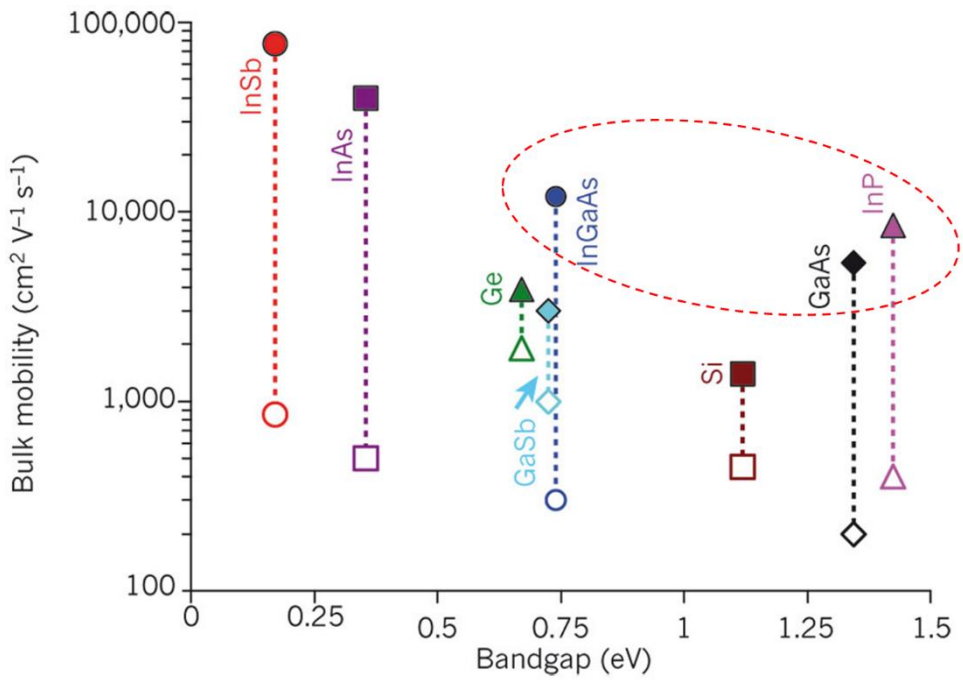


Figure 2.2 A comparison of electron (solid symbols) and hole (hollow symbols) mobility between Silicon and other common III-V compound semiconductors [19]

In the semiconductor world, mobility is a term which determines the ease with which carriers move in a crystal and is in many ways, but indirectly, the primary determinator of the device operation speeds, as it sets access resistances values with the saturated velocity under high electric fields determining how fast carriers move in the crystal. As shown in Figure 2.2, group III-V compound semiconductors have drawn considerable attention due to their higher electron mobility in comparison with Si despite comparable saturated velocities. The high electron mobility nature of III-V materials attracted many interests for military applications such as radar and security communication from the United States government, which lead to rapid expansion of III-Vs for high frequency RF electronics. In the early 1960s, the first Metal Semiconductor Field Effect Transistor was developed by Shockley who utilised a Schottky barrier metal gate on a Gallium Arsenide (GaAs) substrate [17].

However, the potential of III-V materials was not fully unleashed until the invention of the Molecular Beam Epitaxy (MBE) technique, which allows for the formation of heterojunction by growing high-quality interfaces with precise control of material thickness. This ushered in discovering the High Electron Mobility Transistor (HEMT) by Takashi Mimura in 1979 [20], which features enhanced carrier transport property by utilising the concept of a 2-Dimensional Electron Gas (2DEG) channel, with advanced material energy band gap engineering.

From 1980s, applications, such as wireless communications (0.9 to 2.5GHz), satellite communications (20 to 30GHz), traffic radar (77GHz) and different sensor applications (>90GHz) [18], have been widely applied in civilian area in addition to military ones. To accommodate premium market demand, the first pseudomorphic High Electron Mobility Transistor (pHEMT) was introduced by Ketterson who used the strained InGaAs-AlGaAs heterojunction interface in 1985 [20]. The concept of pHEMT is based on the formation of a very thin latticed mismatched but coherent channel layer, which provides a better carrier confinement and enhanced transport behaviour. In the late 1980s, Indium Phosphide (InP), featuring higher electron

mobility and greater degree of control over the band gap engineering, became a successor to GaAs in high frequency and low noise applications such as mm-wave Low Noise Amplifiers, while GaAs-based HEMTs became more widely used in high power application such as power amplifiers (PA), with GaN dominating this later filed in the last 8 years or so.

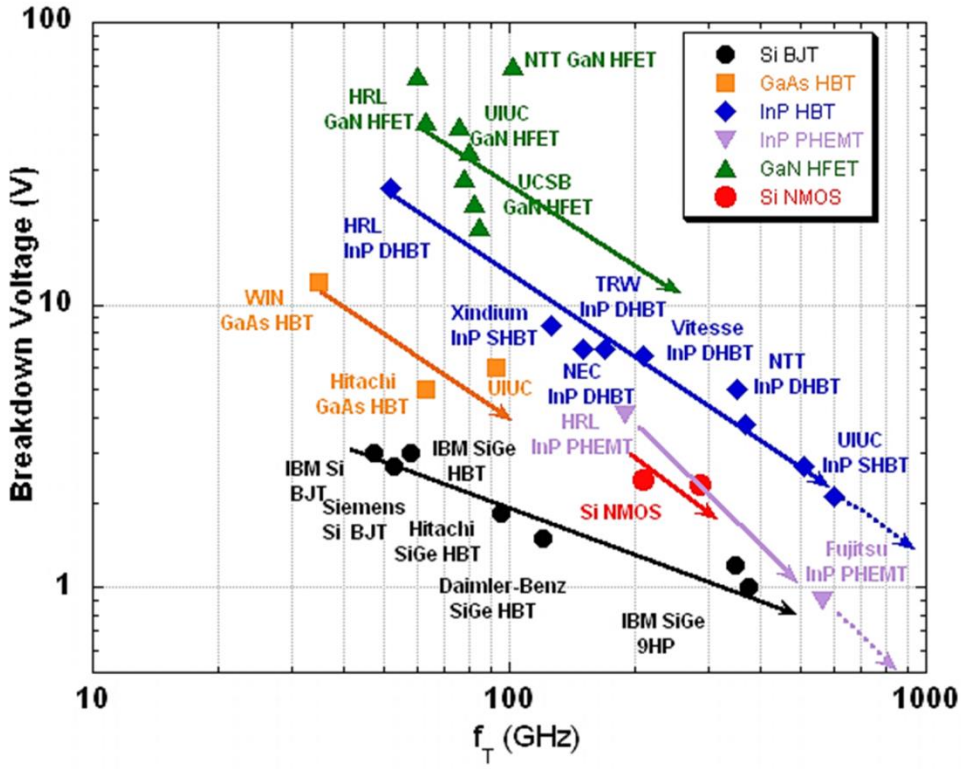


Figure 2.3 The reported cut-off frequency vs breakdown voltage of III-V, Si and SiGe transistors over the years [21]

As shown in Figure 2.3, Si-based devices started to gain grounds in the field of RF electronics recently, thanks to benefits of traditional Very Large Scale Integration (VLSI) electronics. The RF performance was improved significantly with the continuous scaling of the transistor gate dimension, which effectively compensated drawbacks of Si material properties. It also capitalised on heterojunction concepts by introducing Silicon Germanium (SiGe) compound material [22]. This facilitated the integration of digital and analogue electronics in one single chip, resulting in making a System on Chip (SoC) at minimum cost and offering maximum functionality. With continuous shrinkage of gate length, Si MOSFET and SiGe Heterojunction Bipolar

Transistor (HBT) have occupied larger and larger fractions of the wireless communication market, for the application of PAs in receiver and base station designs [23] in the low GHz region. With the VLSI process maturing, it is no doubt that Si or SiGe-based devices will overtake the territory traditionally occupied by GaAs for lower GHz frequency applications (wireless communication), as it enjoys optimal performance at very low cost while InP-based devices such as HEMTS and pHEMT will be the first choice for their low noise performance and high frequency applications in the future.

2.2 Heterostructures

2.2.1 Homojunctions and Heterojunctions

A homojunction defines the interface formed by a semiconductor material combination of different polarity and of course the same band gap. This is often used in Silicon-based devices such as p-n junction diodes. An example of a homojunction pn-diode is shown in Figure 2.4.

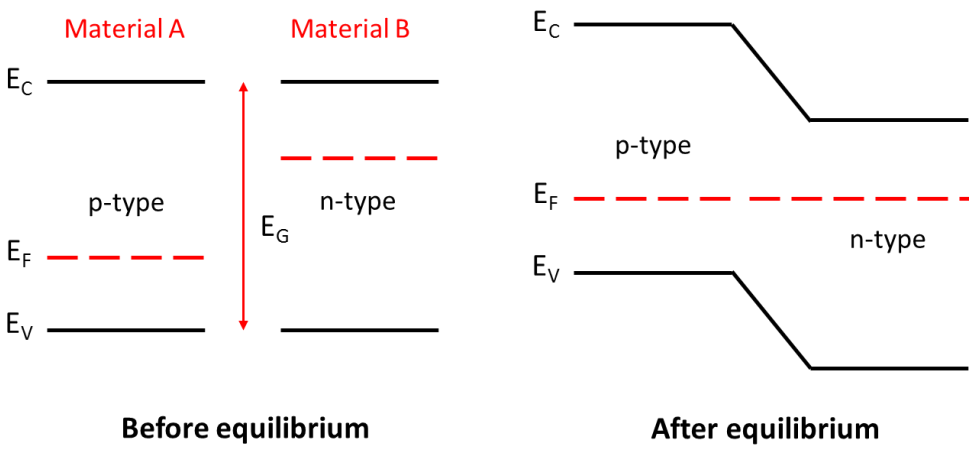


Figure 2.4 Homojunction band structure (pn junction)

This consists of two materials, A and B, having the same band gap (E_G) but different dopants (p-type and n-type); with the position of the Fermi level closer to the corresponding valence band (E_V) or conduction band (E_C) before equilibrium. When thermal equilibrium is established, the alignment of E_F will cause band bending at

either side of the junction, which introduces a built-in electric field barrier for holes and electrons, forcing them to travel in one direction only.

In contrast, a heterojunction is formed by two materials of dissimilar crystalline structure, with different band gaps. The situation in Figure 2.5 indicates the band diagram of a wide and narrow band gap semiconductors doped with donor atoms. n-type, with energy bands (E_G^A and E_G^B) before and after equilibrium.

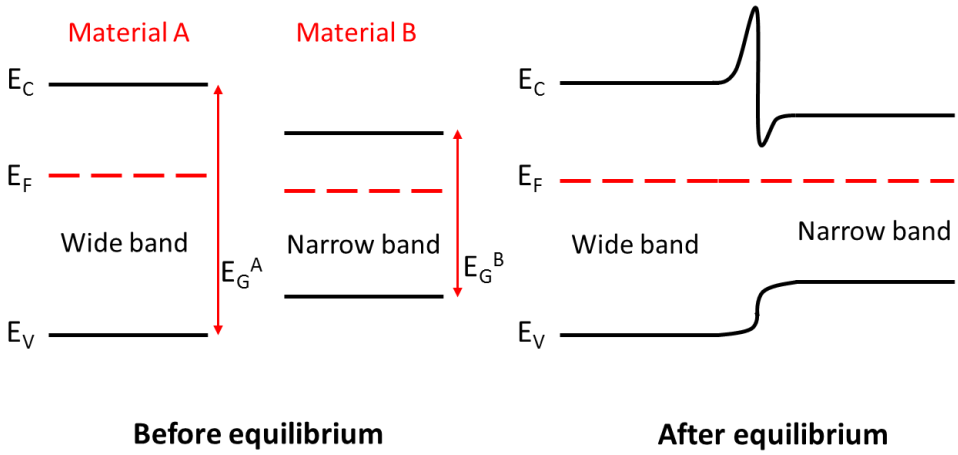


Figure 2.5 Heterojunction band structure (wide-narrow band combination) (n-type)

The difference of the band gaps results in a steep band bending, which leads to the formation of discontinuity regions at the junction. The degree of discontinuity can be very useful in modifying the carriers transport properties and the quality of the junction will have a great impact on the device performance. This is the prime reason why material growth technologies like MBE are often employed in precise heterojunction structures fabrication.

2.2.2 Lattice matched and pseudomorphic materials

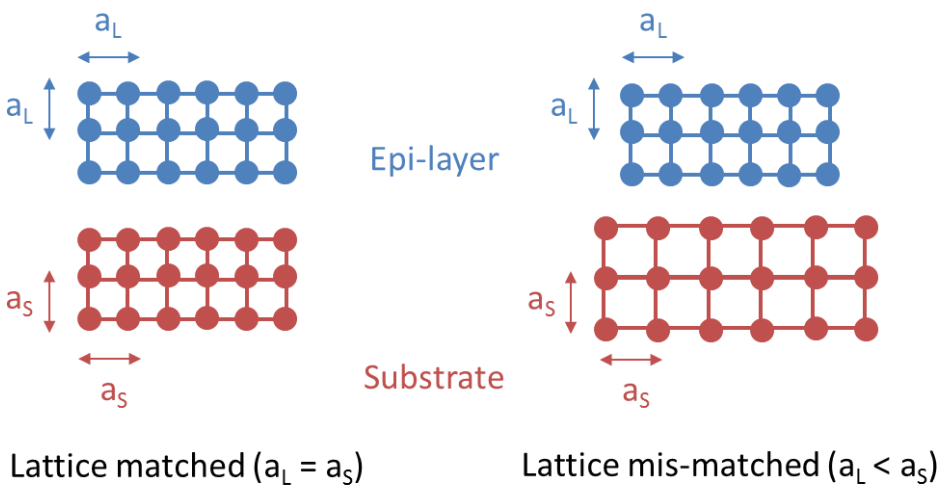


Figure 2.6 Lattice matched and mis-matched materials (Cubic crystal structure)

As mentioned previously, a heterojunction is created by the combination of two materials with different band gap. Formation of heterojunction can be divided into two scenarios, as shown in Figure 2.6. When the lattice constants of the substrate (a_S) and the over-layer (a_L) are similar or the same, their surface atoms are perfectly matched and so are called lattice matched. For lattice mis-matched combination, the surface atoms will shift to try to match each other, and a strain will be introduced onto the junction interface which magnitude is given by Equation 2.1:

$$\varepsilon = \frac{a_L - a_S}{as}$$ **Equation 2.1 [27]**

Where ε = strain between the two layers, a_L = lattice constant of the over-layer, a_S = lattice constant of the substrate layer. Depending on the relative difference of their lattice constants, the strain can be either compressive ($a_L > a_S$), in which the over-layer is compressed to fit the substrate or tensile ($a_L < a_S$), in which the over-layer is stretched.

For the lattice mis-matched case, dislocations occur at atomic level between the junctions if the interface strain is too strong, and this is intimately related to the

thickness of the over-layer, also called the critical thickness, as expressed by Equation 2.2:

$$h_C = \frac{a_S}{2\varepsilon}$$ **Equation 2.2 [27]**

Where h_C = critical thickness of the over-layer, a_S = lattice constant of the substrate and ε = strain between the two layers. Figure 2.7 and Figure 2.8 outline the relationship between over-layer thickness (d) and the critical thickness (h_C). When $d < h_C$, the over-layer will try to stretch itself to fit the substrate. As $d > h_C$, the excessive strain will be released by forming dangling bond between the over-layer and the substrate, which causes dislocations at the interface.

Thanks to advanced material growth technology like MBE, it is possible to grow an over-layer under the critical thickness and avoid the generation of dislocations. This brings about a new material concept called pseudomorphism, where the lattice constants of the two layers are identical in the x-y plane and mis-matched in the z-plane (e.g. InGaAs-InAlAs) and the dislocations problem is addressed by controlling the growth thickness.

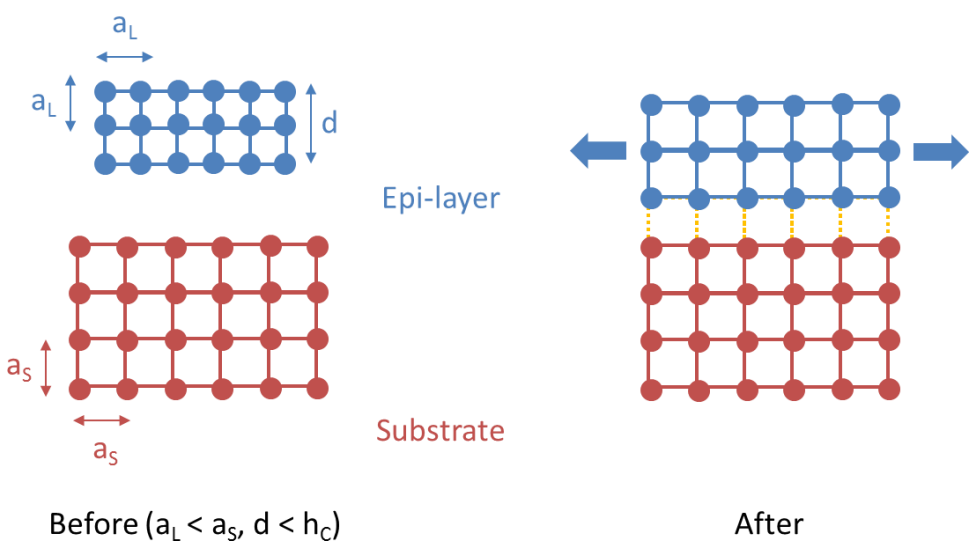


Figure 2.7 Formation of pseudomorphic layers with tensile strain

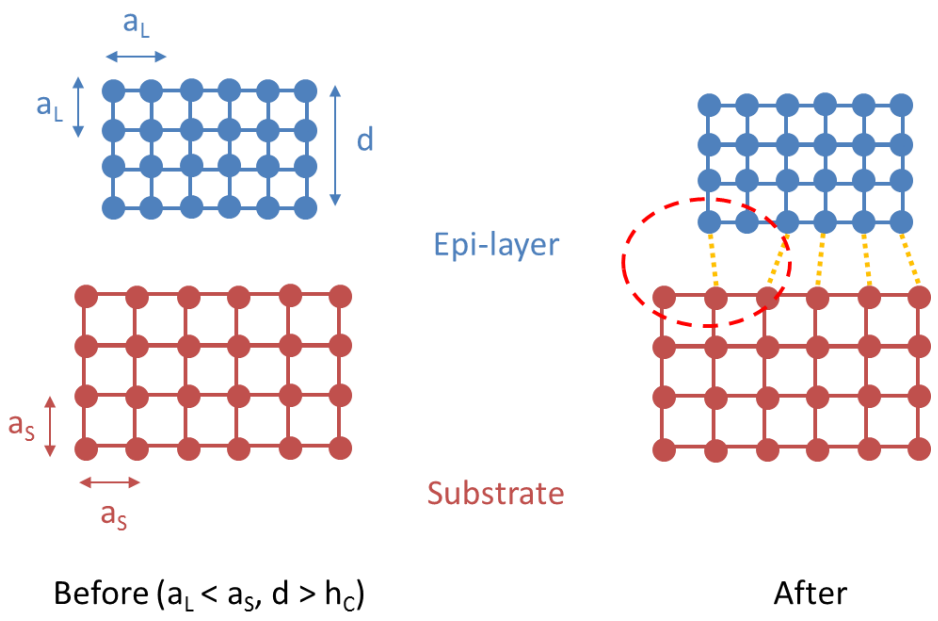


Figure 2.8 Formation of dislocation interface with plastic strain

2.2.3 Band discontinuities

As illustrated in Figure 2.5, the key feature of heterojunction is the formation of band discontinuities at the equilibrium state, which enables the modification of carrier transport properties and is one of the most important building blocks for HEMT-based devices. To understand the relationship between the degree of band bending and the band alignment, the Anderson's rule [24] is particularly useful.

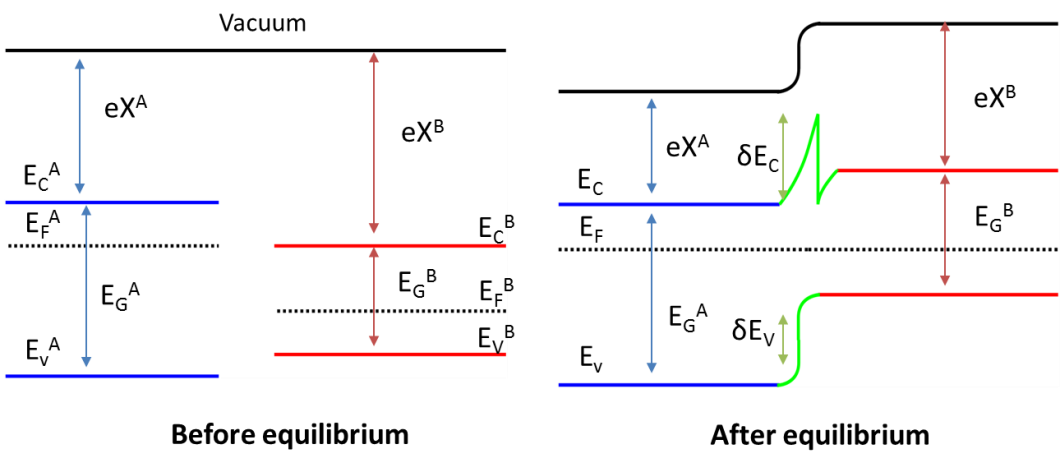


Figure 2.9 Energy band diagrams of a semiconductor heterojunction

Figure 2.9 shows the energy band diagrams of a simple heterojunction between a wide and a narrow band material, A and B. E_C, E_F and E_V are the conduction band, Fermi and valence band energy levels respectively. eX is the electron affinity from the bottom of conduction band to vacuum level. E_G is the difference in energy potential between E_C and E_V for each individual semiconductor material. The magnitude of E_G determines if the material is wide or narrow band, and is given by Equation 2.3:

$$E_G = E_C - E_V$$ **Equation 2.3 [27]**

When thermal equilibrium is established between A and B, energy band discontinuities (ΔE_C and ΔE_V) are introduced at the junction, due to the energy band difference (ΔE_G). Their relationships are given in Equation 2.4, 2.5 and 2.6:

$$\Delta E_C = \mathrm{e}(X_A - \mathrm{X}_B)$$ **Equation 2.4 [27]**

$$\Delta E_\mathrm{V} = \Delta E_G - \Delta \mathrm{eX}$$ **Equation 2.5 [27]**

$$\Delta E_G = E_G^A - E_G^B = \Delta E_C + \Delta E_V$$ **Equation 2.6 [27]**

In practice, the Anderson's rule can only provide an approximation for these parameters as experimental results are often very different, largely because of interface strain and dislocation, which can alter the electron transport properties at the interface.

2.2.4 Quantum well and 2 DEG

As previously discussed, a triangular shape discontinuity (ΔE_C) is formed at the heterojunction as depicted in Equation 2.6 and Figure 2.9, this is actually the simplest form of potential well/quantum well (QW) in heterojunction. Figure 2.10 illustrates the relationship between the triangular QW and a given grown structure in the conduction band diagram:

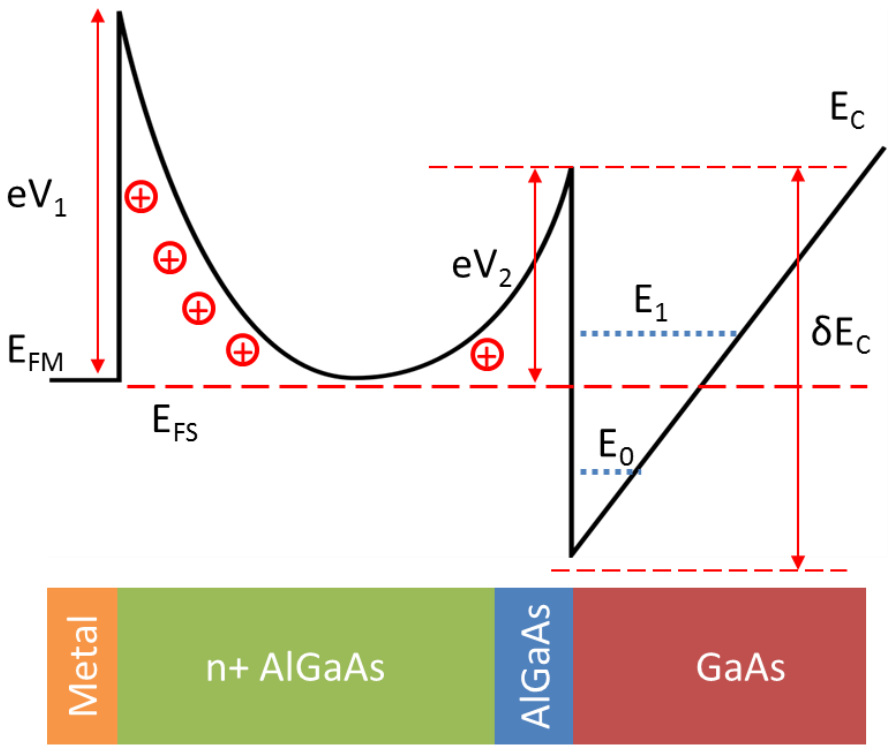


Figure 2.10 Triangular QW in AlGaAs/GaAs heterojunction. (Adapted from [104])

The structure contains a metal layer and three lattice-matched semiconductor layers. At thermal equilibrium, band bending occurs due to the alignment of Fermi levels between metal (E_{FM}) and semiconductors (E_{FS}), leading to the formation of potential barriers. Starting from the left, the metal and the heavily doped AlGaAs layer, with different band gaps, form a potential barrier (eV_1), also called a Schottky barrier. A thin layer of AlGaAs is sandwiched between the heavily doped AlGaAs and the undoped GaAs substrate, resulting in the formation of the second potential barrier (eV_2) and a triangular QW (ΔE_C) at the interfaces.

There are several mechanisms throughout the structure. As the junction thickness is considerably smaller than electron mean free path or the De Broglie wavelength, local energy sub-bands such as E_0 and E_1 will be created inside the QW due to energy quantisation [26]. Besides, the electron carriers sitting on the conduction band of the n+ AlGaAs layer will try to release their energy by jumping into the lower energy level inside the QW. The majority of them will be trapped inside the QW at the lowest energy sub-band (E_0) as it is lower than E_F.

As a result, electrons will be separated from the ionised donors in the n+ AlGaAs layer and effectively confined in a very thin planar surface parallel to the junction. The mobility of electrons is greatly enhanced due to the isolation of any impurity scattering within the n+ AlGaAs layer. The electron rich thin region is called a 2-dimensional electron gas (2DEG), which is vital for HEMT operation.

2.2.5 Bulk doping and δ-doped supply layer

So far all our discussions concentrate on bulk doping on the AlGaAs layer. However, the high doping level under the metal contact will cause severe leakage of the Schottky contact which degrades the device performance. A doping technique called δ-doping, also known as pulse or spike doping, is introduced to resolve this problem [25].

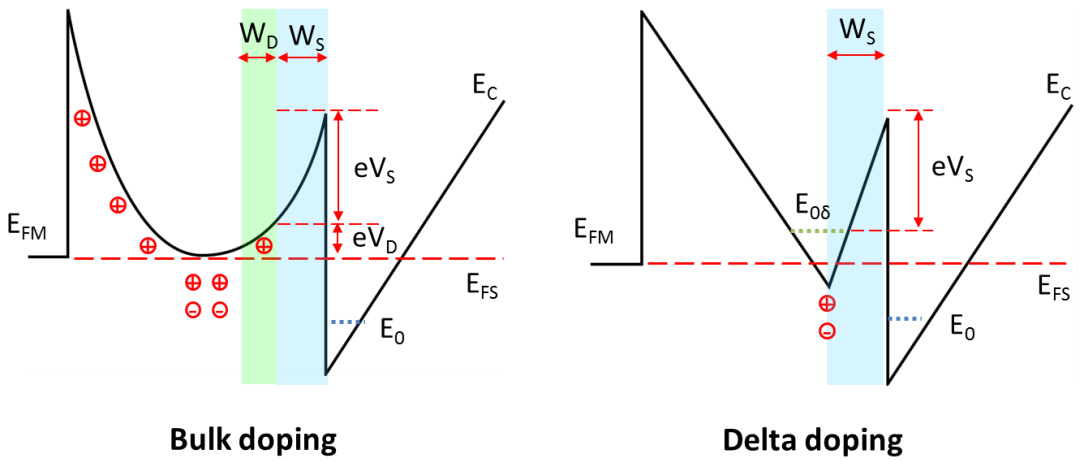


Figure 2.11 Bulk doping vs δ-doping (conduction band only) (Adapted from [25])

Figure 2.11 shows the comparison between bulk doping and δ-doping. In bulk doping structure, the dopants are uniformly distributed across the whole supply layer. The δ-doping scheme consists of a very high doping density (>$2x10^{12}$ cm^{-2}[25]) in a single atomic plane (typically ~10A [104]), which is usually synthesised using the MBE technique. Donor size quantisation occurs due to the formation of small QW, where $E_0\delta$ denotes the lowest energy sub band. The majority of donor impurities are located in $E_0\delta$ rather than being widely spreading over the supply layer in bulk doping.

The bulk doping case contains a very high potential barrier ($eV_D + eV_S$) with an effective depletion width ($W_D + W_S$). Free carriers in the conduction band must overcome this barrier by either jumping over or tunnelling through the barrier before they can settle down at the minimum energy state E_0 in the QW. For the δ-doping case, carriers can easily climb up or tunnel through the barrier due to potential drop in depletion layer. As a result, the carrier concentration in the QW for the δ-doping scheme is higher than that in the bulk doping scheme. Furthermore, the doping level immediately under the gate is much lower than that in bulk doping and this results in a much reduced gate leakage current. All the benefits of δ-doping make it a second-to-none choice for HEMT-based devices.

2.3 Metal to semiconductor interfaces

2.3.1 Schottky contact – rectifying contact

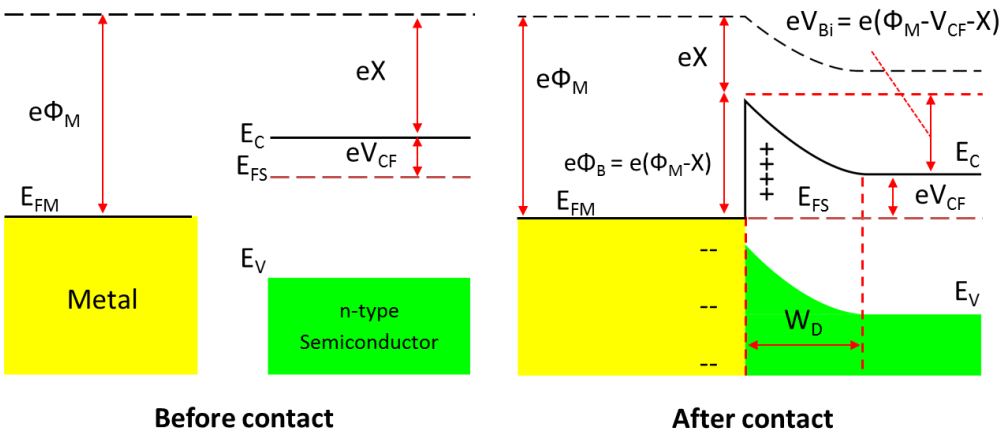


Figure 2.12 Schematic diagram of an ideal Schottky barrier before and after equilibrium

When a metal makes contact with a semiconductor, their respective Fermi-levels align with each other and establish a thermal equilibrium condition. The difference between their energy bands will lead to the formation of a potential barrier called a Schottky barrier. Carriers (electrons or holes) flowing through the metal-semiconductor interface will be rectified by this barrier and such contact is regarded as rectifying contact or Schottky contact.

Figure 2.12 presents the formation process of a metal to semiconductor contact. For simplicity, we assume that there is no surface state or interface between them and the semiconductor is uniformly doped. $e\Phi_M$ is the work function of the metal. eX is the electron affinity and eV_{CF} is the difference between conduction band and Fermi level on the semiconductor side. For simplicity, all formulae in the following discussions will be shown after dividing the electron charge "e", which is an indication of energy level with unit "eV".

It should be noted that there are three important parameters in the diagram, the barrier potential or barrier height ($e\Phi_B$), the build-in potential (eV_{Bi}) and the width of the depletion region (W_D). The barrier height is a measure of potential which restricts the electrons flowing from metal to semiconductor and is given by Equation 2.7:

$$\Phi_B = \Phi_M - X$$ **Equation 2.7 [27]:**

Theoretically, a metal with a higher Φ_M will yield a larger Schottky barrier on the same semiconductor. But this is not quite valid in practice. There is always an intermediate layer, contributing to the surface contamination of the metal or the surface states of the semiconductors such as native oxides or dangling bonds after etching, formed in between the metal and semiconductor. .

As a result, the Fermi level of a semiconductor will pin at the surface of the intermediate layer before equilibrium is achieved. The resulting barrier height is not sensitive to the change of metal work function and such phenomenon is called Fermi-level pinning [27]. Details of Fermi-level pinning effect will not be elaborated in this thesis. Generally speaking, the degree of Fermi-level pinning effect is heavily reliant on material. The relationship is illustrated in Figure 2.13:

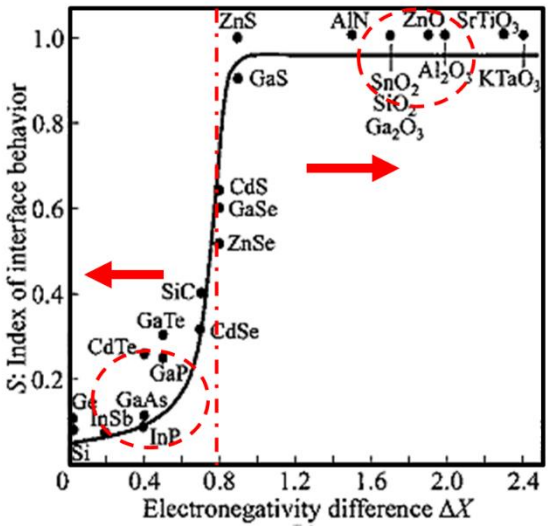


Figure 2.13 Index of interface behaviour S vs the electronegativity difference of common semiconductors [27]

The index of interface behaviour (S) is the change rate of barrier height in response to the change of electronegativity of metal (X_M) for a given semiconductor. X_M is defined as how easily an atom can attract an electron towards itself and is correlated with the metal work function. ΔX is the electronegativity difference between anion and cation of the semiconductor. From the diagram, it is clear that most of the ionic semiconductors (ZnS and GaS) or oxide compounds (Al_2O_3, SiO_2) are strong in electronegativity difference ($\Delta X > 1$) while III-V compound semiconductors (GaAs, InP) are much weaker ($\Delta X < 1$). In other words, the effective barrier height of III-V compound semiconductors is nearly independent of the change of metal work function and hence choices between different Schottky metallisation schemes exert insignificant impact on the resulting Schottky barrier height.

Additionally, electrons travelling from the semiconductor to the metal will need to overcome the depletion region (W_D) due to the built-in potential (V_{Bi}). The built-in potential of an ideal Schottky contact is simply given by Equation 2.8:

$$V_{Bi} = \Phi_M - V_{CF} - X \qquad \textbf{Equation 2.8 [27]}$$

However, due to the Fermi-level pining effect, the build-in potential V_{Bi} depends on the semiconductor properties rather than metal work function (Φ_M). The relationship

is given by Equation 2.9:

$$V_{Bi} = \Phi_B - kTln\frac{N_C}{N_D}$$ **Equation 2.9 [27]**

Where k is the Boltzmann's constant, T is the absolute temperature in Kelvin, N_C is the effective density of states in the conduction band and N_D is the semiconductor doping concentration (Assuming doping is uniform).

Assuming a sharp boundary condition, carriers are totally depleted within this region. The width of depletion region is governed by Equation 2.10:

$$W_D = \sqrt{\frac{2\epsilon(V_{Bi} - V_{\mathrm{GS}} - \frac{kT}{e})}{eN_D}}$$ **Equation 2.10 [28]**

Where ε is the dielectric constant of the semiconductor substrate, V_{GS} is the gate bias voltage and e is the electron charge. At room temperature, the term kT/e is about 0.026V. By applying a gate bias voltage, the depletion width can be adjusted and hence the channel thickness under the depletion region.

The depletion region can also be modelled as a parallel plate capacitor due to the charge interactions on both sides of the depletion region. The depletion region capacitance, also termed as the gate-channel capacitance (C_{GC}) in MESFET, is given by Equation 2.11:

$$C_{GC} = \frac{\varepsilon_D A}{W_D}$$ **Equation 2.11 [28]**

Where ε_D is the dielectric constant of the material in the depletion region and A is the area of the metal plate.

The great advantage of a Schottky contact is the ability to control current transport behaviour by applying forward or reverse biasing voltage. Figure 2.14 shows changes in energy band diagram under different bias conditions:

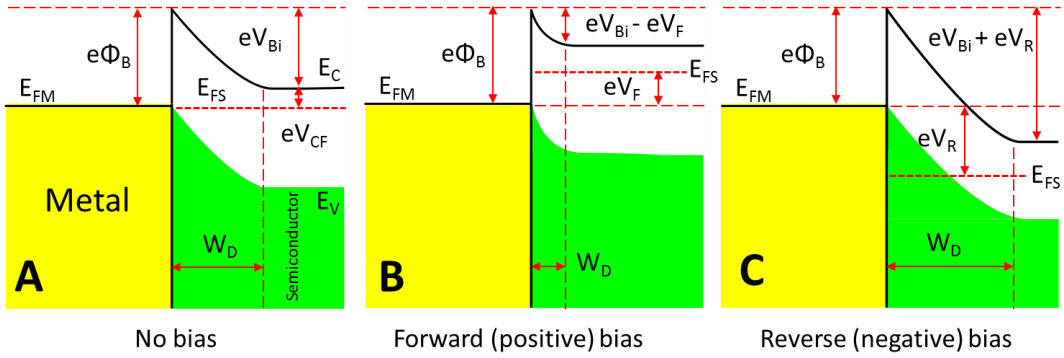


Figure 2.14 Change of ideal Schottky contact under different bias conditions

By default, if there is no bias applied on the junction, electrons travelling from either sides will have to climb up the potential barrier $e\Phi_B$ and eV_{Bi} respectively, where V_{Bi} is usually less than $e\Phi_B$, making the majority of electrons move from the semiconductor to metal direction. Hence, a Schottky contact is usually referred to as a Schottky diode for its rectifying behaviour. When a forward bias (eV_F) is applied, as shown in Figure 2.14B, the built-in potential and the depletion width will be reduced, which enables electrons to more easily climb up the slope. By contrast, a reverse bias (eV_R) will result in a larger potential barrier, which restricts electrons transport from the semiconductor side and facilitates those from the metal side. The unwanted electrons current from the metal is also regarded as a gate leakage current, which will be discussed later.

So far, the electron transport is assumed by clamping the slope in the depletion region. This mechanism is called Thermionic Emission (TE). It occurs when electrons have sufficient amount of activation energy to overcome the built-in potential. However, there is another mechanism called Thermionic Field Emission (TFE). The inherent principle is that when the thickness of potential barrier is thin enough, electrons are able to tunnel through the barrier region even if they do not have sufficient energy. This is an important concept for our next topic, Ohmic contacts.

2.3.2 Ohmic contact – non rectifying contact

As mentioned in the last section, the formation of a Schottky barrier between a metal to semiconductor results in a current rectifying behaviour, this encourages current flow from the semiconductor and blocks it from the metal. Carriers must have sufficient activation energy, in order to overcome the potential barrier from the metal side. Such a transport mechanism is called Thermionic Emissions (TE), and is shown in Figure 2.15A:

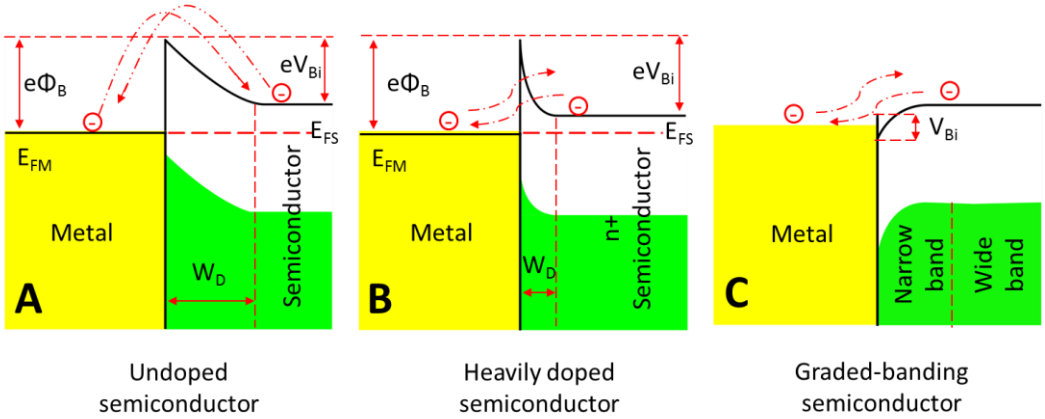


Figure 2.15 Electron transport mechanisms in different metal semiconductor junctions

In contrast, an Ohmic contact is a non-rectifying contact, which facilitates current flow in both forward and reverse direction with a near linear I-V characteristic. Unlike a Schottky contact, it does not control the current flow but provide a low resistance path to the external world. The formation of Ohmic contact can be achieved either by thinning the depletion region (W_D) or lowering the effective Schottky barrier (Φ_B).

For the first approach, the width of depletion region is reduced significantly by increasing the doping concentration in the semiconductor. As shown in Figure 2.15B, carriers are now able to tunnel through the thin barrier in both directions and such carrier transport mechanism is called Thermionic Field Emission (TFE). For the second approach, as illustrated in Figure 2.15C, a graded banding structure is created at the interface by inserting a narrow band gap material in the original junction. This causes band bending in the opposite direction and reduces the Schottky barrier height

significantly. However, the graded material can only be developed precisely by technologies such as MBE and entails more controls on the material growth step.

For fabricating real-life devices, Ohmic contacts can be divided into two categories, alloyed and non-alloyed contacts. Non-alloyed Ohmic contacts are formed by placing a metal layer (e.g. Ti/Au) onto a highly doped (>10^{-19}cm^{-3}) semiconductor, also known as the cap layer in a HEMT-structure [27]. This technique resembles the concept of depletion width thinning, which promotes the carrier tunnelling at the interface. It features several advantages including faster processing and reduced thermal damage during the heat treatment step. In terms of alloyed contacts, metals containing dopants (e.g. AuGe/Au) can be deposited onto the cap layer and subject to heat treatment. The diffusion of metal during the heat treatment process will drive the dopants into the semiconductor, forming a heavily doped region under the metal, which enables carrier tunnelling with reduced depletion width. However, the poor thermal stability of alloyed Ohmic makes it more sensitive to subsequent heat treatment steps during the process.

The selection of Ohmic contact schemes is dependent on structure. For structure with where access resistance to the channel region is relatively low, such as MESFET and HEMT, non-alloyed Ohmic contact would be mostly preferred for its better thermal stability. For pHEMT, the access resistance to a relative 2DEG region is often extremely high, due to the existence of large conduction band discontinuity, which makes alloyed Ohmic contact a compulsory choice, where a direct path will form directly to the 2DEG during the metal diffusion process as shown in Figure 2.16:

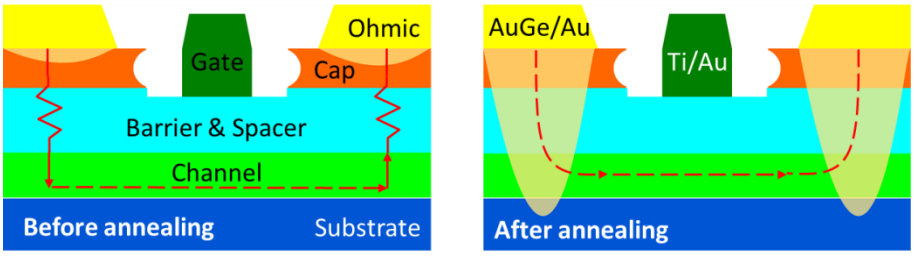


Figure 2.16 Formation of alloyed Ohmic contact in pHEMT structure

2.4 High speed devices and materials

For the past 50 years, the domination of Silicon (Si) in the semiconductor industry has mainly due to the continuous shrinkage of gate size and the lowering cost of production. With the rapid advancement of photolithography technology, Si-based microprocessors with transistor gate length as small as 22nm are now commercially available [29]. However, Si fails to accommodate the ever growing demand of faster electronics and is no longer capable of matching the requirements due its low electron mobility (μ) and so the average electrons drift velocity (V_D). The relationship between μ and V_D is given by Equation 2.12:

$$V_D = \mu E$$ **Equation 2.12 [28]**

Where E = the magnitude of applied electric field strength on the carrier, the electrons for our interests. As previously discussed in Figure 2.2, III-V compound semiconductors display extremely high electron mobility in comparison with Si in general. For example, the bulk mobility of GaAs is almost six times that of Si, making it the most sensible choice to fabricate high-speed devices. In the following sections, some popular GaAs-based device structures, such as MESFET, HEMT and pHEMT for microwave applications will be described. A summary of their device epitaxial structure is shown in Table 2.1:

<table>
<tr><th>Layer</th><th>MESFET</th><th>HEMT</th><th>pHEMT</th></tr>
<tr><td>Cap</td><td colspan="3">n^+ GaAs</td></tr>
<tr><td>Barrier</td><td rowspan="4">n GaAs</td><td>n AlGaAs</td><td>n AlGaAs</td></tr>
<tr><td>Doping</td><td colspan="2">n^+ AlGaAs or Si δ doping</td></tr>
<tr><td>Spacer</td><td colspan="2">Undoped AlGaAs</td></tr>
<tr><td>Channel</td><td>Undoped GaAs</td><td>Undoped InGaAs</td></tr>
<tr><td>Buffer</td><td colspan="3">p^- GaAs</td></tr>
<tr><td>Substrate</td><td colspan="3">Semi-Insulating GaAs</td></tr>
</table>

Table 2.1 Comparison between GaAs-based MESFET, HEMT and pHEMT

2.4.1 Device structures: MESFET

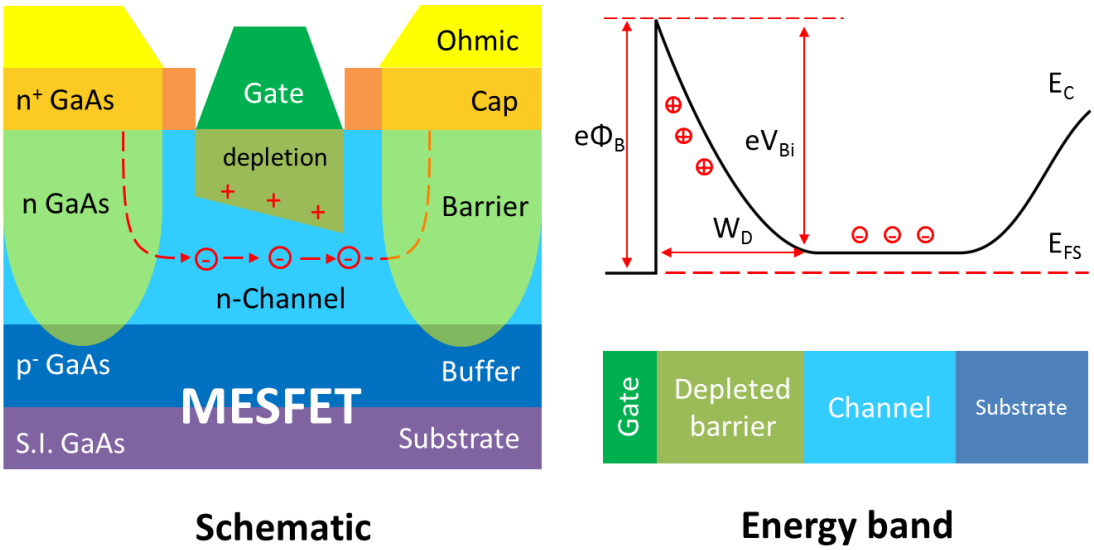


Figure 2.17 Schematic and band diagram of a GaAs MESFET. (Adapted from [104])

The Metal Semiconductor Field Effect Transistor (MESFET) is one of the simplest but mature GaAs-based transistors, and was dominant in RF electronics since its invention. The MESFET structure can be prepared by a simple ion implantation or MBE technique. Figure 2.17 illustrates the schematic diagram of a recess-gate MESFET and the corresponding conduction band diagram. The structure consists of four layers of lattice-matched GaAs with different doping profiles. A p^- buffer is grown on top of the GaAs substrate to isolate the defects and to also act as the device insulation layer to stop potential leakage path. A lightly doped GaAs, called a barrier layer, sits above the buffer to provide an interface for the formation of a metal Schottky gate. Finally, a heavily doped cap layer is provided for Ohmic contacts and to create a low resistance path from the drain and source contacts to the channel region. It should be noted that the MESFET is normally an n-channel device as the mobility of electrons is much higher than that of holes for III-V materials.

Both MESFET and MOSFET rely on the field effect modulation of the current conducing channel underneath the gate. However, in a MESFET, the formation of the electron channel is due to electrons from the n-type dopants, rather than the depletion

of holes as in a MOSFET. The Schottky gate contact, sitting above the n-channel region, controls the flow of electrons between the Ohmic contacts by forming a depletion region underneath itself. The width of the depletion region (W_D) can be altered by applying an electric field between the gate and the channel, where the thickness of the channel is modified accordingly, in response to the change of electric field.

There is no doubt that the design of MESFET is simple and efficient, outperforming the Si MOSFET in microwave electronics. However, as the operating frequency increases, the degradation of device performance becomes significant due to the impact of electrons scattering.

$$\mu = \frac{q\mathrm{t_m}}{\mathrm{m}^*}$$ **Equation 2.13 [28]**

Equation 2.13 illustrates the relationship between the carrier mobility (μ), carrier effective mass (m*) and the carrier mean free time (t_m). The lower m* of GaAs compared with Si is the primary reason for its excellent mobility. The term t_m defines the electron travelling time between successive collisions in the semiconductor crystal structure. It is correlated with the amount of scattering event in a given semiconductor, which is dependent on material and temperature.

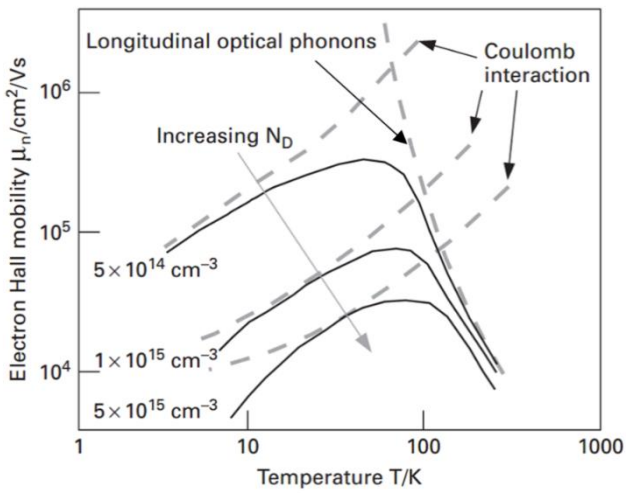


Figure 2.18 The relationship between electron mobility, temperature, doping concentration and scattering mechanisms in GaAs [26]

Figure 2.18 illustrates the change of electron mobility in terms of the variation of temperature and doping profile. At higher temperatures, the electron mobility is mainly affected by vibration of the crystal lattice, called phonon scattering, while when the temperature approaches 0 K, it is suppressed and replaced by another mechanism called coulomb scattering, which is due to the interaction between electron and the ionised dopants. The electron mobility of GaAs also shows an inversely proportional relationship to the amount of dopants. This greatly challenges MESFET design. Firstly, the channel current density is proportional to the doping concentration. Any enhancement of current flow also leads to the degradation of electron mobility. Secondly, the depletion width of the Schottky barrier is inversely proportional to the doping concentration. Any attempt to increase the dopants underneath the metal gate will also cause a higher gate leakage current. The optimisation of MESFET structure is always the result of a balance among the channel current, electron mobility and gate leakage current, hindering enhancement of the device performance.

2.4.2 Device structures: HEMT

The high electron mobility of III-V materials leads to the development of a better device than the MESFET although its performance is constrained by impurity scattering in the device channel. The High Electron Mobility Transistor (HEMT), as a successor to the MESFET, was proposed to resolve the issue and unleash the full potential of III-Vs. The HEMT is also known in many different terms such as Modulation Doped FET (MODFET), Two-Dimensional Electron Gas FET (TEGFET) and Selectively Doped Heterojunction Transistor (SDHT) [27]. It provides high electron mobility with its 2DEG structure thanks to the presence of heterojunctions. A general GaAs-based HEMT structure and its conduction band diagram are illustrated in Figure 2.19.

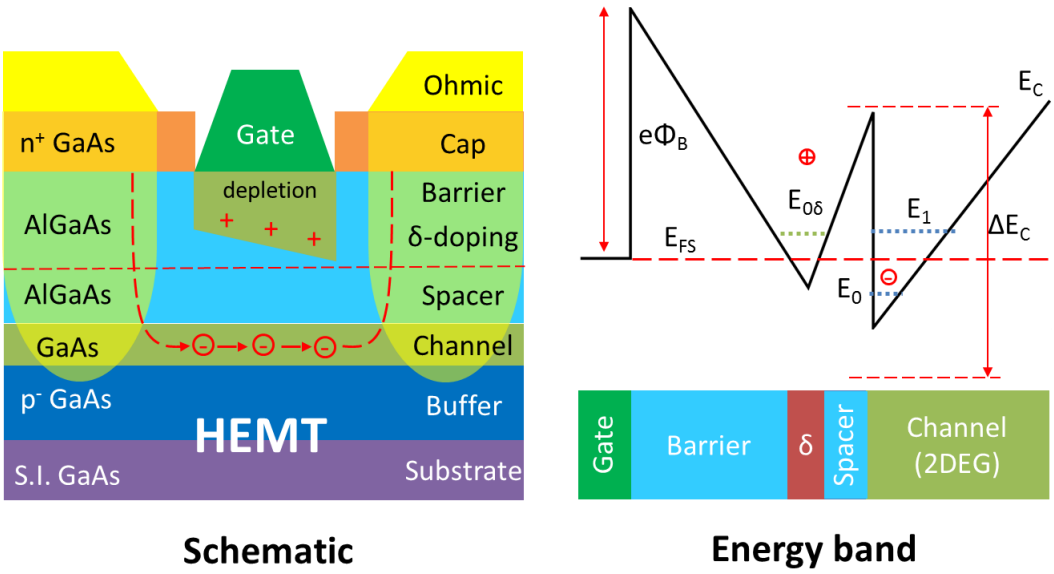


Figure 2.19 Schematic and band diagram of GaAs HEMT. (Adapted from [25])

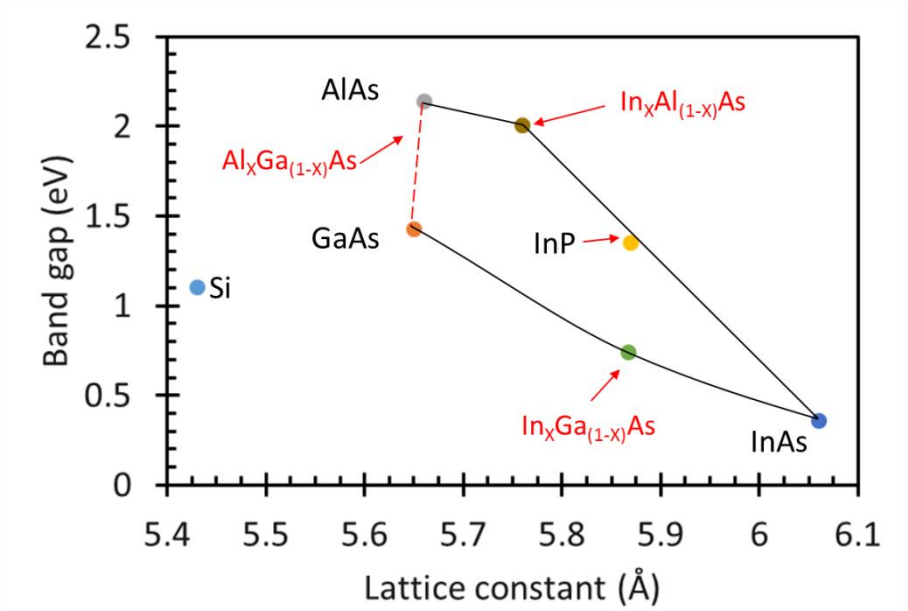


Figure 2.20 Band gap and lattice constant of Si, GaAs and InP-based compounds (300K) (Data source: [30]-[32])

Both HEMT and MESFET share some similarities in the device structure, including the substrate, buffer and cap layers, the use of Ohmic and Schottky contacts. The key difference in the HEMT is the introduction of three new separated device layers (barrier, spacer and channel) and the use of AlGaAs-GaAs materials combination. As shown in Figure 2.20, $Al_xGa_{(1-x)}As$ is a ternary compound of GaAs and AlAs, where x and 1-x represent the composition ratio. By changing the concentration of AlAs

composition, its band gap can be tuned as high as 2.1eV while remaining almost lattice matched to GaAs. The wide band gap property of AlGaAs makes it the optimal material for constructing 2DEG at the channel layer interface and the formation of Schottky gate contact. As previously discussed, the Schottky barrier height mainly depends on the properties of the semiconductor. The use of a wide band gap barrier layer in HEMT enhances the effectiveness of the barrier and hence helps to suppress the gate leakage current.

The formation of 2DEG is outlined in the conduction band diagram in Figure 2.19. The band alignment of the wide and narrow band materials (AlGaAs and GaAs) results in a conduction band discontinuity, where energy quantisation occurs inside the triangular QW. Electrons in the δ-doping layer can then tunnel through the thin potential barrier and get trapped into the triangular QW, leading to the formation of a 2DEG between the channel and the spacer layer. The interaction between ionised donors and the electrons carrier in the 2DEG is isolated by the highly resistive spacer layer. This feature helps to maintain the electron mobility without compromising the electron density in the channel.

2.4.3 Device structures: pHEMT

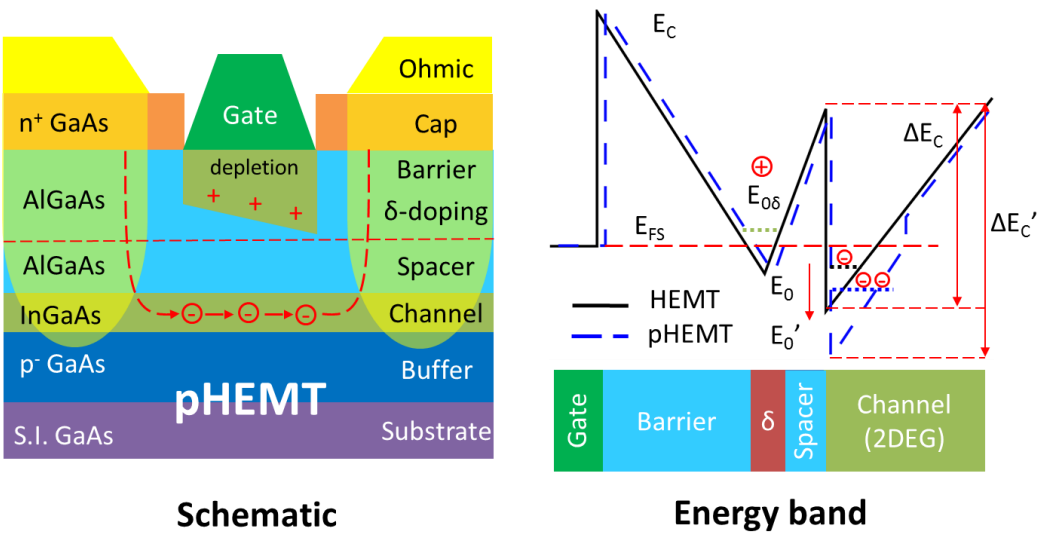


Figure 2.21 Schematic and band diagram (blue dot line) of GaAs pHEMT. (Adapted from [26])

The MBE technique makes it possible to grow high quality lattice mismatched interface on III-V materials. Pseudomorphic High Electron Mobility Transistor (pHEMT) utilities the concept of strained heterojunction to improve the carrier transport in the 2DEG. As shown in Figure 2.21, both the HEMT and pHEMT are very similar in terms of structure, except for the use of the ternary compound $In_xGa_{(1-x)}As$, where x and 1-x denote the compound ratio between InAs and GaAs, as a replacement of the original GaAs. The unique feature of $In_xGa_{(1-x)}As$ is its tuneable band gap and lattice constant. This enables pHEMT to outperform the HEMT in the following aspects:

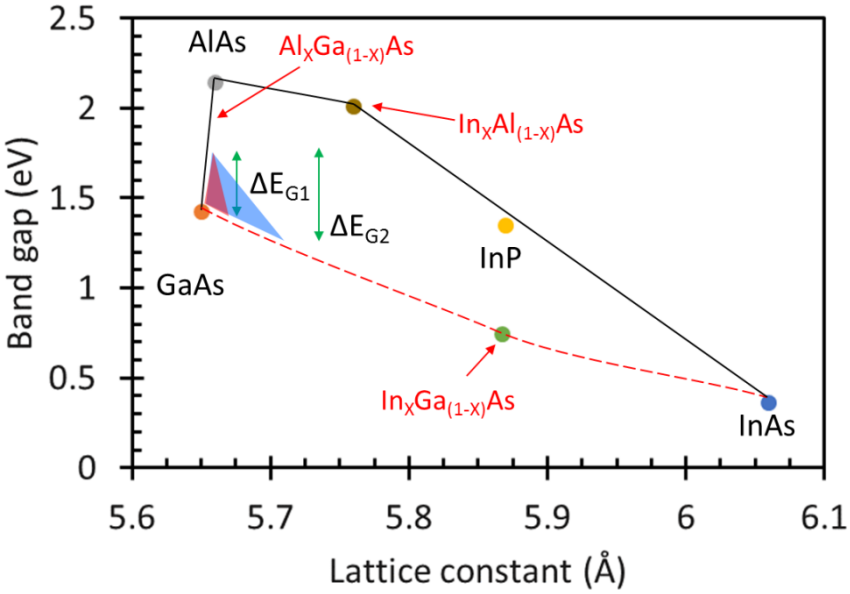


Figure 2.22 Comparison of the energy band gap and lattice properties between AlGaAs-GaAs (red region) and AlGaAs-InGaAs (blue region) heterojunction (300K) (Data source: [30]-[32])

In terms of band structure, by tolerating a pseudomorphic approach, the AlGaAs-InGaAs combination features better flexibility in the energy band optimisation and energy band difference (ΔE_G), compared with the AlGaAs-GaAs pair, as indicated in Figure 2.22. The larger band gap difference leads to a higher conduction band discontinuity at the heterojunction, driving more carriers inside the QW and thus increasing the carrier concentration in the 2DEG. The formation of a double heterojunction, as shown in Figure 2.21, also helps to reduce the carrier injection through the buffer or substrate and also eliminate parasitic leakage paths.

In terms of material properties, the pseudomorphic structure allows the use of lower AlAs content in the supply layer, which helps to reduce DX centres density [26]. More importantly, the effective electron mass (m*) of $In_xGa_{(1-x)}As$ is less than half that of GaAs, suggesting higher electron mobility and enhanced carrier transport properties.

2.4.4 Material systems: AlGaAs-InGaAs and InAlAs-InGaAs

Material	GaAs pHEMT	InP pHEMT
Cap	GaAs	$In_{0.53}Ga_{0.47}As$
Supply	$Al_{0.15}Ga_{0.85}As$	$In_{0.52}Al_{0.48}As$
δ-doping		
Spacer	$Al_{0.15}Ga_{0.85}As$	$In_{0.52}Al_{0.48}As$
Channel	$In_{0.15}Ga_{0.85}As$	$In_{0.70}Ga_{0.30}As$
Buffer	GaAs	$In_{0.52}Al_{0.48}As$
Substrate	GaAs	InP

Table 2.2 Comparison between common GaAs and InP-based pHEMT structures

A comparison is made between practical GaAs and InP-based pHEMT in Table 2.2. For the InP-based compounds, GaAs is replaced by $In_{0.53}Ga_{0.47}As$ compounds and AlGaAs by $In_{0.52}Al_{0.48}As$ respectively, the latter lattice matched to InP substrate. . What makes InP better than GaAs pHEMT is mainly due to the higher electron mobilities and higher band discontinuities in this material system. Figure 2.23 illustrates the latter point.

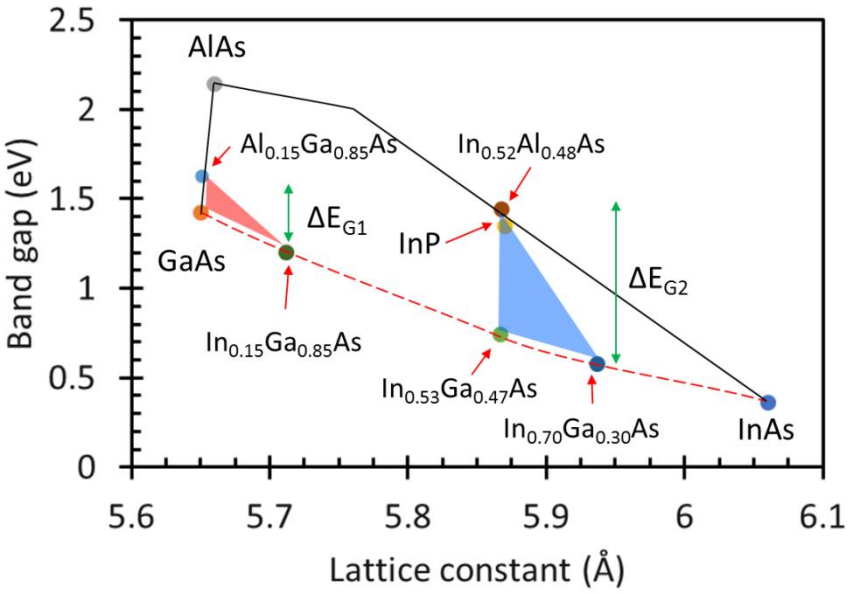


Figure 2.23 Comparison of the energy band gap and lattice properties between AlGaAs-InGaAs and InAlAs-InGaAs heterojunctions (300K) (Data source: [30]-[32])

In the $Al_xGa_{(1-x)}As$-$In_{0.15}Ga_{0.85}As$ system, the Al content must be kept <0.26 in order to avoid serious DX centres and traps. This places a limit on the effective ΔE_G [28]. On the other hand, the InAlAs-InGaAs heterojunction has a larger lattice constant, which allows the use of InGaAs channel with much higher Indium content without compromising the lattice constant. The increment of In content in $In_xGa_{(1-x)}As$ will result in a higher conduction discontinuity and lower electron effective mass, leading to an enhancement of the carrier concentration and mobility. This is one of the main reasons why InP pHEMTs are preferred for high speed microwave electronics.

2.5 Principle of operations

2.5.1 DC-IV output characteristics

In this section, the basic formulae governing HEMT structure operation will be explored, and the discussion will mainly focus on the depletion mode pHEMT, which is the most common type of device. The basic working principle of a pHEMT is very similar to its MESFET counterpart as they both rely on the modulation effect of a gate electric field on the current conducting channel. However, in MESFET, the current modulation is done by changing the channel thickness. In HEMTs, the applied electric field alters the carrier density in the channel while its thickness remains virtually unchanged, and all current flows are constrained on a planar surface as suggested by the name 2DEG [27]. A conduction band diagram, along with its structural plot is shown in Figure 2.24, and this will be a useful reference in the following discussions.

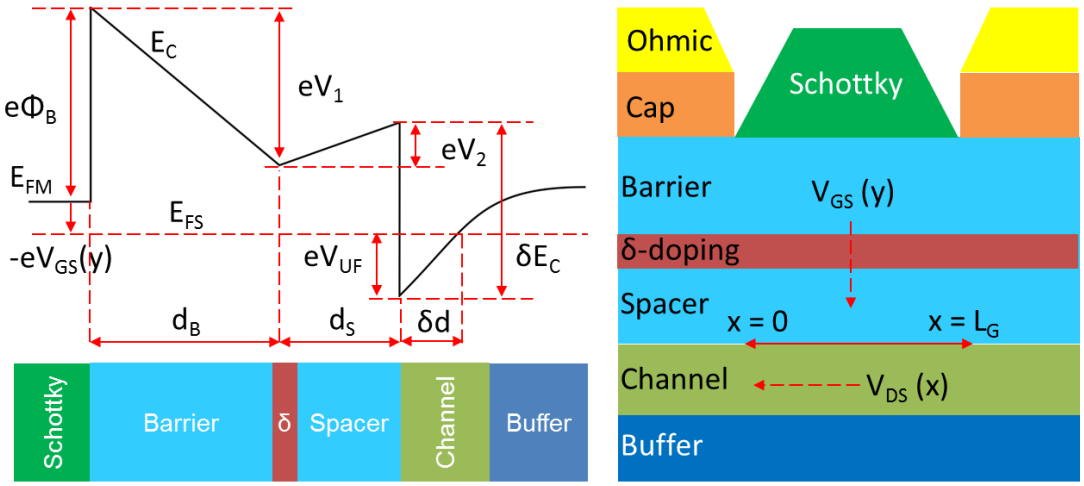


Figure 2.24 Conduction band diagram (left) and structural diagram (right) of typical depletion mode pHEMT with Δ-doping scheme. (Adapted from [104])

In the diagram, eV_1 and eV_2 are the potential drops across the barrier and spacer depletion regions with respect to δ-doping region. eV_{UF} is the amount of QW buried under the semiconductor Fermi level (E_{FS}). Their relationship in related to the band structure is shown in Equation 2.14, Equation 2.15 and Equation 2.16 respectively:

$$V_1 = \frac{en_m d_B}{\varepsilon_B}$$ **Equation 2.14 [104]**

$$V_2 = \frac{en_s d_S}{\varepsilon_B}$$ **Equation 2.15 [104]**

$$V_{UF} = \frac{en_s \Delta d}{\varepsilon_B}$$ **Equation 2.16 [104]**

Where n_m = electron charge density located on the metal gate interface, n_s = sheet electron charge density located in the 2DEG, ε_B = dielectric constant of the barrier layer. d_B and d_S is the thickness of barrier and spacer layers. Δd is the effective thickness of 2DEG within the channel layer, which is sensitive to the change of QW profile [104]. V_{GS} (y) is the vertical voltage potential between the Schottky gate and the conductive channel. V_{DS}(x) is the horizontal voltage potential across the channel between the drain and source contacts.

The net charge distribution of the system can be expressed as below:

$$N_d^+ = n_m + n_{par} + n_s \approx n_m + n_s$$ **Equation 2.17 [104]**

Where N_d^+ = the positive donor charge density in the δ-doping layer, n_{par} = the negative electron charge density located in the barrier layer, which is minimal under normal device operation.

To simplify the discussion, let us consider a system with V_{DS} (x) = 0. After dividing "e", the equilibrium of voltage potential can be expressed as below:

$$-V_{\mathrm{GS}} + \phi_B - V_1 + V_2 - \frac{\Delta E_C}{e} + V_{UF} = 0$$ **Equation 2.18 [104]**

After substituting the Equation 2.14, Equation 2.15, Equation 2.16 and Equation 2.17 into Equation 2.18, the relationship between 2DEG sheet carrier density (n_s) and the gate bias voltage (V_{GS}) can be obtained:

$$n_s = \frac{eN_d^+ d_\Delta + \epsilon_B [V_{GS} - \left(\phi_B - \frac{\Delta E_C}{e}\right)]}{e(d_B + d_s + \Delta d)}$$ **Equation 2.19 [104]**

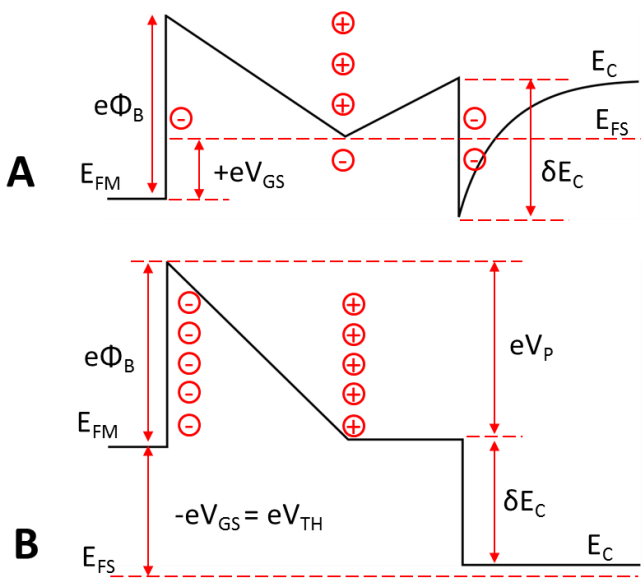


Figure 2.25 Change of energy band diagram for pHEMT with Δ-doping scheme under forward bias (A) and reverse bias (B). (Adapted from [104])

The influence of gate bias on n_S is illustrated through change of band diagram, as shown in Figure 2.25 A and B. By applying a forward bias ($V_{GS} > 0$), band bending occurs and E_{FS} starts shifting upwards by the amount of $+eV_{GS}$. The amount of heterojunction dropping below the E_{FS} will increase accordingly. The resultant n_s will enhance with more electrons trapped in the 2DEG region. Eventually, E_{FS} will be pushed towards the bottom of δ-doping layer, forming a low mobility parasitic channel called 3DEG [27]. This point defines the upper limit of the V_{GS}.

Conversely, the reverse bias ($V_{GS} < 0$) shows the opposite effect, E_{FS} will be pulled downwards, approaching the bottom of 2DEG and significantly reducing n_s. When the term $-eV_{GS}$ is sufficiently large, all electron charge in the 2DEG will be depleted, leaving only a potential barrier between the metal and the δ-doping layer. This point defines the lower limit of the V_{GS} as the channel is totally pinched-off when $n_s = 0$. The corresponding $-V_{GS}$ is referred to as threshold voltage (V_{TH}), which can be calculated by deriving from Equation 2.19:

When $n_s = 0$, $-V_{GS} = V_{TH}$

$$V_{\mathrm{TH}} = \Phi_B - \frac{\Delta E_C}{q} - V_P$$ **Equation 2.20 [104]**

$$V_{\mathrm{P}} = \frac{qN_d^+ d_{\mathrm{B}}}{2\varepsilon_0 \varepsilon_B}$$ **Equation 2.21 [104]**

Where V_P = the potential drop across the barrier and spacer layers. V_P is also called pinch-off voltage in the MESFET case [27, 104], which is rather confused. It is important to point out that the concept of pinch-off and threshold voltage are the same; they refer to the point where the conductive channel is completely pinched-off due to the depletion of sheet carrier in the 2DEG.

Therefore, V_{TH} determines the device ON or OFF behaviour. If $V_{TH} < 0$, which means that the device is normally ON when $V_{GS} = 0$, and this is also called a depletion mode (D-mode) device. By contrast, when $V_{TH} > 0$, this is classified as an enhancement mode (E-mode) device.

Next, the relationship between the channel current (I_{DS}) and drain-source electric voltage potential (V_{DS}) will be investigated by applying the velocity-field assumption [27]. When $V_{DS}(x) > 0$, Equation 2.19 can be rewritten:

$$n_s(x) = \frac{\epsilon_B[V_{GS} - V_{TH} - V_{DS}(x)]}{e(d_B + d_s + \Delta d)}$$ **Equation 2.22 [26]**

The I_{DS} at any location along the x-direction is given by:

$$I_{DS}(x) = W_G Q_n(x) V_D(x) = eW_G n_s(x) \mu_n E_x(x)$$ **Equation 2.23 [26]**

Where W_G = gate width, $Q_n(x)$ = the channel charge and V_D = the carrier drift velocity, they are governed by the following equations:

$$Q_n(x) = en_s(x)$$ **Equation 2.24 [26]**

$$V_D(x) = \mu_n(x)E(x) = \mu_n \frac{dV_{DS}(x)}{dx}$$ **Equation 2.25 [26]**

Where E(x) = the electric field strength along the x-direction.

Substituting Equation 2.22 and Equation 2.25 into Equation 2.23, the I_{DS} can be obtained by integrating from x = 0 to x = L_G, as shown in Equation 2.26:

$$I_{DS} = \frac{\epsilon_B \mu_n W_G}{(d_B + d_S + \Delta d) L_G} \int_0^{L_G} [V_{GS} - V_{TH} - V_{DS}(x)] \frac{dV_{DS}(x)}{dx} dx$$ **Equation 2.26 [26]**

The gate-channel capacitance (C_{GC}) can be expressed in terms of n_s, V_{GS} and Δd:

$$C_{GC} = eW_G L_G \frac{dn_s}{dV_{GS}} = \frac{\epsilon_B A}{d_B + d_s + \Delta d}$$ **Equation 2.27 [26]**

By substituting Equation 2.27 into Equation 2.26, the I_{DS} equation finally becomes:

$$I_{DS} = \frac{W_G}{L_G} \mu_n C_{GC} [(V_{GS} - V_{TH}) V_{DS} - \frac{V_{DS}^2}{2}]$$ **Equation 2.28 [26]**

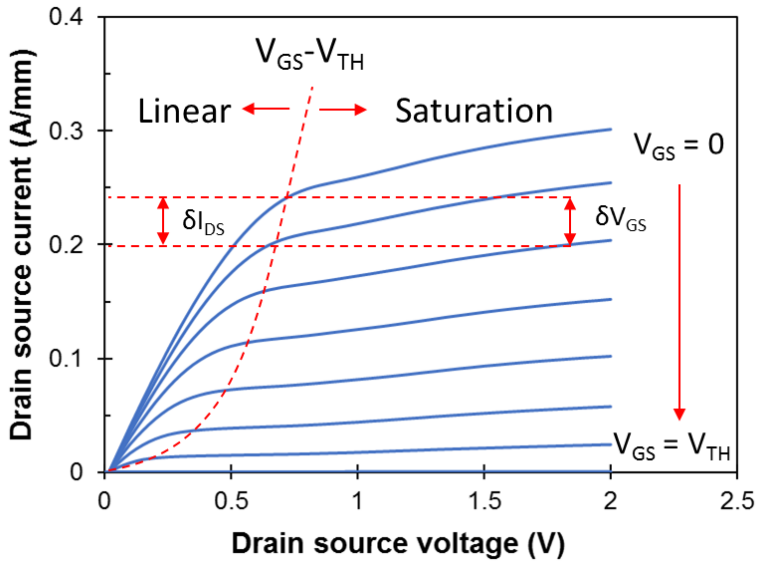


Figure 2.26 I-V characteristic of D-mode pHEMT – normalised I_{DS} against V_{DS} and V_{GS}

A graphical representation of Equation 2.28 is illustrated in Figure 2.26. The channel current behaviour under the influence of V_{DS} can be described using a two-region model [26]. At low-field, the electron carrier in the channel travels with a field-dependent drift velocity (V_D), which corresponds to the constant mobility region. When $V_{DS} < (V_{GS}-V_{TH})$, the term $V_{DS}^2/2$ can be ignored and the device acts like a voltage controlled resistor, and this is referred to as the linear region. The corresponding channel current (I_{DS}) and resistance (R_{DS}) are given by Equation 2.29 and Equation 2.30 respectively:

$V_{DS} < (V_{GS}-V_{TH})$, $$I_{DS} = \frac{W_G}{L_G}\mu_n C_{GC}(V_{GS} - V_{TH})V_{DS}$$ **Equation 2.29 [26]**

$$R_{DS} = \frac{L_G}{W_G \mu_n C_{GC}(V_{GS} - V_{TH})} = \frac{\Delta V_{DS}}{\Delta I_{DS}}$$ **Equation 2.30 [26]**

At high-field, in which case $V_{DS} \geq (V_{GS}-V_{TH})$, the V_D becomes saturated; the device starts operating in the constant velocity region. The channel is pinched-off at the drain terminal, resulting in a very small change in I_{DS} with further increase in V_{DS}. This corresponds to the saturation region in Figure 2.26 and it is given by Equation 2.31:

$V_{DS} \geq (V_{GS}-V_{TH})$, $$I_{DS} = \frac{W_G}{L_G}\mu_n C_{GC}(V_{GS} - V_{TH})^2$$ **Equation 2.31 [26]**

Figure 2.27 illustrates the change of I_{DS} as a function of V_{GS} in the saturation region. The channel is completely pinched off at $V_{GS} = V_{TH}$, where $I_{DS} \approx 0$ and $n_s = 0$. As the V_{GS} increases, I_{DS} shows a parabolic profile as predicted in Equation 2.31, which is attributed to population of electron carriers in the 2DEG. When $V_{GS} >> V_{TH}$, I_{DS} shows a linear profile eventually.

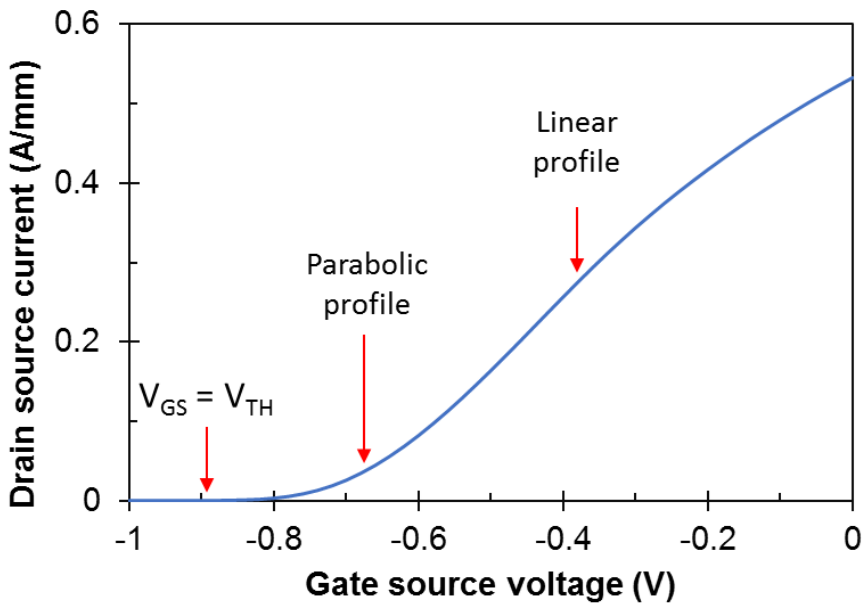


Figure 2.27 I-V characteristic of D-mode pHEMT at saturation region– normalised I_{DS} against different V_{GS} under a fixed V_{DS}

The transconductance (G_m) is defined as the change in I_{DS} in accordance with the change in V_{GS} for a given V_{DS}, which also a measure of the device current gain. G_m is governed by Equation 2.32 and it can be further derived into Equation 2.33 and Equation 2.34, depending on the operating conditions:

In general,
$$G_m = \frac{\Delta I_{\mathrm{DS}}}{\Delta V_{GS}}$$
Equation 2.32 [26]

In the linear region,
$$G_m = \frac{W_G \mu_n C_{\mathrm{GC}} V_{DS}}{L_G}$$
Equation 2.33 [26]

In the saturation region,
$$G_m = \frac{W_G \mu_n C_{GC} (V_{GS} - V_{\mathrm{TH}})}{L_G}$$
Equation 2.34 [26]

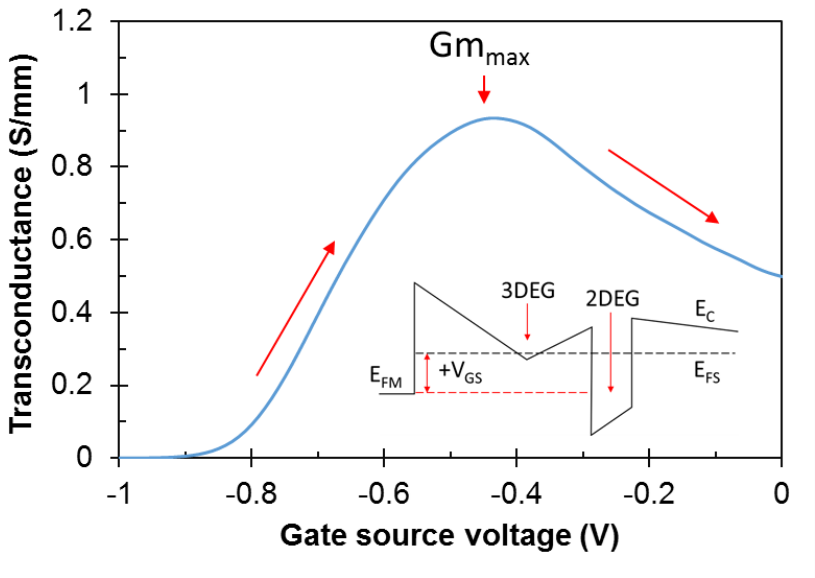


Figure 2.28 I-V characteristic of D-mode pHEMT – normalised transconductance against different V_{GS} at a fixed V_{DS} and the change of conduction band diagram at maximum Gm

Figure 2.28 illustrates the bell shape profile of the transconductance for a typical pHEMT device operating in the saturation region. Initially, G_m fallows the similar trend as I_{DS} with stronger V_{GS} but it declines very rapidly after the peak value (Gm_{max}). Recall the discussion in Figure 2.25A, the E_{FS} will touch the bottom of δ-doping layer, resulting a low mobility parasitic channel called 3DEG [26], as shown in the inset of Figure 2.28. The formation of low mobility 3DEG will shield the interaction between the gate and the 2DEG, causing significant degradation in the overall mobility and the effective Gm.

2.5.2 Small signal equivalent circuit

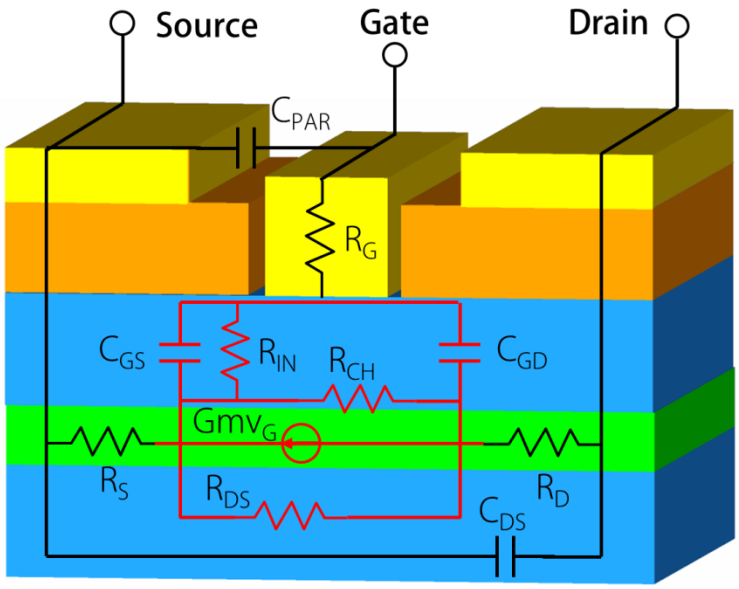


Figure 2.29 Small signal equivalent circuit for pHEMT (Intrinsic elements in red, extrinsic elements in black) (Extrinsic inductors are neglected) (Adapted from [27])

In real world applications, the device DC and RF performances are often deteriorated due to parasitic elements. A small signal equivalent circuit model, as shown in Figure 2.29, is employed to describe the origin of the parasitic elements in relation to the physical structures for pHEMT in the saturation operation mode. In the model, there are many different circuit elements and they are mainly categorized into two groups, intrinsic and extrinsic elements.

The intrinsic elements, highlighted with red line, are located in the device active channel space underneath the gate. They are identified as bias dependent parameters because of their sensitivity to V_{DS} and V_{GS}. R_{CH} is the channel resistance; R_{DS} is the output resistance; R_{IN} is the input resistance; C_{GS} and C_{GD} are the gate-source and gate-drain capacitances and their sum is equal to the total gate-channel capacitance (C_{GC}). The magnitude of the intrinsic elements is dependent on the material properties of underneath the metal gate and more importantly, the gate length (L_G).

The extrinsic elements are however bias independent. R_G, R_S and R_D are the gate, source and drain resistances; C_{PAR} is the parasitic input capacitance between the gate

and source terminals; C_{DS} is the parasitic output capacitance between the drain and source terminals. Their magnitudes are closely related to the Ohmic and Schottky contact metallisation scheme and the device geometry

2.5.3 RF and Noise characteristics

The RF performance of pHEMT devices is accessed by two important figures of merits (FOM), the cut-off frequency (f_T) and the maximum frequency of oscillation (f_{max}). Their relationship is illustrated in Figure 2.3. Both short-circuit current gain (h_{21}) and unilateral power gain (U) are frequency dependent, but one rolls off at 20dB/decade while the other exhibits two different profiles (10 and 20dB per decade).

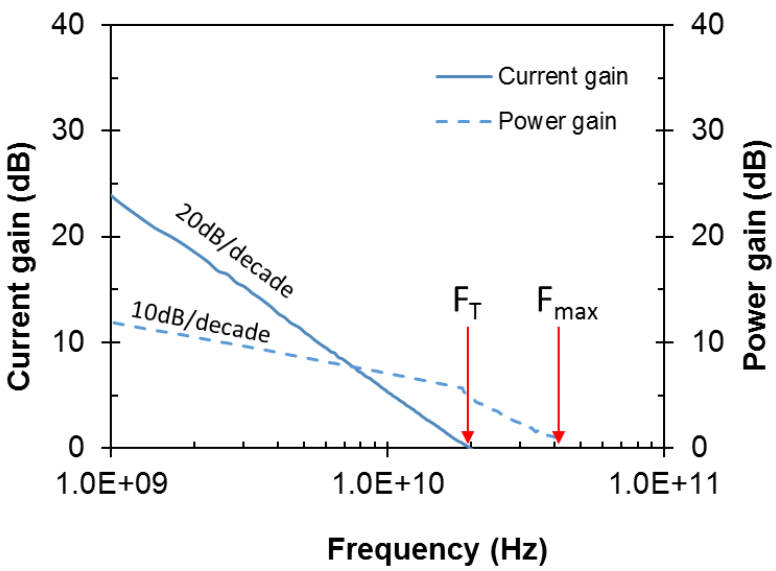


Figure 2.30 frequency responses of a D-mode pHEMT under a fixed V_{DS} and V_{GS}

The cut-off frequency, also called the unity current gain frequency, is defined as how fast an electron transit time within a distance (L_G). In other words, it is the maximum frequency at which the device can operate with a current gain factor of "1", in which case the input gate current (I_{GS}) is equal to the output drain current (I_{DS}) in the small-signal equivalent circuit. Equation 2.35 describes the relationship between f_T, G_m and other parasitic elements:

$$f_T = \frac{G_m}{2\pi((C_{GD} + C_{GS})\left(1 + \frac{R_D + R_S}{R_{DS}}\right) + C_{GD}G_m(R_D + R_S) + C_{PAR})}$$ **Equation 2.35 [27]**

For a conventional device with relatively large L_G, the magnitude of the parasitic capacitances C_{GD}, C_{PAR} and resistances R_D, R_S and R_{DS} are insignificant. The equation can be further simplified as the following:

$$f_T = \frac{G_m}{2\pi C_{GS}} = \frac{v_S}{2\pi L_G}$$ **Equation 2.36 [27]**

Where v_S = electron drift velocity and it will thus approach the saturation velocity (v_{SAT}) as L_G becomes shorter. Clearly, f_T is dependent on L_G and this is the primary reason why the shrinkage of L_G is so important for high performance microwave devices.

The maximum frequency of oscillation, also known as the power gain frequency, is defined as the point in the frequency spectrum when the unilateral power gain (U) = 0 dB ("1"). It is given by Equation 2.37 and it can be rewritten as Equation 2.38 for the same reasons. Unlike f_T, the parasitic elements R_G and C_{GD} must be taken into account for the optimisation of f_{max}.

$$f_{max} = \frac{f_T}{2\sqrt{\frac{R_G + R_{CH} + R_S}{R_{DS}} + f_T(2\pi R_G C_{GD})}}$$ **Equation 2.37 [27]**

$$f_{max} = \frac{f_T}{\sqrt{8\pi R_G C_{GD}}}$$ **Equation 2.38 [27]**

The two FOMs define the upper limit of the device maximum operating frequency (f_{op}). These are however indeed applications dependent. As a general rule, f_{max} is preferred over f_T for higher power gain and high power applications, such as Power Amplifiers (PA). For noise sensitive applications, such as LNA, a higher priority will be given to f_T for better noise performance [18].

The noise performance of pHEMT can be summarised in two FOMs, Minimum Noise Figure (NF_{min}) and Noise Figure (NF). NF_{min} describes the minimum level of noise achievable from a given device structure. It can be calculated by the Fukui's noise expression, as shown in Equation 2.39 [34]. On the other hand, NF indicates the actual noise level on a given device under a set of operating parameters, such as input matching conditions, bias conditions and operating frequency. The relationship between NF_{min} and NF is shown in Equation 2.40 [35, 36]. It correlates NF with NFmin under the influence of source power reflection. The magnitude of NF depends on the magnitude of NF_{min} and also the degree of impedance matching at the inputs.

$$NF_{min} = 1 + \mathrm{k}_1 \frac{f}{f_T}\sqrt{G_m(R_S + R_G)}$$ **Equation 2.39 [34]**

$$\mathrm{NF} = NF_{min} + \frac{4R_n}{Z_O} * \frac{|\Gamma_S - \Gamma_{opt}|^2}{|1 + \Gamma_{opt}|^2 * (1 - |\Gamma_S|^2}$$ **Equation 2.40 [35,36]**

$$\mathrm{R}_n = \frac{k_2}{Gm^2}$$ **Equation 2.41 [35]**

Where f = the operating frequency, k_1, k_2 = the Fukui fitting parameters (k_1 = 0.016, k_2 = 0.030) [34], R_n = noise resistance, Z_O = the system impedance (50 Ω), Γ_S = complex source reflection coefficient and Γ_{opt} = complex optimum reflection coefficient.

From the device point of view, there are two ways to reduce the NF. Firstly, the noise resistance (R_n) is dependent of Gm; the enhancement of Gm could greatly decrease the amount of reflection, making the NF come closer to NF_{min}. Secondly, the improvement of f_T will result in lower NF_{min}, which in effect pulls down the base level of the NF. By referring to Equation 2.34 and Equation 2.36, it is obvious that both Gm and f_T are inversely proportional to the transistor gate length (L_G), which emphasises the importance of gate dimension in low noise applications.

2.6 Experimental techniques and procedures

2.6.1 Measuring instruments

This section describes the measuring instruments that were utilised in this thesis. There were two sets of system available and each of them was built for different purposes. The first system is called System A, as illustrated in Figure 2.31, is proposed for the process monitoring during device fabrication. It consists of mainly five key components. The core part is the Agilent B1500A Parameter Analyser [108], which can be configured as a fixed current source or a gate bias source to the Karl Suss PM5 Cascade Prober [109] via the two Source Measurement Units (SMU1 and SMU2). The B1500A is connected to a control PC via a General Purpose Bus Interface (GBIB). The data acquisition and the SMUs configuration are performed though a software called Integrated Circuit Characterization and Analysis Program (ICCAP) 2010 from Agilent [105]. The PM5 prober has four probe arms in total, which is equipped with 2 µm probe tip and acts as a connection in between the Device Under Test (DUT), the B1500A and the digital mutli-meter.

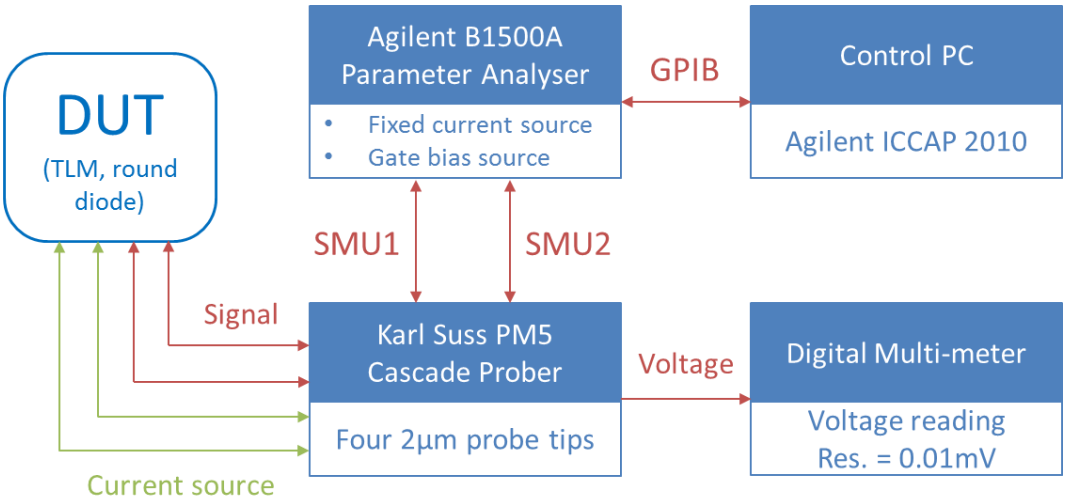


Figure 2.31 System A diagram for TLM and round diode measurement.

This system is very flexible as it can be configured for different applications, such as Transmission Line Measurement (TLM), Schottky diode and others, etc. For the

Schottky diode measurement, only two probe arms are required to supply voltage towards the DUT and measure the output current response. For TLM, the two SMUs will be configured as a fixed current supply towards the DUT while the voltage output is sensed by the digital multi-meter via the other two probes.

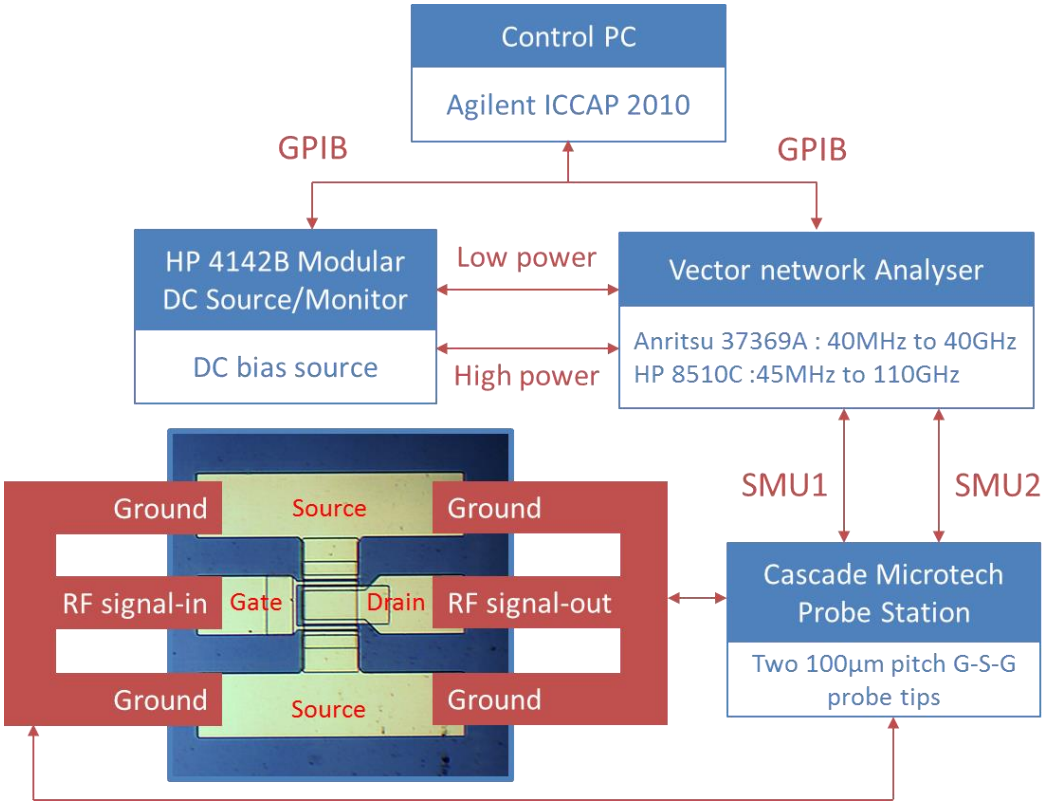


Figure 2.32 System B diagram for on-wafer DC and RF measurements.

System B, as shown in Figure 2.32, is built for on-wafer DC and RF characterisation after the device fabrication stage. There are five key building blocks in the system. The Vector Network Analyser (VNA) is most important bit as it is responsible for the S-parameter measurement. There are two VNA kits available for the device measurements with different frequency range requirements. The Antritsu 37369A machine [110] is designed for the S-parameters measurement with a frequency range from 40MHz to 40GHz while the HP 8510C machine [111] is capable to work up to 110GHz. The DC bias signal is provided by a HP 4142B Modular DC Source/Monitor [112], which is connected to the VNA by an internal bridge network. Both of them are controlled by a PC via the GPIB ports. The configuration of the two SMUs and the

data collection are performed by the ICCAP 2010 software package.

A Cascade Microtech Prober is connected to the VNA kit via the two SMUs. It consists of two probe arms. Each of them equips with a 3-fingers probe tip in the configuration of Ground-Signal-Ground (G-S-G) with a pitch separation of 100 µm. The inset diagram on the bottom left corner demonstrates the arrangements of the two probe tips on the fabricated transistor. The two source pads are grounded with the Ground fingers of the two probe tips. RF signals are sent and received through the gate and drain pads via the two Signal fingers.

2.6.2 Measuring procedures

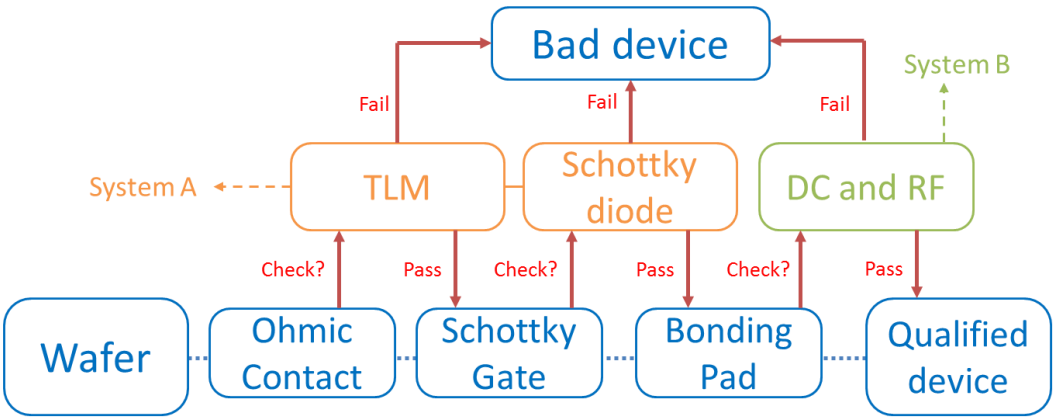


Figure 2.33 Work flow for the processing and characterisation of the pHEMT.

Figure 2.33 illustrates the relationship between the fabrication process flow and the different stages of device characterisation for the purpose of process monitoring. The layout of corresponding process monitoring block on the 15x15mm wafer is also shown in Figure 2.34. The quality of Ohmic and Schottky contacts are accessed by performing the on-wafer TLM and Schottky diode measurements, and they are detailed in section 2.6.4 and 2.6.5 respectively. They are conducted by employing the aforementioned System A on the corresponding TLM ladder and round diode structures. The related parameters such as the contact resistance (R_C), the 2DEG sheet resistance (R_{SH}), the Schottky barrier height (Φ_B) and the ideality factor (n) can be extracted from those measurements, which helps to detect any process related issue.

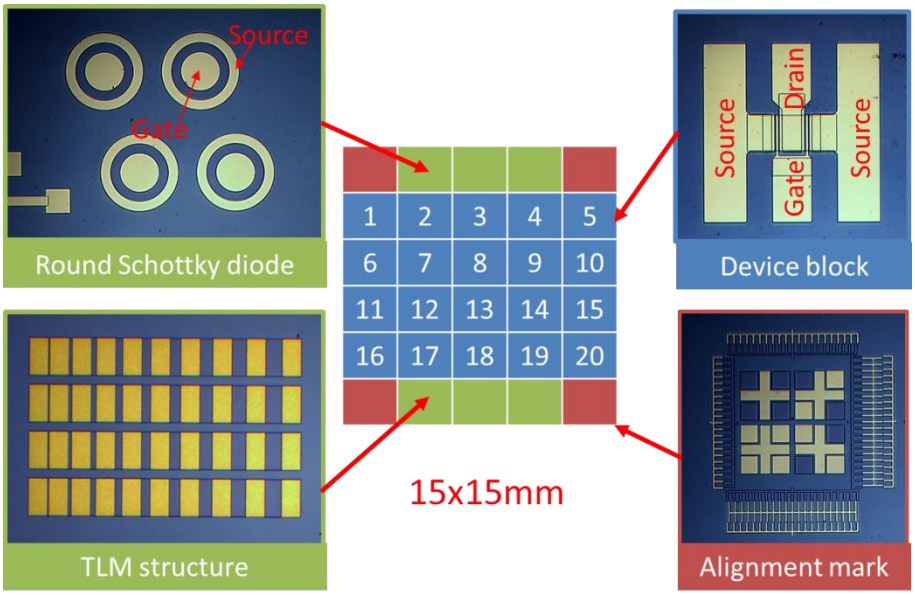


Figure 2.34 Mask layout of different component blocks on a standard diced sample.

On the other hand, the on-wafer transistor measurements are employed to access the DC and RF performances of the fabricated device. Depending on the process complexity, there are typically 80 to 90% of qualified devices on a 15x15mm sample. The results presented in the following chapters are the average of the measured values, with a typical 10% tolerance, which is essential for meaningful device modelling and simulation. Measurements are conducted by employing the System B. The transistor's IV-characteristics are investigated by applying different bias conditions at the gate and drain terminals during the DC measurement, which helps to reveal some key parameters such as the on and off-state gate leakage (I_{GS}), the threshold voltage (V_{TH}) and the maximum transconductance (Gm), etc.

The RF measurements, also called S-parameter measurements, are carried out on the same system by measuring the device frequency responses via a two-port network, which is detailed in section 2.6.5. The correct device bias points are chosen from the DC data accordingly. However, calibration must be performed to remove the errors associated with the VNA, RF cables and probe tips. In this work, a standard Line-Reflect-Reflect-Match (LRRM) calibration technique is employed by using a calibration substrate and the WinCal software from Cascade Microtech [106 - 107].

2.6.3 Transmission Line Measurement (TLM)

A common technique called Transmission Line Measurement (TLM) was employed for the characterisation of Ohmic contacts during the device fabrication, which was developed by Berger [99] and Murrman and Widmann [100] in 1969. The principle of TLM is to extract the resistance elements (R_A and R_B) between two metal pads sitting on an isolated MESA structure, as shown in Figure 2.35. Each metal pad has a thickness of t and a width of W. R_A is the access resistance of the point contact located at a distance of L_T from the edge of the metal pad. R_B is the access resistance between the two metal pads separating by a distance of d.

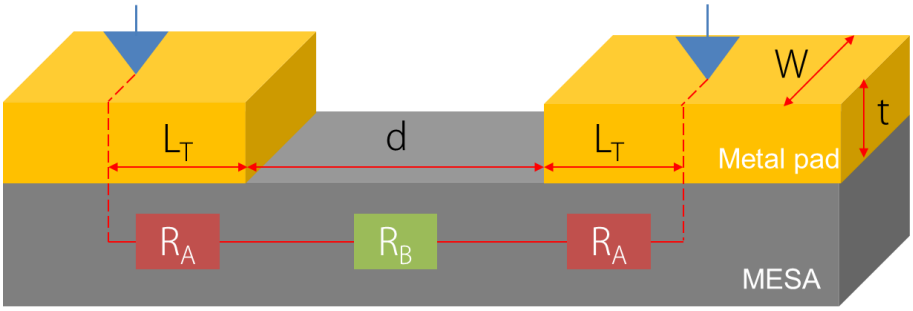


Figure 2.35 Schematic diagram of simple TLM structure.

As a first principle, the resistance of a transmission line is determined by several parameters, including material resistivity (ρ), length of the transmission line (L), cross-sectional area of the transmission line (A) and also the sheet resistance (R_{SH}). Their relationships are shown in Equation 2.42.

$$R = \rho \frac{L}{A} = \rho \frac{L}{Wt} = R_{SH} \frac{L}{W}$$ **Equation 2.42 [27]**

The total resistance (R_T) between the two pads is essentially a sum of the resistance elements (R_A and R_B). By substituting Equation 2.42, this can be further expanded into Equation 2.43. R_{SHA} is the sheet resistance of the region under the metal pad. R_{SHB} is the sheet resistance of the semiconductor region between the pads.

$$R_T = 2R_A + R_B = 2R_{SHA} \frac{L_T}{W} + R_{SHB} \frac{d}{W}$$ **Equation 2.43 [27]**

By assuming $R_{SHA} = R_{SHB}$ [101], the equation of R_T can be simplified in term of specific contact resistance (R_C) and semiconductor sheet resistance (R_{SH}), as shown in Equation 2.44:

$$R_T = 2R_C + R_{SH}\frac{d}{W}$$ **Equation 2.44 [27]**

In practice, the extraction of R_C (in Ω.mm) and R_{SH} (Ω/□) involves the use of a TLM ladder structure as shown in Figure 2.36. It consists of a series of metal pads separating in distance (d_1, d_2, d_3 to d_i). A current source with 1mA supply is attached at the first and last metal pads. The voltage drop (V_1 to V_i) across between the neighbouring pads is recorded by a voltage meter, which results in the corresponding resistance (R_i).

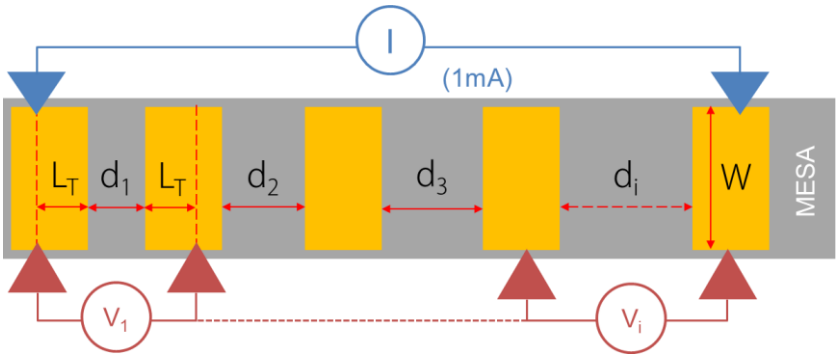


Figure 2.36 Schematic diagram of the TLM ladder configuration.

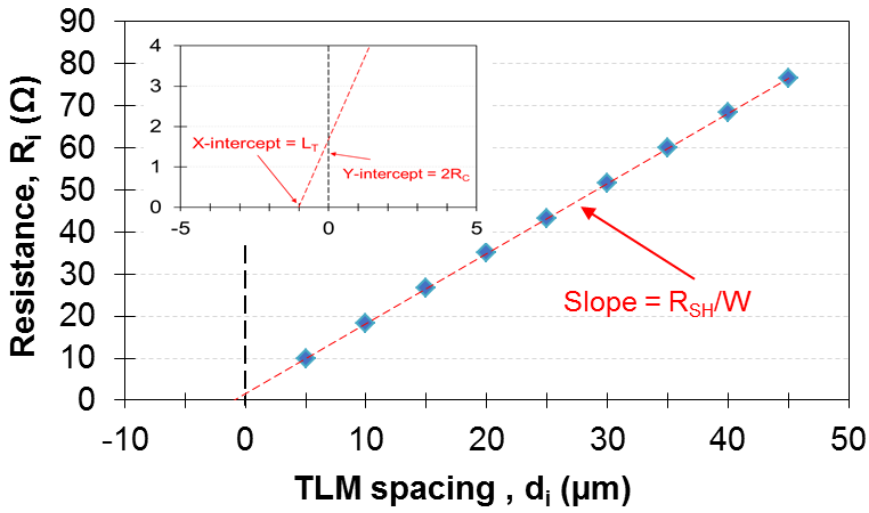


Figure 2.37 Typical linear TLM plot for a TLM structure with pad spacing from 5 to 45 μm.

The relationship between R_i and d_i can be represented in a linear equation as shown in Equation 2.45 and its graphical representation is depicted in Figure 2.37. The TLM plot is generated based on a set of resistances (R_i) with a series of pad spacing (d_i = 5 to 45 µm). The R_{SH} can be extracted from the slope of the plot. The X and Y-intercept of the plot represents the terms L_T and $2R_C$.

$$R_i = \frac{R_{SH}}{W} d_i + 2R_C$$ **Equation 2.45 [27]**

2.6.4 Extraction of Schottky barrier height and ideality factor

The Schottky diode measurement is normally carried out after the formation of Schottky contact, which is vital for monitoring of the gate quality. In this process, a gate voltage (V) is applied across the gate and source terminals and the change of gate current (I) is recorded. Depending on the polarity of gate voltage, the Schottky diode behaviour can be divided into forward (V>0) and reverse (V<0) bias regions. The Schottky barrier height (Φ_B), ideality factor (n) and series resistance (r_s) can be extracted from the log scale plot of the forward region, as shown in Figure 2.38:

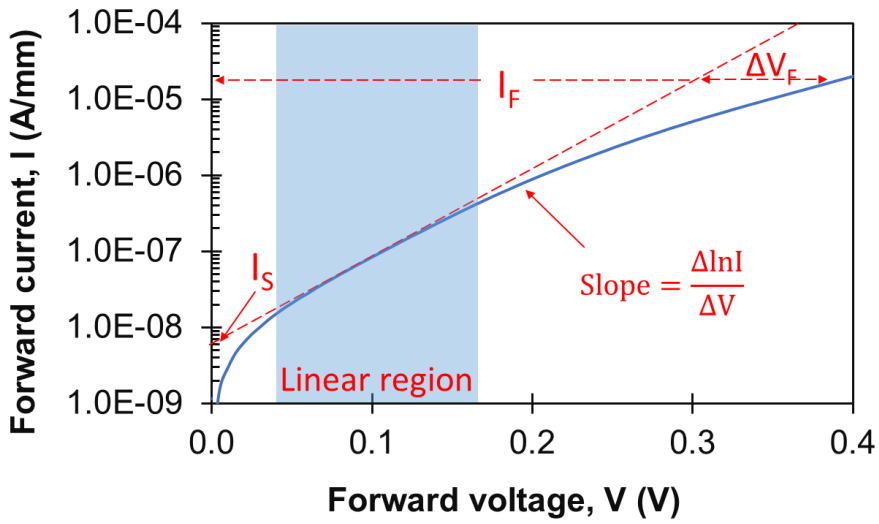


Figure 2.38 Typical Schottky diode forward current characteristic of InP pHEMT.

The technique for the parameters extraction is based on the thermionic current-voltage relationship given by:

$$I = I_S(e^{\frac{qV}{nkT}} - 1)$$ **Equation 2.46 [102]**

$$I_S = AA^*T^2 e^{\frac{-q\Phi_B}{kT}}$$ **Equation 2.47 [102]**

Where I_S = saturation current for V = 0, A = area of the Schottky diode contact, A* = Richardson constant (material dependent, $In_{0.52}Al_{0.48}As$ = 10.26 [12]). When the applied voltage (V) is >3kT/q, the output current (I) will exhibit a linear behaviour in the log plot, which is labelled in Figure 2.38. In the linear region, the output current is given by Equation 2.48 and it can be rearranged into Equation 2.49:

$$I = I_S(e^{\frac{qV}{nkT}})$$ **Equation 2.48 [102]**

$$lnI = \frac{q}{nkT}V + lnI_S$$ **Equation 2.49 [102]**

The slope of the linear region and thus the ideality factor (n) are given by Equation 2.50 and Equation 2.51:

$$slope = \frac{\Delta lnI}{\Delta V} = \frac{q}{nkT}$$ **Equation 2.50 [102]**

$$n = \frac{\mathrm{q}}{\mathrm{kT}} * \frac{1}{slope}$$ **Equation 2.51 [102]**

Also, by rearranging Equation 2.47 into Equation 2.52, Φ_B can be obtained:

$$\Phi_B = \frac{kT}{\mathrm{q}} \ln\left(\frac{AA^*T}{I_S}\right)$$ **Equation 2.52 [102]**

Where I_S = the Y-intercept of the linear extraction on the plot. Finally, the series resistance (r_S), which contributes to the saturation of the output current at higher applied voltage, is governed by Equation 2.53:

$$\mathrm{r}_S = \frac{\Delta V_F}{\mathrm{I}_F}$$ **Equation 2.53 [102]**

Where ΔV_F = the voltage difference between the linear extrapolation line and the saturation region, I_F = the current at the beginning of the saturation region.

2.6.5 Two-port network and S-parameters

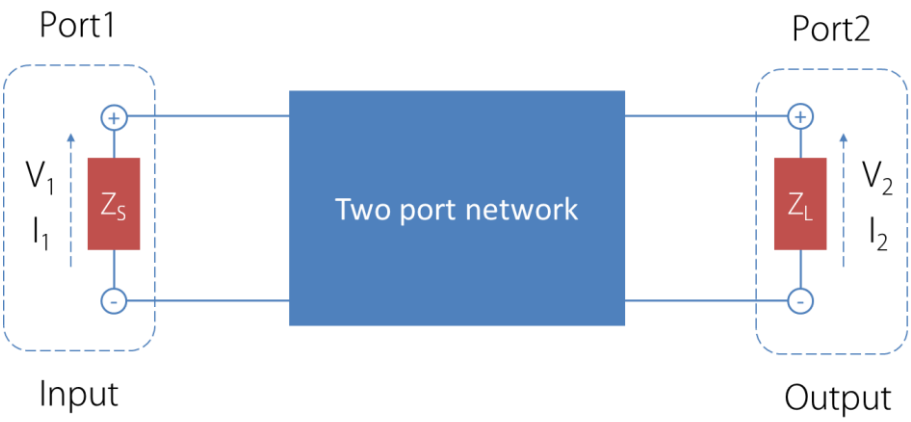


Figure 2.39 Schematic diagram of simple two-port network

A two-port network is a mathematical model which provides an abstract description of a complex electrical network. In the network, the device under test (DUT, a transistor) is treated as a black box, in which the responses of the applied signal can be resolved easily without worrying about the complicated internal circuit elements. Figure 2.39 illustrates a simple diagram of the two-port network model. V_1, V_2 are the input and output voltage at port 1 and port 2. I_1 and I_2 are the corresponding the input and out current. Z_S and Z_L are the source and load impedance.

There are many different parameters available to describe the performance of the network, such as impedance (Z-parameters), admittance (Y-parameters), hybrid (H-parameters) and scattering (S-parameters), etc. The S-parameters are preferred to describe the microwave characterisation as they are relatively easier to measure at high frequency although they are inter-convertible with each other. With the S-parameters, the network is modelled as a set of incident (a_n) and reflected (b_n) signal waves travelling across the DUT, where n = port number. The relationship between a_n, b_n and the corresponding port voltage (V_n) and current (I_n) is given by:

$$a_n = \frac{V_n + Z_0 I_n}{2\sqrt{Z_0}} = \frac{\text{Incident voltage wave on port n}}{\sqrt{Z_0}}$$ **Equation 2.54 [103]**

$$b_n = \frac{V_n + Z_0 I_n}{2\sqrt{Z_0}} = \frac{\text{Reflected } voltage\ wave\ on\ port\ n}{\sqrt{Z_0}}$$ **Equation 2.55 [103]**

Where Z_0 = reference impedance (50Ω normally).

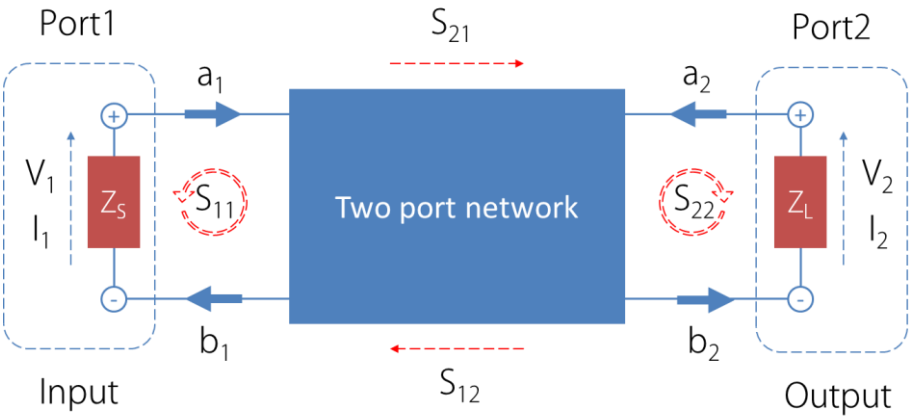


Figure 2.40 Schematic diagram of two-port network in S-parameter representation.

The relationship between the S-parameters and the travelling waves can be expressed in a matrix form [103] and their relationship is illustrated in Figure 2.40:

$$\begin{bmatrix} b_1 \\ b_2 \end{bmatrix} = \begin{bmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \end{bmatrix}$$ **Equation 2.56 [103]**

Where S_{11} = reflected signal at the input port, S_{12} = reverse transmission gain, S_{21} = forward transmission gain and S_{22} = reflected signal at the output port, at a matched condition as the following:

When Z_L=Z_0, a_2=0, $$S_{11} = \frac{b_1}{a_1}$$ **Equation 2.57 [103]**

When Z_S=Z_0, a_1=0, $$S_{12} = \frac{b_1}{a_2}$$ **Equation 2.58 [103]**

When Z_L=Z_0, a_2=0, $$S_{21} = \frac{b_2}{a_1}$$ **Equation 2.59 [103]**

When $Z_S=Z_0$, $a_2=0$, $$S_{22} = \frac{b_2}{a_2}$$ **Equation 2.60 [103]**

In practice, a series of device microwave performance parameters, such as maximum available gain (MAG), stability factor (K) and short-circuited current gain (h_{21}) can be calculated from the S-parameters as shown below:

$$\mathrm{MAG} = \left|\frac{S_{21}}{S_{12}}\right| K - \sqrt{K^2 - 1}$$ **Equation 2.61 [12]**

$$K = \frac{1 + |S_{11}S_{22} - S_{12}S_{21}|^2 - |S_{11}|^2 - |S_{22}|^2}{2|S_{12}S_{21}|}$$ **Equation 2.62 [12]**

$$|h_{21}| = |\frac{-S_{21}}{(1 - S_{11})(1 + S_{22}) + S_{12}S_{21}}|$$ **Equation 2.63 [12]**

MAG defines the power gain obtained at input and output impedance matched condition, which is equal to the unilateral power gain (U) given that there is no feedback network used. K defines the boundary of device's stability, where a device with K>1 is regarded as unconditionally stable. This is why the power gain in Figure 2.30 shows two different rolling off profiles. Recalled the two FOMs f_T and $f_{max,}$ they are the corresponding cut-off frequency at the unity of MAG and h_{21} respectively.

In addition to those parameters, the S-parameters can be converted into Z, Y and H parameters for the extraction of intrinsic and extrinsic small signal parameters mentioned previously, which is the key for the design and modelling of microwave circuitry, but their detail is not within the scope of this work.

3 STUDY OF SUBMICRON GATE STRAINED InGaAs-InAlAs pHEMT using ELECTRON-BEAM LITHOGRAPHY

3.1 Introduction: The importance of shorter gate length

In the last decade, the University of Manchester has made significant progress in the design and development of high performance LNA for the SKA project. This is largely attributed to the works that have been done by the group and also the excellent material growth facility at Manchester [12]-[16]. By incorporating the power of high quality epitaxial structure, the optimised fabrication process and advanced device features, excellent InP-based pHEMT devices with low noise characteristic and also high breakdown voltage have been demonstrated [37].

As discussed in section 2.5, the transistor DC and RF performance, in particular f_T, f_{max} and NF_{min}, show a strong dependence on its gate length (L_G). Figure 3.1 depicts the relationship between f_T and L_G for common III-V and Si transistors.

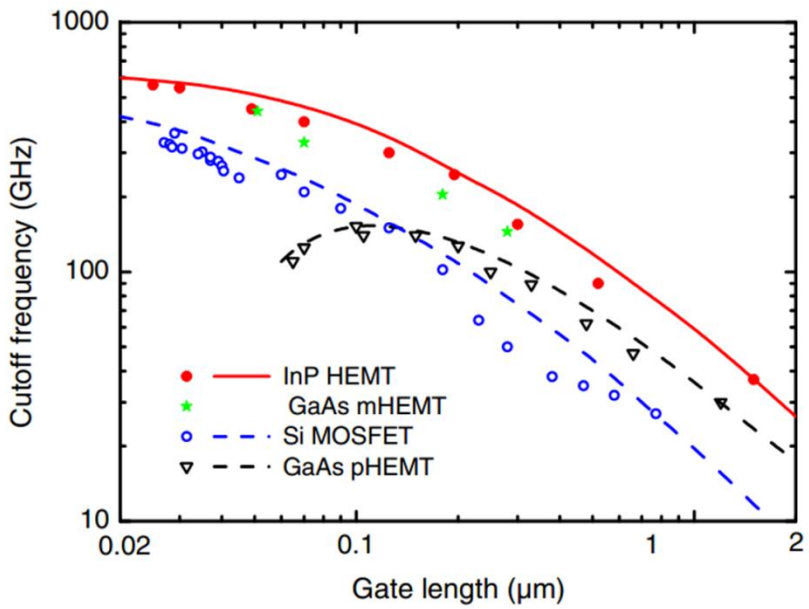


Figure 3.1 Reported cut-off frequency vs gate length for common III-V and Si transistors [23]

The shrinkage of gate length has led to a tremendous increase on the magnitude of f_T and this is more pronounced for InP-based transistors. So far, all work at Manchester has been conducted on InP-based pHEMT using a conventional 1 µm gate process. The restriction of large gate dimension has become a major issue towards the future development of our LNA designs. The implementation of submicron gate pHEMT is now becoming mandatory to further improve the LNAs in meeting current SKA phase requirement and subsequent next generation LNA for phase 2 and 3 of SKA.

Therefore, the objective of this study is to explore the potential of submicron gate pHEMT for future LNA applications for SKA. There are two parts to this study. The first part begins with a comparison between the conventional 1 µm and the 0.1 µm gate devices on conventional InGaAs-InAlAs epitaxial structure, fabricated by optical and E-beam lithography respectively. Part two will elaborate more on the combination of the submicron gate together with some modified device geometries and the optimised material growth, which results in a drastic improvement compared with the 1 µm gate device.

3.2 Principle of common lithography techniques

Lithography is one of most important steps in microelectronic fabrication, where desired pattern can be printed onto the semiconductor substrate by means of an electrochemically sensitive polymer. It can be performed with or without a mask depending on the lithography techniques. For simplicity, the discussion will start with the most popular option, namely photolithography or optical lithography, as illustrated in Figure 3.2.

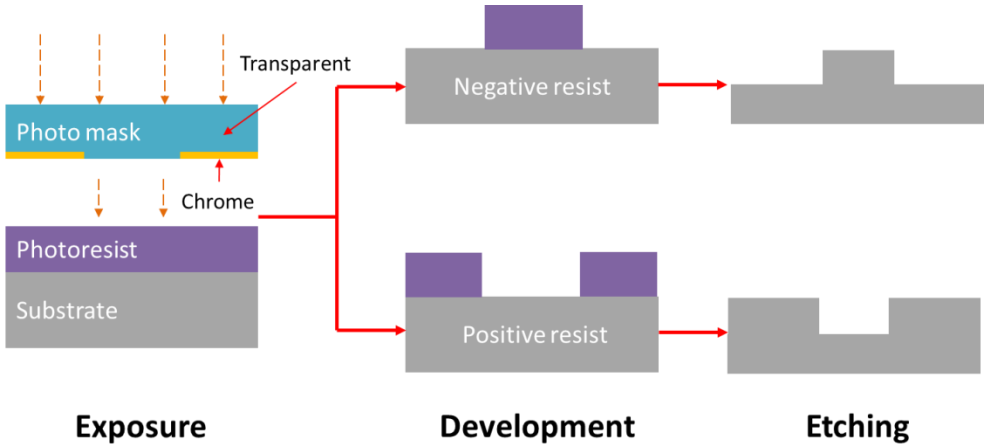


Figure 3.2 General work flow of photolithography

This is a pattern transfer process in three steps, namely exposure, development and etching. In the first step, a photosensitive polymer called photoresist or resist is coated onto the semiconductor substrate. A transparent photomask with predefined pattern in chrome coating, sits on top of the resist. The pattern is transferred by exposing the unmasked area to a photo light source. Chemical reaction then takes place in the exposed area. In step two, the exposed sample is developed in a chemical solution called a developer. The photoresist region under exposure can also be hardened or softened depending on the exact type of photoresist. Finally, the pattern is transferred from the photoresist to the substrate by means of chemical or plasma etching.

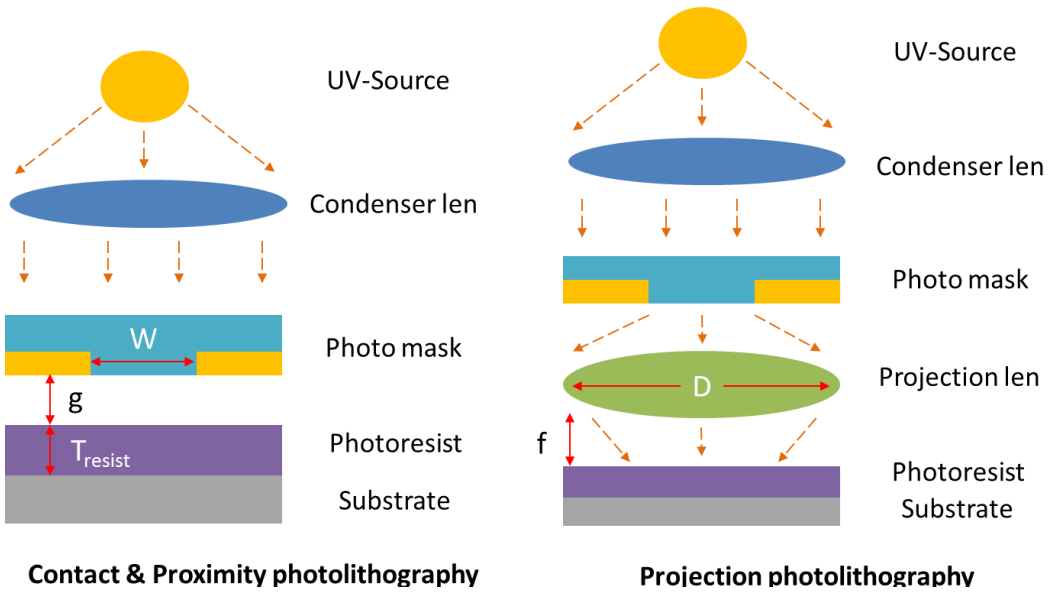


Figure 3.3 Comparison between contact, proximity and projection systems

Photolithography can be implemented in three different systems and their details are depicted in Figure 3.3. As shown on the left hand side, the concept of contact and proximity photolithography is very similar as they share the identical physical features including UV-source, condenser lens and photomask. The unique difference between the two is the gap (g) between the photomask and the photoresist coating. In the proximity system, the mask is separated from the photoresist layer by a small distance (typically g = few µm). On the contrary, for contact lithography, the mask and the photoresist are brought into direct contact, leaving no gap (g = 0) between them. The key advantage of a contact system is its capability to produce very small feature. This can be understood via Equation 3.1 [38].

$$\text{Assume } \lambda < g < \frac{w^2}{\lambda}, \qquad W_{min} \approx \sqrt{k_3 \lambda g} \qquad \textbf{Equation 3.1}$$

Where W_{min} = minimum width of the feature, k_3 = process dependent constant, which relates to the properties of photoresist and light source and λ = wavelength of the light source. From the equation, it can be concluded that the minimum resolution of the system is constrained by g and this is why contact photolithography is more popular. However, the improved resolution is achieved with a cost of damaging the substrate and the photomask, due to the high contact pressure at the interface.

An alternative approach called projection photolithography, shown on the right hand side of Figure 3.3, is capable of resolving the aforementioned issue. In the projection system, a projection lens is inserted in between the mask and the substrate. It acts as an image reducer, which reduces the incoming pattern from the photomask by a certain scale (typically 4 to 5x) and projects it onto the photoresist layer. For a given light source, its resolution is significantly improved from the contact based system. The minimum feature size of this system is governed by Equation 3.2 [38].

$$\text{Assume } W^2 < \lambda\sqrt{g^2 + \frac{D^2}{2}}, \qquad W_{min} = \frac{1.22f\lambda}{D} \approx k_4 \frac{\lambda}{NA} \qquad \textbf{Equation 3.2}$$

Where f = focal distance between the projection lens and the photoresist, D = diameter of the projection lens, k_4 = process dependent constant and NA = numerical aperture of the projection lens. For a given light source, the resolution of the system is inversely proportional to the magnitude of NA, which is also dependent on the terms f and D. Unlike the other two systems, the projected image can only cover a finite region on the substrate. This is why projection photolithography is often utilised in a system called a “Stepper”, where the substrate is shifted from the exposed location to another one with the help of high precision computing control. The cost of projection photolithography is therefore extremely high, which hinders its application in research and academic projects.

By referring to the equations of the above-mentioned systems, it is clear that λ is the key determinant of the system resolution. Table 3.1 highlights the relationship between λ and resolution:

Source	Spectrum	Wavelength (nm)
Hg lamp (g-line)	UV	436
Hg lamp (h-line)		405
Hg lamp (i-line)		365
Excimer laser (KrF)	DUV	248
Excimer laser (ArF)		193

Table 3.1 Relationship between the source wavelength and the resolution [39]

The UV-based light sources (g-line and i-line) are economical option with moderate resolution, making them very popular for research and development work. Excimer laser sources can push the resolution into the sub-100nm scale. However, they are restricted in industrial applications due to the extremely high cost of ownership.

To resolve any feature size below the capability of conventional i-line lithography, a rather economical option called electron-beam lithography (E-beam) is widely employed for the purpose of research works. As illustrated in Figure 3.4, it relies on the interaction between the electron sensitive photoresist and the electron beam projection from the electron gun source on the top.

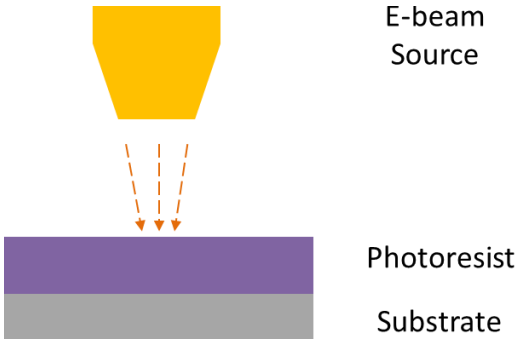


Figure 3.4 Principle of E-beam lithography system

The advantage of E-beam lithography is that it requires no mask for pattern generation. Pattern can be generated from the movement of an E-beam point source, where the resolution is controlled by the voltage potential of the source electron, as shown in Equation 3.3 [40].

$$\lambda = \frac{h}{\sqrt{2me\Delta V}} \qquad \textbf{Equation 3.3}$$

Where h = Planck constant, m = mass of electron and ΔV = voltage potential of the electron gun. The major advantage of E-beam lithography over the optical one is that it requires no mask but remains capable of resolving nanoscale images, though the production speed is very slow owing to the fact that it can only write a single line at a time. In the next sections, its application will be demonstrated through the fabrication of submicron gate pHEMT devices.

3.3 Baseline process for pHEMT fabrication

The fabrication of pHEMT is a device manufacturing process which requires state of the art facilities and carefully handling. It is a link between the material growth and the device characterisation stages that implies a direct impact on the device performance. At the University of Manchester, epitaxial growth on the GaAs or InP substrate wafers is done by an in-house solid source MBE machines, RIBER V90H and V100H. The completed wafers are checked by employing Hall Effect and C-V measurements before they are qualified for device fabrication. The wafers are then diced into blocks of 15x15mm^2 for different studies. The fabrication process is conducted with the economical i-line photolithography in a class-1000 cleanroom environment. A summary of the baseline process for the 1 µm planar gate pHEMT fabrication is depicted in Figure 3.5. For further details, refer to the full process flow attached in APPENDIX A1.

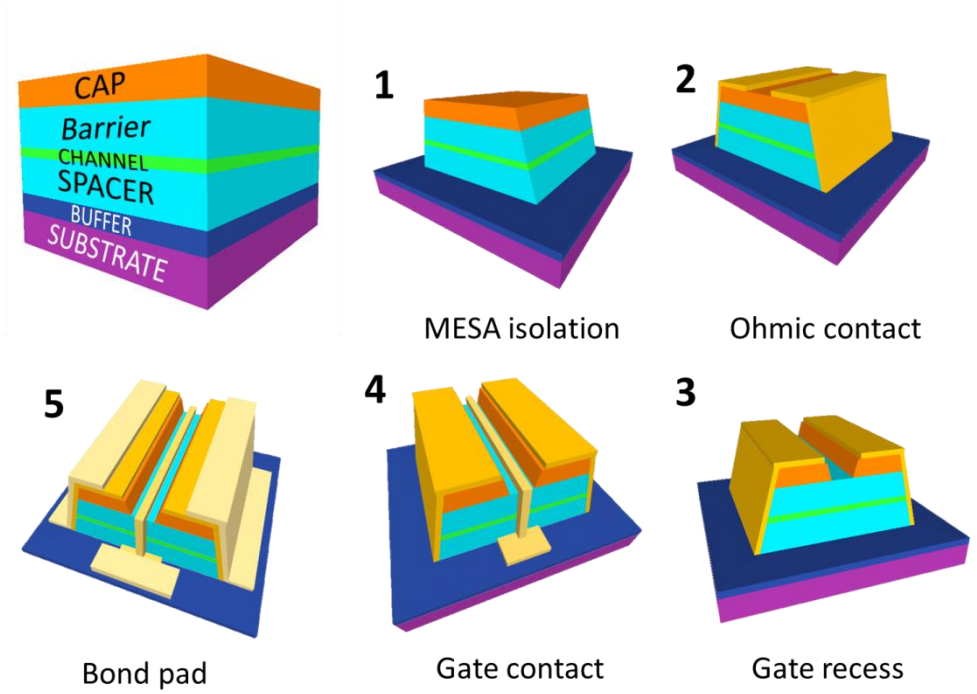


Figure 3.5 Illustration of 1 µm planar gate pHEMT baseline process steps for one or two fingers devices

The process contains four major steps, MESA, Ohmic, gate and bond pads. The sample is cleaned using Trichloroethylene, Acetone and Isopropyl Alcohol (IPA) solvents before and after each step to ensure it is free from particles during the transport or any other forms of contamination especially hydrocarbons. Firstly, the active device region is defined by the MESA isolation step. The MESA area is patterned using S1805 positive photoresist followed by a chemical wet etch using phosphoric solution in the ratio of H_3PO_4:H_2O:H_2O_2 = 3:1:50. The depth of the MESA region is typically 160nm. To prevent the potential leakage path through the side-wall of the InGaAs channel, a further side-wall etching step is performed by using a selective Succinic etchant which also improves the thermal stability of the Schottky gate during succeeding heat treatment [65]. This etchant provides a very high selectivity between InGaAs and InAlAs (250:1), which was improved by a past co-worker Dr. M. M. Isa [15].

Next, the drain and source regions are patterned using a MicroChem 2μm grade negative photoresist, which has an undercut profile after development. Alloyed Ohmic contacts using AuGe/Au = 50/400nm are formed by thermal evaporation followed by life-off in hot N-Methyl-2-pyrrolidone (NMP) solvent. The contact resistance is then monitored and used as a check for the integrity of the contacts.

Thirdly, MicroChem 1μm grade Aznlof negative photoresist with an undercut profile is employed to define the 1μm gate recess opening. The removal of the cap layer in the gate recess region is done by the aforementioned highly selective Succinic etchant to ensure a high quality Schottky interface. A gate metal consisting of Ti/Au = 50/450nm is deposited by thermal evaporation, followed by lift-off in hot NMP to create the 1μm planar metal gate on the InAlAs layer.

Finally, bonding pad metallisation with Ti/Au = 50/450nm is deposited on top of the original drain, source and gate region, by employing a similar process as in the Ohmic step. The function of the bond pads is to enhance the contact between the probers of measuring instrument and the hardened metal contacts due to the heat treatment process.

Devices after the bond pad stage are essentially ready for characterisation. But for the low noise or high gain applications, multi-fingers device configuration (number of gate fingers > 2) is compulsory to suppress the noise resistance and maintain a high power gain. The connection between separated source contacts is achieved by employing a metal bridge structure consisting of Au = 500nm with SF11 dielectric bridge support developed earlier. Figure 3.6 illustrates the fabricated pHEMT devices with four and six gate fingers.

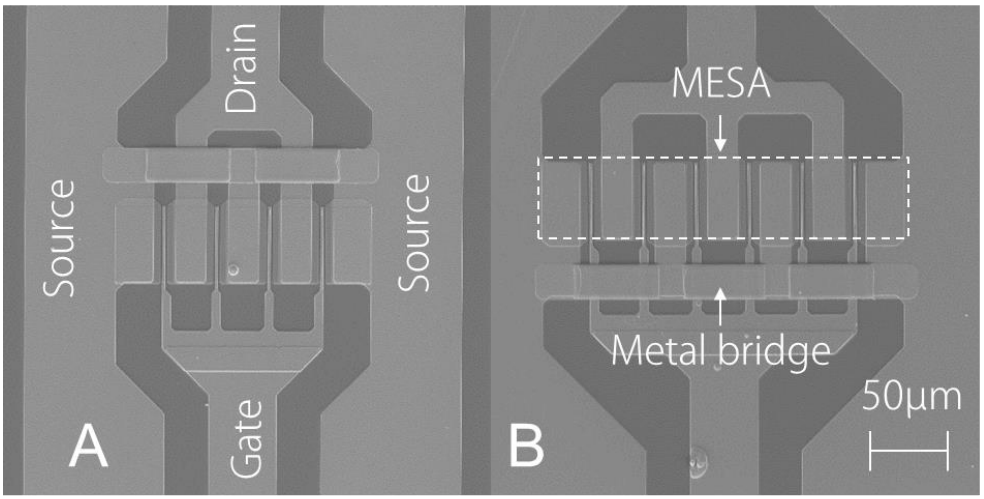


Figure 3.6 Digital images of fabricated pHEMT devices with four (A) and six (B) gate fingers

3.4 Fabrication of conventional 1µm and 0.1µm planar gate pHEMT

3.4.1 Introduction

In this study, two different sets of pHEMT device with planar gate structures of 1µm and 0.1µm are fabricated, by means of conventional i-line photography and E-beam lithography respectively. Its purpose is to demonstrate the benefits and challenges on the fabrication of submicron gate device, by combining the power of our in-house growth material and pHEMT process flow with the advanced E-beam lithography in a modified SEM system in our Lab at Manchester.

3.4.2 Material growth

The epilayer employed in the study is sample XMBE#171, where the prefix XMBE indicates that the material growth was done in the RIBER V100H system. It includes a conventional InGaAs/InAlAs/InP structure for the fabrication of 1µm and 120nm planar gate pHEMT in this study. The structure profile is shown in Table 3.2:

Epi-structure data - XMBE#171		
Epilayer	**Material**	**Thickness (Å)**
Cap	$In_{0.53}Ga_{0.47}As$	50
Barrier	$In_{0.52}Al_{0.48}As$	300
δ-doping (Si-doped)		
Spacer 1	$In_{0.52}Al_{0.48}As$	50
Channel	$In_{0.70}Ga_{0.30}As$	140
Buffer	$In_{0.52}Al_{0.48}As$	4500
Substrate	InP (Fe-doped)	

Table 3.2 Epitaxial structure data of XMBE#171

From the bottom, a thick lattice-matched $In_{0.52}Al_{0.48}As$ is grown on top of the semi insulating InP to smooth down structural defects from the substrate. The strained $In_{0.70}Ga_{0.30}As$ channel is sandwiched between the wide band gap spacer and buffer, leading to the formation of 2DEG. The top δ-doped layer supplies the electron carriers,

where the coulomb scattering between ionised donors and carriers of 2DEG is prevented by the spacer layer. A rather thick (30nm) barrier layer, sitting on top of the δ-doped layer, is utilised for the formation of a Schottky barrier. Finally, a lattice-matched $In_{0.53}Ga_{0.47}As$ is used as the cap layer to protect the underneath Al-rich material from oxidation and also to lower the barrier for the formation of low resistance Ohmic contact.

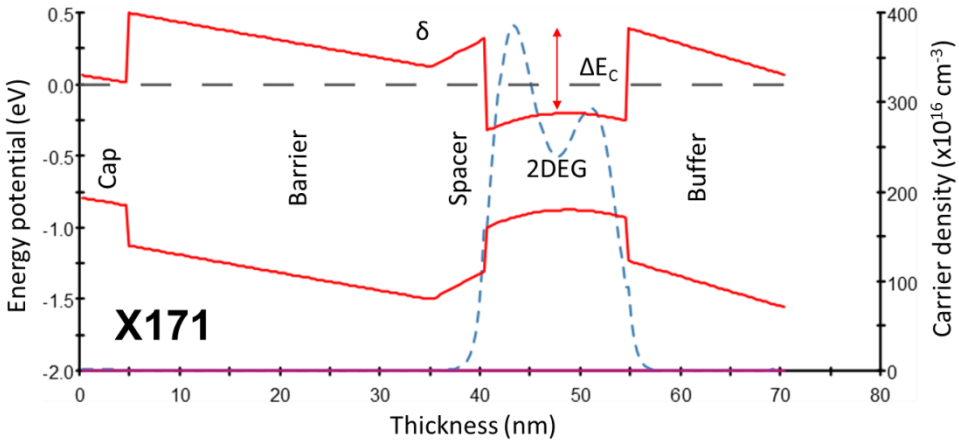


Figure 3.7 Energy band diagram of XMBE#171 structure (V_{GS} = 0, V_{DS} = 0)

Quantum transport simulation is carried out using the WinGreen® simulation package, based upon the epilayer profile stated in Table 3.2. The energy band diagram and carrier distribution of the epitaxial structure under zero bias conditions (V_{GS} = 0 V, V_{DS} = 0 V) is illustrated in Figure 3.7. With the help of a high concentration δ-doping scheme and thin spacer layer, the bottom QW is pulled deeply under the Fermi level (Energy = 0eV), which implies stronger confinement of the electron carrier and thus increases the effective carrier concentration of the 2DEG.

The material was characterised by Hall Effect measurement at room temperature and 77K conditions and the data is shown in Table 3.3:

Hall effect measurement data - XMBE#171		
Sheet carrier concentration (n_H) (x10^{12} cm^{-2})	300K	3.16
	77K	3.56
Hall mobility (μ_H) (cm^2/V.s)	300K	10653
	77K	24649
Sheet resistance (R_{SH}) (Ω)		185

Table 3.3 Hall Effect measurement data of XMBE#171

The thin spacer profile encourages more electrons to be trapped inside the 2DEG region, resulting in extremely high carrier concentration in the channel. However it does lead to degradation in the carrier mobility due to the strong coulomb scattering effect between the δ-doped and channel layers. Device fabrication

For comparison purpose, there were two types of different devices fabricated in the study:

1. Conventional 1 µm planar gate pHEMT by i-line photolithography.
2. Submicron (0.1 µm) planar gate pHEMT by E-beam lithography.

Both transistors produced were based on the same epitaxial structure XMBE#171 and shared similar process steps as stated in section 3.3. Device active area with a depth of 160nm was defined by MESA etching using a phosphoric etchant, followed by a sidewall etching to eliminate the leakage path between the channel sidewall and the Schottky gate metal in the later stage. Alloyed Ohmic contact consisting of AuGe/Au = 50/400nm for the drain and source regions was formed by thermal evaporation with additional heat treatment in N_2 filled furnace.

The formation of Schottky gate contact was considerably different between the 1 µm and the 0.1 µm gate devices. For the 1 µm gate device, the patterning of the gate recess region was done by i-line photography with Aznlof 1 µm grade negative photoresist. As shown in Figure 3.8, a 1 µm gate recess opening with undercut photoresist profile

was created after development. A highly selective Succinic etchant was employed to remove the top 5nm cap layer for the metal gate contact. A Schottky gate metallisation consisting of Ti/Au = 50/450nm was formed by thermal evaporation and lift-off in hot NMP solvent.

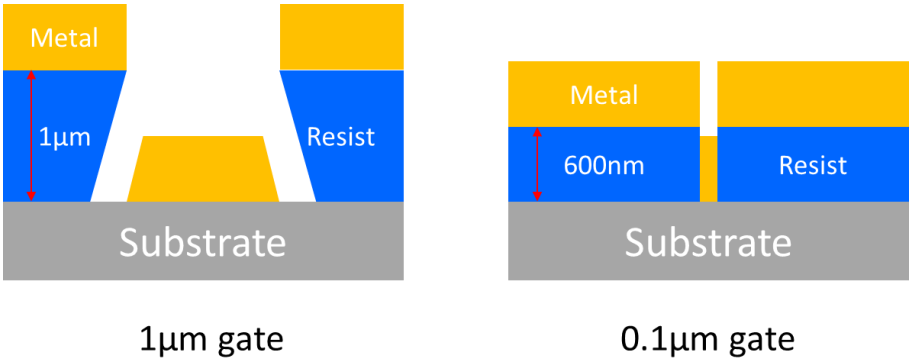


Figure 3.8 Illustration of the photoresist profile between 1 µm and 0.1 µm gate recess opening after development and gate deposition

The fabrication of the submicron gate device was conducted in joint collaboration with Dr. Scott Lewis, from the Photon Science Research Institute (PSI) at the University of Manchester. After completion of the Ohmic contact stage, the sample was sent to the PSI department for the patterning of 0.1 µm gate footprint by E-beam lithography. A Cambridge S360 Scanning Electron Microscope (SEM) and a special high aspect ratio SML600 0.6 µm grade E-beam sensitive photo resist were employed for the gate footprint pattering. After development, the process shows a vertical profile in the gate recess region, as depicted in Figure 3.8.

The removal of the cap layer in the gate recess and the evaporation of Schottky gate metal were carried out in our cleanroom as before. However, the gate deposition step needed a great deal of precaution. Firstly, the total thickness of the gate metallisation must be less than that of the photoresist (600nm), in order to avoid problems in the lift-off process. Therefore, a gate metallisation scheme with Ti/Au = 50/450nm was selected. Secondly, the sample with nanoscale gate footprint has to be located right at top of the metal source during the evaporation process; otherwise, the formation of gate contact may be deformed due to the narrow incident angle of the photoresist

profile, as shown in Figure 3.9.

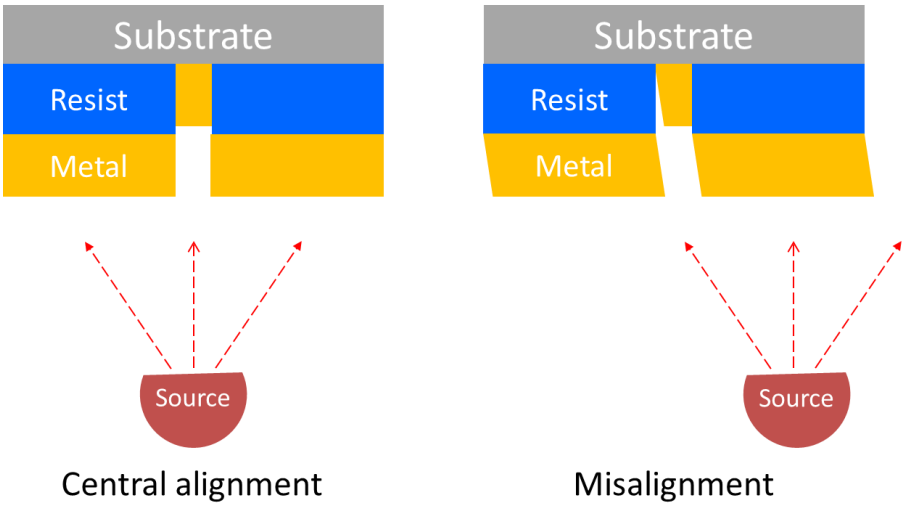


Figure 3.9 Change of metal profile with different sample locations in the evaporator

The lift-off process was carried out with Acetone. The fabrication processes were completed with bond pads metal deposition by thermal evaporation of 50/450nm of Ti/Au. SEM images of devices with 5 µm drain source separation from the two samples are shown in Figure 3.10 for comparison purpose. The size of the gate lengths were 1.1 µm and 0.12 µm as labelled on the diagram.

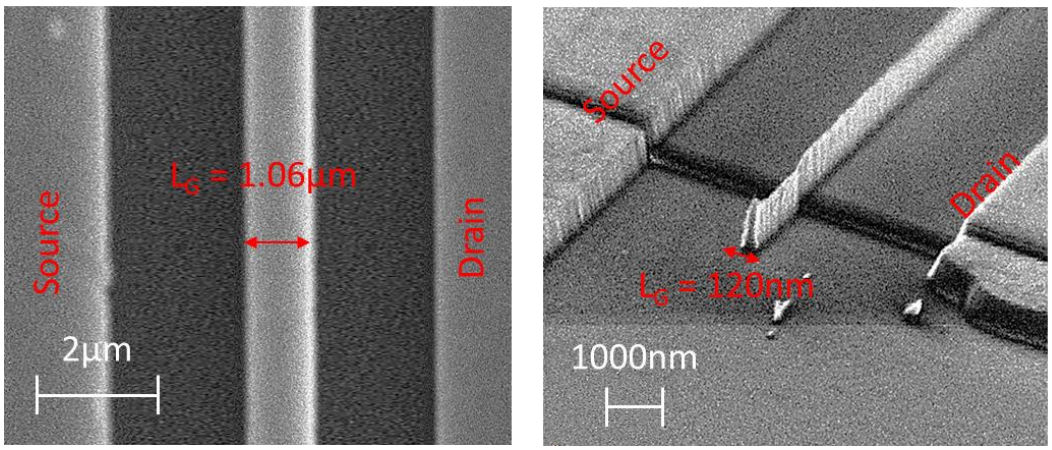


Figure 3.10 SEM images of XMBE#171 1 µm (left) and 0.1 µm (right) planar gate devices

3.4.3 Results and discussions

3.4.3.1 TLM measurement

Upon completion of the Ohmic contact deposition, samples were annealed at 280˚C for 90s in a temperature controlled furnace with N_2 purging gas. Transmission Line Measurements (TLM) were conducted on the predefined TLM structures at room temperature after the heat treatment. The structures consist of a series of 50x100 µm probing pads, with spacing from 5 to 45 µm in between. For details of the TLM structures and the principle of measurement, refer to section 2.6.3.

The Ohmic contact resistance (R_C) and 2DEG sheet resistance (R_{SH}) extracted from the measurement are listed in Table 3.4. The R_C of both samples is less than 0.2Ω.mm, which is an indication of good Ohmic contacts considering the undoped nature of the cap layer. To complete the comparison, the R_{SH} of the TLM data and Hall data are contrasted. The very close matching between the TLM and Hall data implies that there was no sign degradation of the 2DEG carrier during the annealing process for both samples.

Sample name		**1 µm gate pHEMT**	**0.1 µm gate pHEMT**
R_C (Ω.mm)		0.12	0.15
R_{SH} (Ω/□)	**TLM**	173	179
	Hall	186	

Table 3.4 TLM measurement data for XMBE#171 1 µm and 0.1 µm gate devices (R_{SH} from the Hall data attached for comparison)

3.4.3.2 DC measurement

DC measurements were performed at room temperature by employing a Cascade Ground-Signal-Ground (GSG) probe station with 100 µm pitch separation and a HP 4142B DC parameter analyser. On the 0.1 µm gate sample, there were 16 devices with gate width of 2x100 µm written by E-beam lithography but only 3 of them were working properly, giving a device yield of ~18%. This was attributed to the poor E-beam patterning process due to unfamiliarity with the devices and inherent difficulty in using an essentially manual, mechanically scanned E-beam tool. Nevertheless, the best device was selected from the 0.1 µm sample to compare with the 1 µm ones which had a very high yield (> 90%). All measurements were conducted on devices with gate width of 2x100 µm and drain source separation of 5 µm and were normalised to their gate width. The average results of the measured 1 µm and 0.1 µm devices were presented.

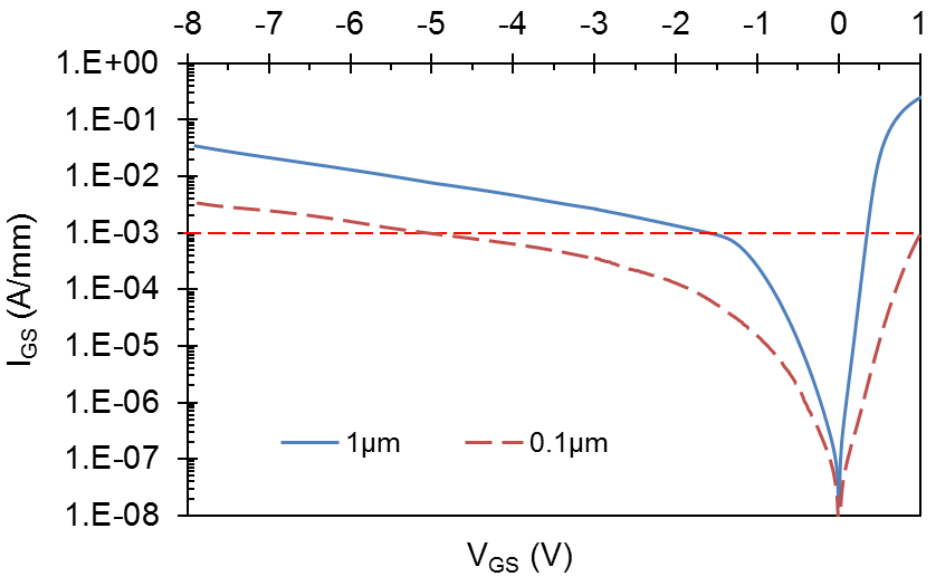


Figure 3.11 Schottky diode measurement for XMBE#171 devices with 1 µm and 0.1 µm planar gate

The results of Schottky diode measurements are depicted in Figure 3.11. The devices were biased from V_{GS} = 1V to -8V. The corresponding values of ideality factor (n) and Schottky barrier height (Φ_B) are extracted from the forward bias region by using the

technique described in section 2.6.4. The Φ_B of the 1 µm and 0.1 µm devices are 0.55eV and 0.57eV respectively. The similar value of Φ_B is due to the fact that both were based on the same epilayer structure and the fact that scaling is well maintained. However, their ideality factors show a significant discrepancy, with the 1 µm device having a value of 1.51 and the 0.1 µm device having a value of 2.16. For the 0.1 µm device, the high value of the ideality factor is an indication that the Schottky gate is unable to modulate the current as a proper diode.

In the reverse bias region, the off-state leakage current of the 0.1 µm device at V_{GS} = -8V is 3.2mA/mm, which is lower than that of 1 µm device by approximately an order of magnitude. Taking the value of 1mA/mm as a reverse breakdown reference, the 1 µm device shows a lower breakdown voltage of -1.5V while the 0.1 µm device pushes this further to V_{GS} = -5V.

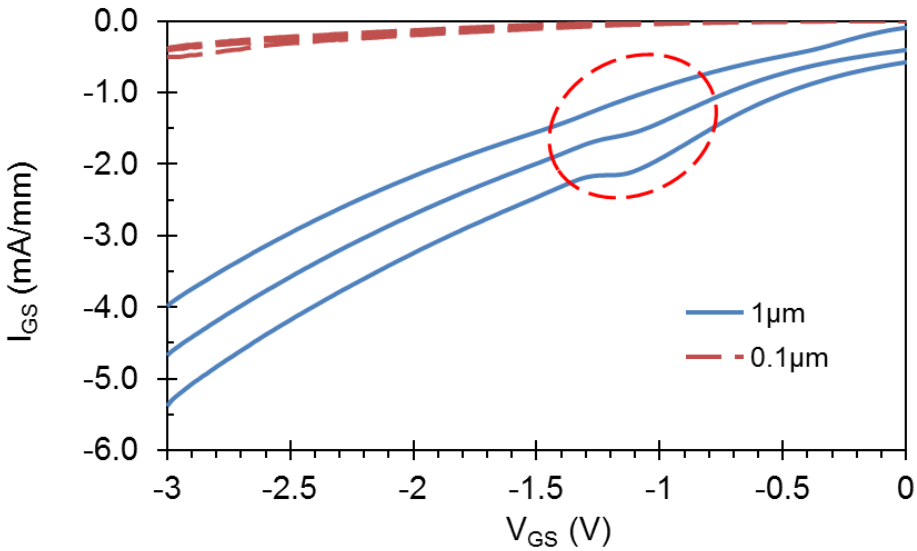


Figure 3.12 On-state leakage measurements for XMBE#171 devices with 1 µm (V_{DS} = 1 to 2V, step = 0.5V) and 0.1 µm planar gate (V_{DS} = 0 to 1V, step = 0.5V)

Figure 3.12 illustrates the on-state leakage for the 1 µm and 0.1 µm gate devices. The 1 µm device was biased at V_{DS} = 1 to 2V with step of 0.5V. The gate leakage current due to impact ionisation [41]-[42], as highlighted in the bell shape region, is rather insignificant in comparison with the leakage current due to tunnelling at all V_{DS}

values. The rather thin spacer layer and extremely high carrier concentration in the 2DEG are the key cause of the serious tunnelling current between the channel and Schottky barrier. For the 0.1 µm device, the device was biased at V_{DS} = 0 to 1V with step of 0.5V. The on-state leakage shows a similar trend to the case of off-state. At V_{DS} = 1V, its leakage current is about 10 times lower than that of 1 µm device.

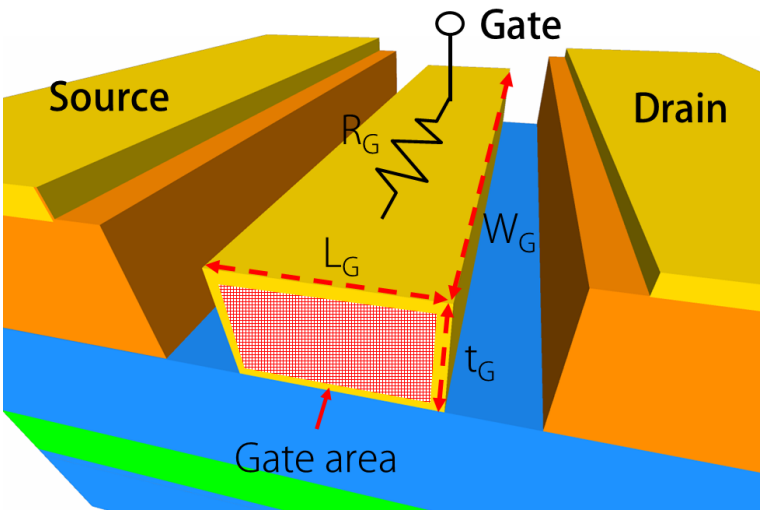


Figure 3.13 Relationship between the parasitic gate resistance and the gate geometry

As both 1 µm and 0.1 µm devices share the same epitaxial structure and similar values of Φ_B, this suggests that the reduction of the off-state and on-state leakage does not originated from their material or structural properties. However, the difference in their L_G implies a major effect on the parasitic gate accesses resistance (R_G), which is illustrated in Figure 3.13 and the much higher electric fields for the 0.1 µm device. The metal gate acts as a conductive path for the leakage current flowing between the semiconductor and the external circuitry. The magnitude of the gate current is directly proportional to R_G. R_G also depends on the gate geometry and the cross-sectional area in the direction of current propagation. The reduction of gate length from 1 µm to 0.1 µm results in an increase of R_G by almost 10 times. The tremendous access resistance of the submicron planar gate causes a significant decline on the current flow in both forward and reverse direction. This is the reason why the on-state and off-state leakage are improved on the 0.1 µm gate device.

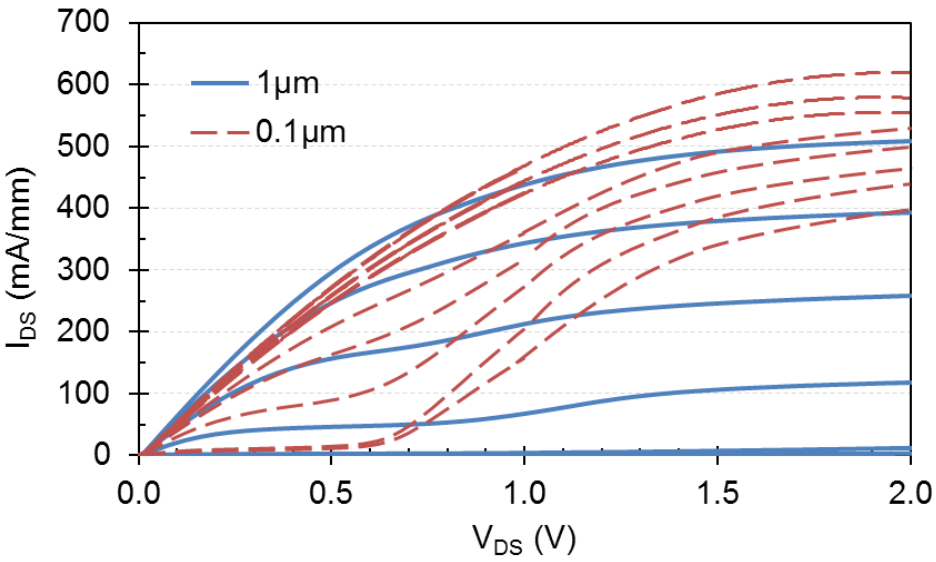


Figure 3.14 I_{DS} vs V_{DS} at different V_{GS} for XMBE#171 devices with 1 µm (V_{GS} = 0 to - 1.2V, step = - 0.3V) and 0.1 µm planar gate (V_{GS} = 0 to - 3.8V, step = - 0.475V)

The 1 µm and 0.1 µm devices behave very differently with regard to output characteristics, as depicted in Figure 3.14. Even though L_G was scaled down by a factor of 10, the 0.1 µm device exhibits only a minor improvement as indicated by its saturated drain-source current (I_{DSS}) of 620mA/mm when compared with the 1 µm device which had an I_{DSS} of 500mA/mm. Moreover, the I_{DS} of 0.1 µm device does not show a good saturation behaviour like the 1 µm device at all V_{GS} and the channel pinch off characteristic starts vanishing at V_{DS} = 0.6V.

The aforementioned phenomena can be explained by the short-channel effect (SCE). When the gate length scales down to very small dimensions, the gate-channel aspect ratio L_G/d, where d is the total distance from the top of barrier layer to the bottom of the channel layer, has to be maintained in order to accommodate the increasing drain-source electric field in the nanoscale channel [27].

According to [43], L_G/d should uphold a minimum value which is equal or greater than 5 (though values as small as 3 are sometimes acceptable) for preserving the device performance with nanoscale gate dimension. Our 0.1 µm gate device has an aspect ratio of 2.04 which is far below the desired aspect ratio value. The

controllability of the metal gate to the channel will become very weak due to the imbalance between gate-channel and drain-source electric fields. Under the influence of strong drain-source electric fields, the phenomenon of channel punch-through takes place, where a leakage current is developed in the channel due to the effect of drain-induced barrier lowering (DIBL) [27]. As a net result, the output characteristics of the 0.1 µm device exhibits poor pinch-off behaviour and increased output conductance (G_d), due to the large I_{DS} at higher V_{DS} as shown in Figure 3.14.

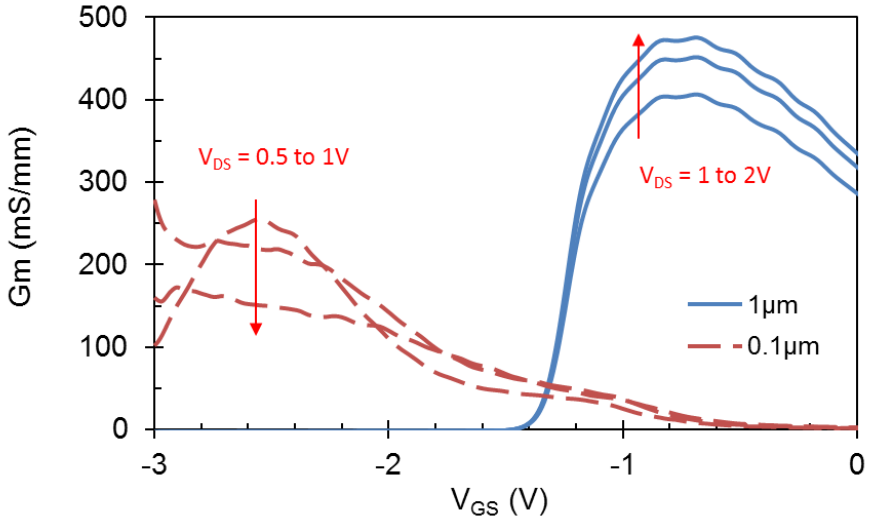


Figure 3.15 Gm vs V_{GS} at different V_{DS} for XMBE#171 devices with 1 µm (V_{DS} = 1 to 2V, step = 0.5V) and 0.1 µm planar gate (V_{DS} = 0.5 to 1V, step = 0.25V)

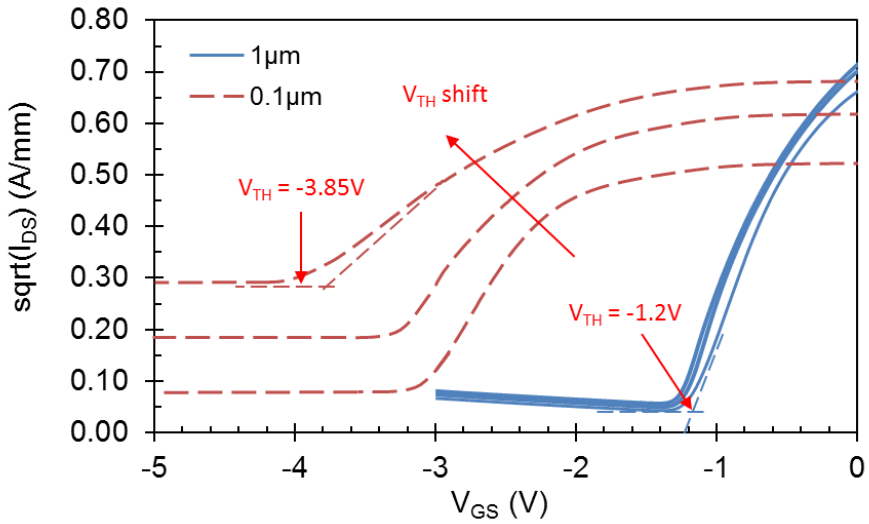


Figure 3.16 Square root of I_{DS} vs V_{GS} at different V_{DS} for XMBE#171 devices with 1 µm (V_{DS} = 1 to 2V, step = 0.5V) and 0.1 µm planar gate (V_{DS} = 0 to 1V, step = 0.5V)

The negative impact of SCE is also replicated on the transconductance (Gm) and sub-threshold voltage characteristics, as illustrated in Figure 3.15 and Figure 3.16 respectively. Compared with the 1 µm device, the 0.1 µm device exhibits a reduction of about 50% on its peak Gm and it becomes even lower at higher V_{DS}. In addition, a significant amount of threshold voltage (V_{TH}) shift is observed in the 0.1 µm device. Its V_{TH} is shifted by about -2V from the value of the 1 µm device even at V_{DS} = 0V and it becomes more serious with increased V_{DS}. This is in line with the findings in [44] that the amount of V_{TH} shift is dependent on L_G and V_{DS} on device showing SCE.

3.4.3.3 RF measurement

RF measurements were performed on the two samples using an HP 8510C VNA with a standard LRRM calibration technique. They were characterised at room temperature in a frequency range from 45MHz to 110GHz. The average results of the measured 1 µm and 0.1 µm devices were presented. Figure 3.17 shows their current gain and power gain as a function of frequency. Their bias points were set to the V_{GS} located at the maximum Gm in order to achieve the lowest NF_{min} and the corresponding bias conditions with measurement error are shown in Table 3.5.

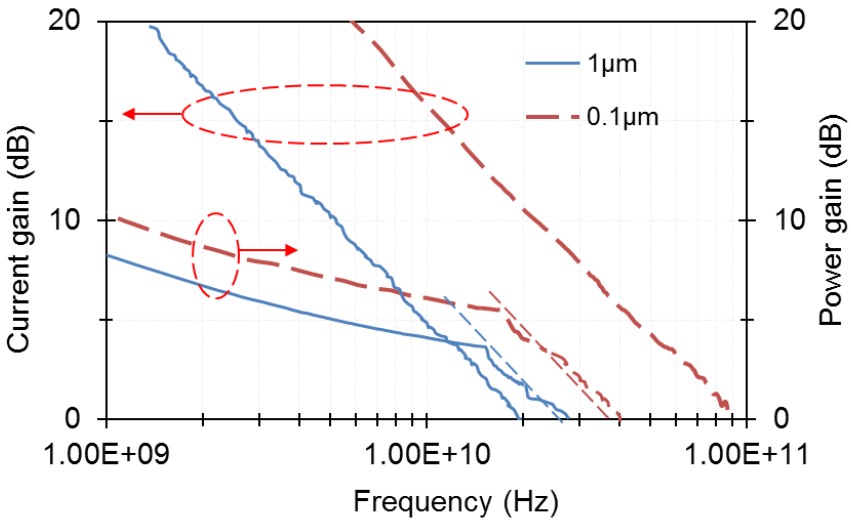


Figure 3.17 Frequency response of current and power gains for XMBE#171 devices with 1 µm (V_{DS} = 1V) and 0.1 µm planar gate (V_{DS} = 0.5V) (Unity current gain and power gain are obtained by the 0dB crossing at x-axis)

Device summery – Gate width = 2x100µm, drain-source spacing = 5µm		
Parameters	**1µm planar gate**	**0.1µm planar gate**
V_{DS} (V)	1.0	0.5
V_{GS} (mV)	-580 ±5	-2380 ±5
Gm (mS/mm)	464 ±5	257 ±5
I_{DS} (mA)	39.6 ±1	48.8 ±1
f_T (GHz)	20 ±0.5	94 ±1
f_{max} (GHz)	28 ±0.5	38 ±0.5

Table 3.5 Bias conditions for the for XMBE#171 devices with 1µm and 0.1µm planar gate device

Even with the aggressive scaling of L_G, the current gain cut-off frequency (F_T) on 0.1µm device is only about 4.5 times that of the 1µm device, increasing from 20GHz to 94GHz. The degradation of Gm from the SCE, together with the arising parasitic elements R_S and R_D due to the large drain-source separation, have hindered the f_T from further enhancement.

The impact of parasitic elements is more substantial on the maximum oscillation frequency (f_{max}). The 0.1µm device demonstrates a f_{max} of 38GHz, which is less than 2 times the value of the 1µm device and is only about half of its f_T. The degradation of f_{max} in the 0.1µm device is attributed to the increased R_S, R_D and in particular R_G due to the massive reduction on the cross-sectional area of the gate metallisation. The optimisation of device geometries is evidently a very important task to explore the full potential of the submicron device.

3.5 Fabrication of 128nm T-gate pHEMT with optimised geometry

3.5.1 Introduction

The aim of the study was to address the performance issue encountered in our first attempt at making submicron pHEMTs. The problems of SCE and parasitic elements could be resolved by means of an improved material epitaxial layer profile and optimised device geometries. The improvements were demonstrated through the fabrication of 128nm T-gate pHEMT as shown in the following sections.

3.5.2 Material growth

Two new wafers, XMBE#129 and XMBE#169, were prepared to solve the problem of SCE for the fabrication of 128nm T-gate devices in this study. Their layer profiles are shown in Table 3.6 and Table 3.7, along with the Hall Effect data shown in Table 3.8. Both structures are based on an InP substrate with a thick $In_{0.52}Al_{0.48}As$ buffer layer. Unlike the conventional structure XMBE#171 mentioned previously, they both contain a double δ-doping scheme, and the isolation of ionised impurity is done by sandwiching the channel within two separate spacers, resulting in a better carrier confinement of the 2DEG. To make an equal comparison, their sheet carrier concentration and mobility were deliberately matched by adjusting the individual doping parameters.

Epi-structure data - XMBE#129		
Epilayer	**Material**	**Thickness (Å)**
Cap	$In_{0.53}Ga_{0.47}As$	50
Barrier	$In_{0.52}Al_{0.48}As$	150
δ-doping 1 (Si-doped)		
Spacer 1	$In_{0.52}Al_{0.48}As$	100
Channel	$In_{0.70}Ga_{0.30}As$	150
Spacer 2	$In_{0.52}Al_{0.48}As$	100
δ-doping 2 (Si-doped)		
Buffer	$In_{0.52}Al_{0.48}As$	4500
Substrate	InP (Fe-doped)	

Table 3.6 Epitaxial structure data of XMBE#129

Epi-structure data - XMBE#169		
Epilayer	**Material**	**Thickness (Å)**
Cap	$In_{0.53}Ga_{0.47}As$ (Si-doped, n= $9.5x10^{11}$ cm^{-2})	50
Barrier	$In_{0.25}Al_{0.75}As$	70
δ-doping 1 (Si-doped)		
Spacer 1	$In_{0.25}Al_{0.75}As$	28
Channel	$In_{0.70}Ga_{0.30}As$	150
Spacer 2	$In_{0.52}Al_{0.48}As$	100
δ-doping 2 (Si-doped)		
Buffer	$In_{0.52}Al_{0.48}As$	4500
Substrate	InP (Fe-doped)	

Table 3.7 Epitaxial structure data of XMBE#169

Hall effect measurement data		XMBE#129	XMBE#169
Sheet carrier concentration (n_H) ($x10^{12}$ cm^{-2})	**300 K**	1.72	1.70
	77 K	1.90	1.90
Hall mobility (μ_H) (cm^2/V.s)	**300 K**	10000	10000
	77 K	25200	26000
Sheet resistance (R_{SH}) (Ω)		363	367

Table 3.8 Hall Effect measurement data for XMBE#129 and XMBE#169

However the two structures differ greatly in their layer profiles, especially in the Schottky and spacer layers. This is illustrated in their energy band diagram simulation in Figure 3.18 and Figure 3.19.

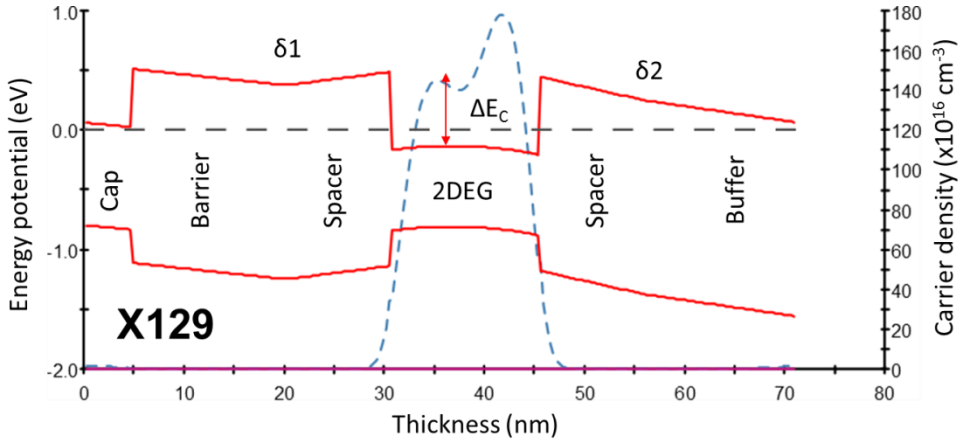


Figure 3.18 Energy band diagram of XMBE#129 structure (V_{GS} = 0, V_{DS} = 0)

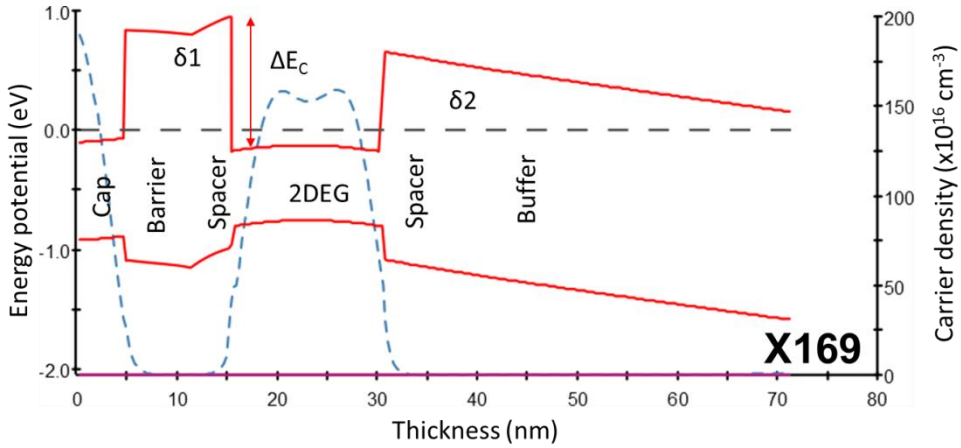


Figure 3.19 Energy band diagram of XMBE#169 structure (V_{GS} = 0, V_{DS} = 0)

For XMBE#129, it has a relatively thin $In_{0.52}Al_{0.48}As$ layer sitting on top of the $In_{0.70}Ga_{0.30}As$ channel, forming a heterojunction with ΔE_C = 0.62eV [45]. This results in a L_G/d = 3.2 with L_G = 128nm, which is approximately 60% improvement from the previous sample.

For XMBE#169, this aspect ratio is elevated further to 5.2 with its thinner barrier and spacer layers. A very wide band gap $In_{0.25}Al_{0.75}As$ material with a ΔE_C = 1.27eV [46], is employed to compensate for the increased tunnelling current due to the scaling

down of the barrier layer. In order to minimise the contact resistance, a heavily doped cap layer is grown on the XMBE#169, which does help with reducing R_S and R_D.

In this study, the benefit of increased the gate-channel aspect ratio, together with the improvement of the doped cap layer and the ultra-wide band gap barrier layer on the improved epitaxial structures will be demonstrated through the fabrication of a high performance 128nm T-gate pHEMT.

3.5.3 Device fabrication

The studies for the 128nm T-gate pHEMT device was conducted in a joint collaboration with the Institut d'Electronique, de Microélectronique et de Nanotechnologie (IEMN) in the University of Lille Nord de France. The device fabrications were carried out with our in-house growth wafers (XMBE#129 and XMBE#169) at the IEMN facilities and the completed devices were returned for characterisation. Both laboratories share a similar fabrication process, as described in section 3.3, but their geometries are optimised for submicron device using a different mask set to the Manchester ones. On the new devices, two major improvements have been made to tackle the performance issue due to parasitic elements.

Figure 3.20 Comparison between planar gate and T-gate structure

Firstly, a T-gate structure, also referred to as mushroom gate, was employed to replace the original planar gate design in the previous study. A comparison between them is depicted in Figure 3.20. Unlike the planar gate structure, the T-gate has a separate gate head and gate stem elements. The size of the gate footprint (L_G) is defined by the gate stem. The gate head is designed to lower the value of R_G by increasing the effective cross-sectional area of the gate metal contact. Hence, f_{max} will be expected to be greatly improved in the new devices.

Secondly, R_S and R_D have serious influences on both f_T and f_{max} on small scale devices. As shown in Figure 3.21, the contact resistance (R_C) and the path access resistance to the intrinsic gate region (R_A) are fundamental elements for R_S and R_D. The R_C can be suppressed by introducing a doped cap layer scheme as used on the

XMBE#169 wafer. To minimise the R_A, the drain-source separation (L_{SD}) was reduced from 5 µm in the previous study to 1.5 µm. This is illustrated in the digital image shown in Figure 3.22. With the improvement of the two new features, the effective R_S and R_D were reduced to a great extent, hence maximising the device microwave performance.

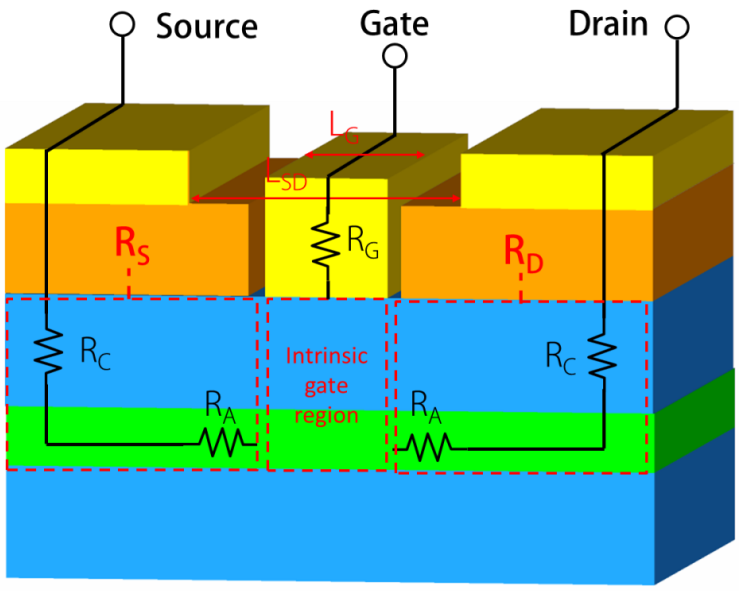


Figure 3.21 Parasitic resistance paths to the intrinsic gate region between the drain and source contacts

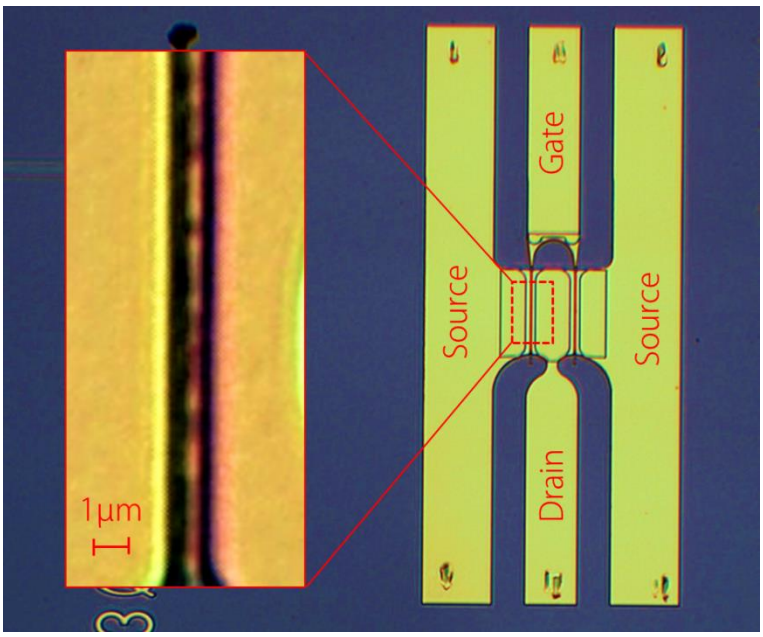


Figure 3.22 Digital image of the fabricated 2x100 µm device with 1.5 µm drain-source separation

3.5.4 Results and discussions

3.5.4.1 DC measurement

The fabricated devices with different gate widths were characterised using the same configuration as mentioned in the previous study. For simplicity, only devices with gate width of 2x100 µm and drain-source separation of 1.5 µm will be shown in the study. Measurements were conducted at room temperature and the results are normalised to the device gate width. The average results of the measured devices were presented. The legends (X129 and X169) on the plots are referred to devices fabricated on the XMBE#129 and the XMBE#169 wafers respectively.

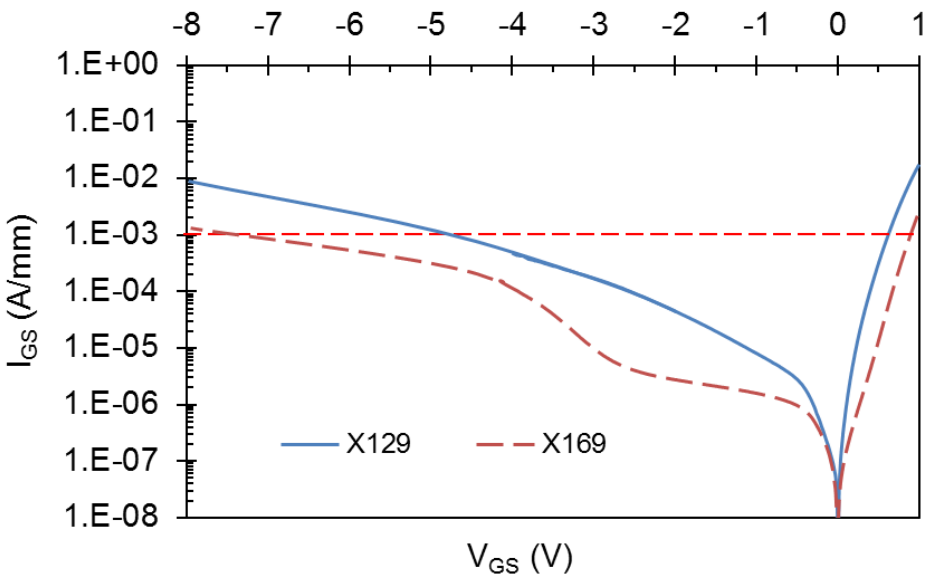


Figure 3.23 Schottky diode measurement for the 128nm T-gate pHEMTs on XMBE#129 and XMBE#169 epitaxial structures

The results of Schottky diode measurement are shown in Figure 3.23. In the forward bias region, the X129 and X169 devices show a barrier height value (Φ_B) of 0.57eV and 0.93eV and an ideality factor (n) of 1.78 and 1.54 respectively. On the X169 device, the highly strained barrier layer of $In_{0.25}Al_{0.75}As$ results in a higher Φ_B in comparison with the lower lattice matched material $In_{0.53}Al_{0.47}As$ in the X129 device. The benefit of the enhanced Φ_B is even more apparent in the reverse bias region,

where the off-state leakage current of the X169 device is about an order of magnitude lower than that of the X129 device. By taking the breakdown reference as 1mA/mm, the X129 device yields a breakdown voltage of -4.8V and this is extended to -7.5V on the X169 device.

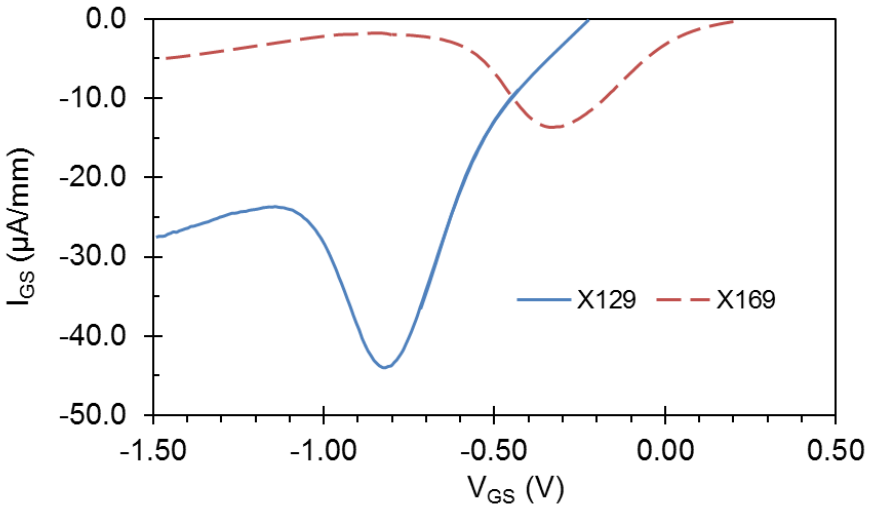


Figure 3.24 On-state leakage measurement for the 128nm T-gate pHEMTs on XMBE#129 and XMBE#169 epitaxial structures (V_{DS} = 1.0V)

A similar situation is also observed in the on-state leakage measurement as depicted in Figure 3.24. In terms of the peak leakage due to impact ionisation at V_{DS} = 1.0V, the X169 device demonstrates a leakage current (I_{GS}) of 15 µA/mm when compared with the X129 device which has I_{GS} = 45 µA/mm. The suppression of both off-state and on-state leakage on the X169 device is a result of the improved material growth with reduced Indium content in the barrier layer.

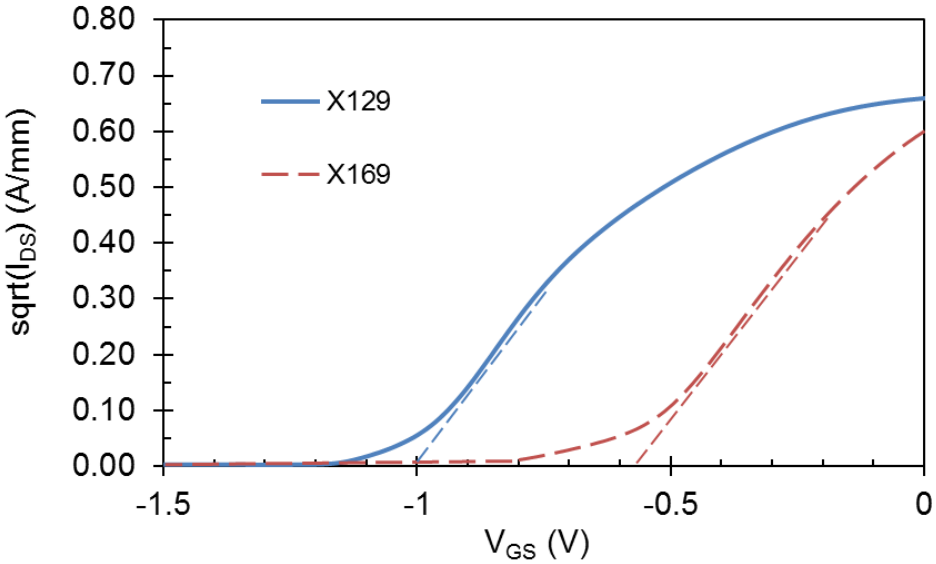


Figure 3.25 Square root of I_{DS} vs V_{GS} at different V_{DS} for the 128nm T-gate pHEMTs on XMBE#129 and XMBE#169 epitaxial structures (V_{DS} = 1.0 V)

The square root of I_{DS} dependence on V_{GS} at V_{DS} of 1.0V is illustrated in Figure 3.25. The suppression of the SCE is evidenced from the good pinch-off characteristic on both X129 and X169 devices. However, the latter exhibits a better current drivability from the comparison of their linear regions. In addition, from the linear extrapolation of the threshold voltage (V_{TH}), the X169 device demonstrates a more positive V_{TH} of -0.6V from the X129 device which has a V_{TH} = -1.0V. The aforementioned features of the X169 device are attributed to the reduced gate-channel separation (d) in its structural profile.

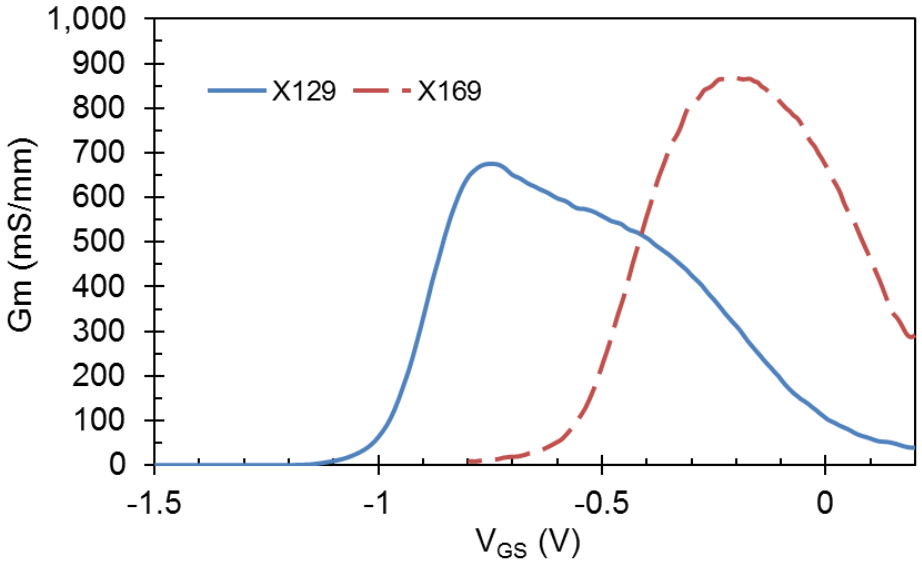


Figure 3.26 Gm vs V_{GS} at different V_{DS} for the 128nm T-gate pHEMTs on XMBE#129 and XMBE#169 epitaxial structures (V_{DS} = 1V)

The enhancement with thinner barrier structure is also shown in the transconductance (Gm) measurement, as depicted in Figure 3.26. The peak Gm at V_{DS} = 1.0V is 670mS/mm and 860mS/mm for the X129 and X169 devices respectively. By employing a thinner barrier structure, they yield an improvement of 1.7 to 3.5 times Gm over the flat gate 0.1 µm device described in the previous study.

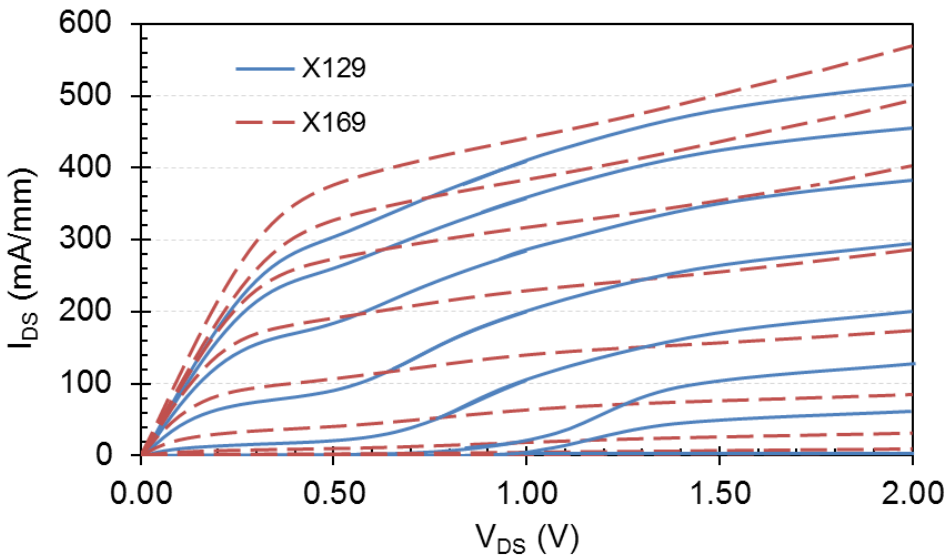


Figure 3.27 I_{DS} vs V_{DS} at different V_{GS} for the 128nm T-gate pHEMTs on XMBE#129 (V_{GS} = - 1 to -0.2, step = - 0.1V) and XMBE#169 (V_{GS} = -0.6 to +0.2V, step = - 0.1V) epitaxial structures

Figure 3.27 is a comparison of the output characteristics between the X129 and X169 devices. Considering the difference in their V_{TH}, all the devices were biased with a fixed overdrive voltage (ΔV) from their V_{TH}, where the range of the gate bias voltage was varied from $V_{GS} = V_{TH}$ to $V_{GS} = V_{TH} + \Delta V$, and $\Delta V = 0.8V$, ensuring that their channel will be modulated by an equal amount of gate electric field.

It is not surprised that they both show a peak I_{DS} over 500mA/mm at $V_{DS} = 2V$, as their sheet carrier concentration and mobility are virtually the same from the Hall data. But, a moderate level of SCE is observed on the X129 device, where the I_{DS} suddenly increases from $V_{DS} = 0.5V$ to 1.0V, leading to a degradation of the device output conductance. On the contrary, no change is observed across the whole range of V_{GS} in the saturation region of the X169 device. Clearly, the reduced gate-channel aspect ratio ($L_G/d = 3.2$) on the X129 device results in its channel punch-through characteristic at higher V_{DS}. This suggests that the L_G/d should be maintained at a minimum value of 5, as employed in the X169 device, in order to eliminate the SCE completely.

3.5.4.2 RF measurement

Microwave measurements were conducted on the 2x100 µm device with L_{SD} = 1.5 µm drain-source separation using the same configuration as mentioned in the last study. The devices were characterised from 45MHz to 110GHz with LRRM calibration standard at room temperature using the similar setup as in section 3.4.3. The average results of the measured devices were presented. Figure 3.28 illustrates the dependence of the current gain and power gain as a function of frequency and the corresponding bias conditions with error are tabulated in Table 3.9.

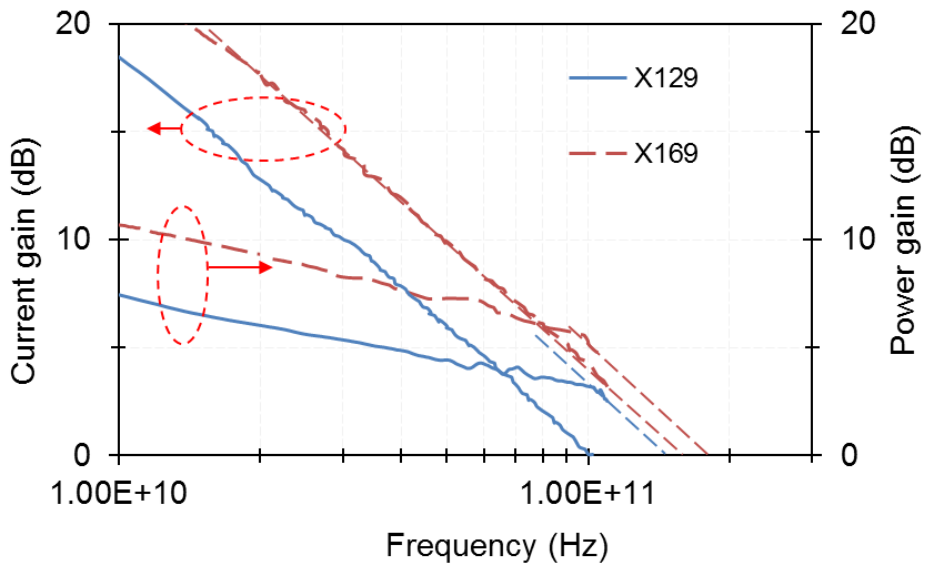


Figure 3.28 Frequency response of current and power gains for the 128nm T-gate pHEMTs on XMBE#129 and XMBE#169 epitaxial structures. (V_{DS} = 1V) (Unity current gain and power gain are obtained by the 0dB crossing at x-axis)

Device summery – Gate width = 2x100 µm, drain-source spacing = 1.5 µm		
Parameters	**XMBE#129**	**XMBE#169**
V_{DS} (V)	1.0	1.0
V_{GS} (mV)	-750 ±5	-220 ±5
Gm (mS/mm)	672 ±5	867 ±5
I_{DS} (mA)	13.9 ±1	16.4 ±1
f_T (GHz)	103 ±1	163 ±2
f_{max} (GHz)	151 ±1	182 ±2

Table 3.9 Bias conditions for the for the 128nm T-gate pHEMTs on XMBE#129 and XMBE#169 epitaxial structures

Due to the limited capability of the VNA kit, the f_T and f_{max} were extrapolated by using the best linear fitting with a slope of 20dB/decade as shown in the plots. The f_T and f_{max} of the X129 device are 103GHz and 151GHz respectively and these figures are increased to 163GHz and 182GHz for the X169 device. The microwave performance of the T-gate devices are far better than that of the planar gate device from the previous work, even though the latter had a slightly smaller gate dimension (L_G = 128nm vs L_G = 120nm for the previous study), which is a net result of the following three factors:

Firstly, on the new devices, L_{SD} is dramatically reduced from 5 µm to 1.5 µm, which is 70% less than the previous study. This yields a higher f_T due to the reduction of R_S and R_D. Secondly, R_G is suppressed with the new T-gate structure by increasing the cross-sectional area of the metal conductive path. The reduction of R_G leads to a massive improvement of f_{max}. These two items, along with the improved SCE by means of the increased L_G/d ratio, are responsible for the significant enhancement of f_T and f_{max} observed in the study.

It should be noted that the X129 device shows worse performance against the X169 device, even though they share similar device structures and material properties. The X169 device has a L_G/d = 5.2, which is higher than that of the X129 device. The increased L_G/d suppresses the SCE and improves the DC and RF performance on small scale devices. Also, unlike the X129 device, it also has a heavily doped cap layer scheme, which reduces the contact resistance (R_C) and the parasitic access resistances (R_S and R_D), and consequently increases its f_T and f_{max}. Therefore, it is essential to maintain a good gate-channel aspect ratio (L_G/d >5) and minimise R_C to achieve the optimised performance from submicron pHEMT devices.

3.6 Issues and improvements

The fabrication of submicron gate pHEMTs by means of advanced E-beam lithography has demonstrated a great potential for the enhancement of DC and RF performance. However, a few issues were encountered with regard to device yield and processing time, as summarised on Table 3.10.

Epi-layer	XMBE#171	XMBE#129	XMBE#169
Gate structure	120nm planar gate	128nm T-gate	128nm T-gate
Device yield	18%	31%	91%
Processing time	> 2 months	> 6 months	> 6 months

Table 3.10 Comparison between the yields of the submicron gate devices

The device yield was defined as the number of working device, which was determined from DC measurements. This is mainly dependent on the fabrication process, in particular the E-beam patterning step, in which factors such as the E-beam exposure condition and the operator experience could affect the end product performance. The large variation between their yields outlines the poor repeatability of the E-beam process in this study. It also introduced extra uncertainty in the measured data and consumed too many expensive materials.

The next issue is about the long processing time, which is the total time it takes to complete a process cycle from wafer to devices. For a typical 1 μm gate process with i-line photography, the process timescale is less than a month in comparison with the submicron process of E-beam lithography, which can take up to 6 months to complete. The slow writing speed of E-beam lithography and the long transit time for the sample transfer between different institutes mainly contribute to the long process.

The main purpose of the joint collaboration was to perform study in an effective and economical way by harnessing the strengths of the different institutes. However, those benefits were, to some extent, negated by the yield and time issues arising from the process. It is evident that a solution has to be developed for the fabrication of submicron pHEMT in a way that is far more efficient and independent and this the

subject of a study presented in chapter 5.

3.7 Summary

In this chapter, we presented our effort on the fabrication of submicron gate InP-based pHEMT by employing advanced E-beam lithography in joint collaborations with other research institutes. The study began with a 120nm planar gate pHEMT based on a conventional epitaxial structure with single δ-doping scheme. The DC and RF characteristics of the fabricated device were severely compromised by the serious SCE and also the influence of parasitic elements, R_G, R_D and R_S. The planar devices only yielded a relatively low f_T of 94GHz and a poor f_{max} of 38GHz.

These problems were addressed by optimising the material epitaxial layer profiles and also the device geometry in the second part of the study. The SCE were suppressed by increasing the L_G/d ratio through thinning the barrier layer, and a highly strained $In_{0.25}Al_{0.75}As/In_{0.30}Ga_{0.70}As$ channel scheme was employed to compensate the increased tunnelling current due to the barrier thinning. Regarding the problem of the parasitic elements, a heavily doped cap layer structure along with the reduced L_{SD} was proposed to minimise R_S and R_D. The increased R_G due to the shrinkage of gate dimension was offset by means of a T-gate structure.

With all these improvements, a 128nm T-gate with highly strained $In_{0.25}Al_{0.75}As$ $/In_{0.30}Ga_{0.70}As$ barrier and channel and heavily doped cap layer scheme has been fabricated, exhibiting an excellent f_T of 163GHz and f_{max} of 182GHz. The potential of submicron InP-based pHEMT for high speed and low noise applications has thus been demonstrated through this work and for phase 2 and 3 of the SKA, leaving cost effectiveness as the only remaining issue and this is addressed in chapter 5.

4 HEAT TREATMENT STUDY OF STRAINED InGaAs-InAlAs pHEMT USING A PALLADIUM GATE METALLISATION

4.1 Introduction: The importance of enhancement mode pHEMT

III-V compound semiconductors have been recognised as prospective contenders for the International Technology Roadmap for Semiconductors (ITRS) post 22nm node due to their high intrinsic electron mobility and high carrier densities [47]-[48]. The excellent DC and RF performance of III-V materials attracts many interests in such various applications as LNA, PA and sub-millimetre wave circuits approaching THz regimes in nanoscale devices. In particular, pHEMT, utilising the high carrier mobility property of a 2DEG structure, has been drawing a great deal of attention by researchers.

Unlike the MOSFET or HBT competitors, conventional pHEMTs are used mainly in depletion mode (D-mode). The requirement for an additional negative bias source for the circuit design showcases many drawbacks in terms of cost, power consumption and circuit complexity [49]-[50].

In the past, many efforts were devoted to the development of enhancement mode (E-mode) pHEMT. A number of different approaches such as deep gate recess etching, use of strained barrier layer or use of higher barrier gate metals have been studied in the GaAs-based devices. Unlike the InGaAs-AlGaAs E-mode pHEMT, the InP-based system has much higher mobility and results in an impressive improvement in device performances. In this work, heat treatment studies were carried out on InGaAs - InAlAs pHEMTs by resorting to a palladium (Pd) buried gate, in an effort to investigate the change of devices characteristics during the annealing process.

4.2 Solutions for enhancement mode pHEMT

The depletion nature of pHEMT structure makes it rather difficult to achieve enhancement mode (E-mode) device. To understand the origin of the problem, it is necessary to know about the threshold voltage (V_{TH}) equation mentioned in section 2.5.1. It contains three key items, barrier height (Φ_B), conduction band discontinuity (ΔE_C) and pinch-off voltage (V_P). An E-mode device should have a $V_{TH} \geq 0$ by definition. This can be achieved by increasing Φ_B while keeping the other two items as small as possible by introducing modifications in the material properties and device structures. In this section, a detailed discussion on the fabrication of E-mode device with InGaAs-InAlAs and InGaAs-AlGaAs strained channels will be presented.

$$V_{\mathrm{TH}} = \Phi_B - \frac{\Delta E_C}{q} - V_P$$

For the InGaAs-InAlAs system, a reduction of ΔE_C can be accomplished by decreasing the Indium content in the InGaAs channel at the expense of lowering the sheet carrier density of the 2DEG, and also increasing the tunnelling leakage current. Moreover, the origin for the difficulty of making E-mode device on InP substrate is that they have a ΔE_C about twice larger than that of the GaAs-based HEMTs ($In_{0.3}Ga_{0.7}As$-$In_{0.53}Al_{0.47}As$ = 0.62eV and of $In_{0.15}Ga_{0.85}As$-$Al_{0.15}Ga_{0.85}As$ = 0.3eV) [26, 45].

The term V_P is correlated to the dielectric constant of the barrier layer material (ε), the doping concentration of the supply layer (n) and the gate-channel separation (d). The reduction of V_P by decreasing ε and n is not a feasible option as it will also decrease the sheet carrier density of the 2DEG. An alternative solution is to reduce d by employing gate recess techniques with deep etching as reported in [51]-[52]. However, this it often accompanied by problems of low device uniformity due to the absence of an etch stop layer, and as a consequence the device exhibits a large V_{TH} variations and poor Schottky characteristics. Furthermore, the device reliability is greatly affected by

the increasing tunnelling leakage with reduced d.

Another approach is to increase Φ_B by employing a barrier layer with increased energy band gap or a metal with higher work function. The first item can be achieved by adjusting the composition of the barrier layer and the fabrication of E-mode pHEMT has been demonstrated by using strained $In_{0.7}Ga_{0.3}As$-$In_{0.45}Al_{0.55}As$ channel [53]. In theory, it is also possible to increase Φ_B by employing an underlying metal with higher metal work function, such as Palladium (Pd) or Platinum (Pt), as shown in Figure 4.1, but this only exerts a slight impact on the effective Φ_B in practice, due to the presence of Fermi-level pinning effect in III-V materials.

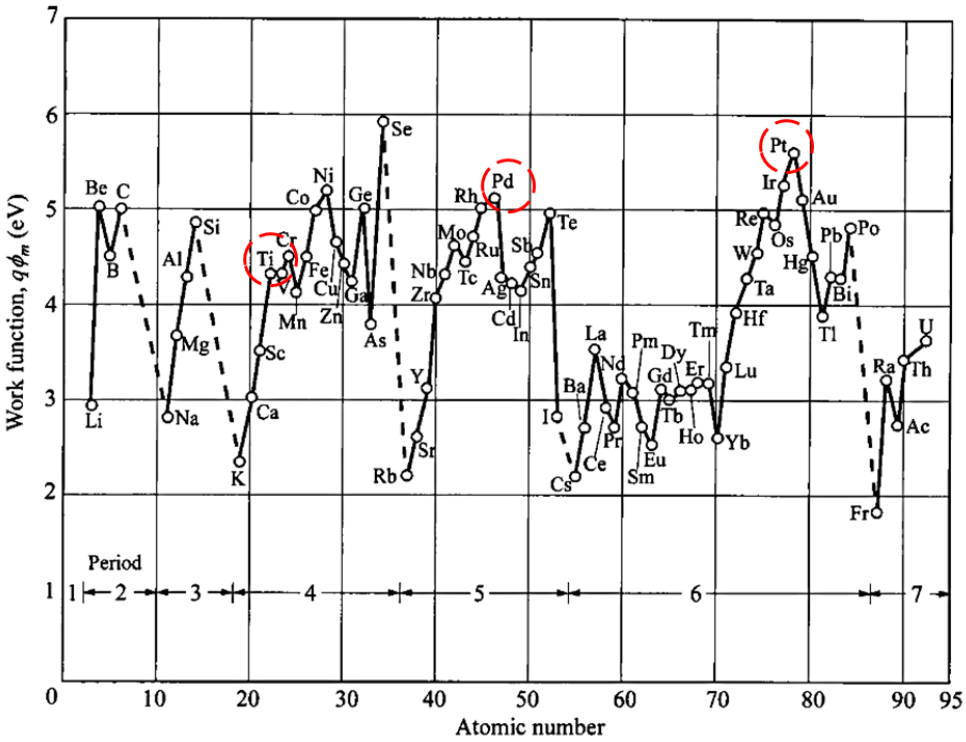


Figure 4.1 Dependence of metal work function on atomic number [27]

Refractory metals such as Pd and Pt, which have rather strong diffusion in III-V compounds, are widely employed to reduce d and decrease the effective V_P in E-mode pHEMT fabrication. Extensive studies have been done on E-mode devices by resorting to Pt buried gate structure on GaAs and InP substrates respectively [53]-[58]. However, Pd has become a better alternative than Pt for the following reasons. Firstly, the

commodity price of Pt is more than double that of Pd which is important in some cost sensitive applications. Secondly, Pd has a higher vapour pressure and a lower melting point compared with Pt [62], which makes conventional thermal evaporation possible with Pd and nearly impossible with Pt. Therefore the use of Pd leads to a process friendly option without necessarily using the high cost of electron-beam evaporators, and associated undesirable damage to the Schottky interface during the evaporation. More importantly, the continuous diffusion of Pt is always a concern as device performance might degrade during successive heat treatment or during high power operation [55].

On the contrary, the Pd diffusion tends to saturate after annealing at a given temperature [59]. A Pd buried gate E-mode pHEMT with Φ_B = 0.58eV has been reported in InGaAs-AlGaAs device [60]. Recently, Pd buried gate E-mode pseudomorphic hetero-junction structure (Hall bar) with Φ_B = 0.60 to 0.68eV have been studied using InGaAs-InAlAs structures in [61] but that particular structure is not suitable for transistor operation due to its low carrier density and poor transconductance.

In this work, quasi-enhancement mode $In_{0.7}Ga_{0.3}As$-$In_{0.52}Al_{0.48}As$ pHEMTs were fabricated by using a Pd/Ti/Au gate metallisation. By combining the Pd-gate diffusion effect and the benefit of highly strained InGaAs-InAlAs channel, the device performance has shown over 50% enhancement to those reported in [60]-[61]. The interrelations between Pd metal thickness, heat treatment time and temperature are explored to evaluate the thermal stability of these Pd-gate devices.

4.3 Fabrication of 1 μm gate pHEMT using Pd/Ti/Au and Ti/Au metallisation

4.3.1 Material growth

An epitaxial structure called XMBE#131 was employed for this study as shown in Table 4.1.

Epi-structure data - XMBE#131		
Epilayer	**Material**	**Thickness (Å)**
Cap	$In_{0.53}Ga_{0.47}As$	50
Barrier	$In_{0.52}Al_{0.48}As$	150
δ-doping 1 (Si-doped)		
Spacer 1	$In_{0.52}Al_{0.48}As$	100
Channel	$In_{0.70}Ga_{0.30}As$	140
Spacer 2	$In_{0.52}Al_{0.48}As$	100
δ-doping 2 (Si-doped)		
Buffer	$In_{0.52}Al_{0.48}As$	4500
Substrate	InP (Fe-doped)	

Table 4.1 Epitaxial structure data of XMBE#131

The structure is grown on semi-insulating InP substrates, incorporating a 4500Å lattice-matched InAlAs buffer layer which isolates the surface states and decouples the potential impurity out-diffusion from the substrate. A double δ-doping scheme is employed to enhance the sheet carrier density in the channel. The strained $In_{0.7}Ga_{0.3}As$ channel is sandwiched by two $In_{0.52}Al_{0.48}As$ spacer layers to protect the 2DEG from ionised donor scattering. A 150Å wide band gap $In_{0.52}Al_{0.48}As$ barrier (supply) layer was grown for the formation of a Schottky barrier. Finally, an undoped 50Å narrow band gap $In_{0.53}Ga_{0.47}As$ cap layer was grown to protect the barrier layer from oxidation and lower the resulting barrier height for the formation of Ohmic contacts.

Hall effect measurement data - XMBE#131		
Sheet carrier concentration (n_H) ($x10^{12}$ cm^{-2})	300 K	2.40
	77 K	2.50
Hall mobility (μ_H) (cm^2/V.s)	300 K	13896
	77 K	47829
Sheet resistance (R_{SH}) (Ω)		155

Table 4.2 Hall Effect measurement data of XMBE#131

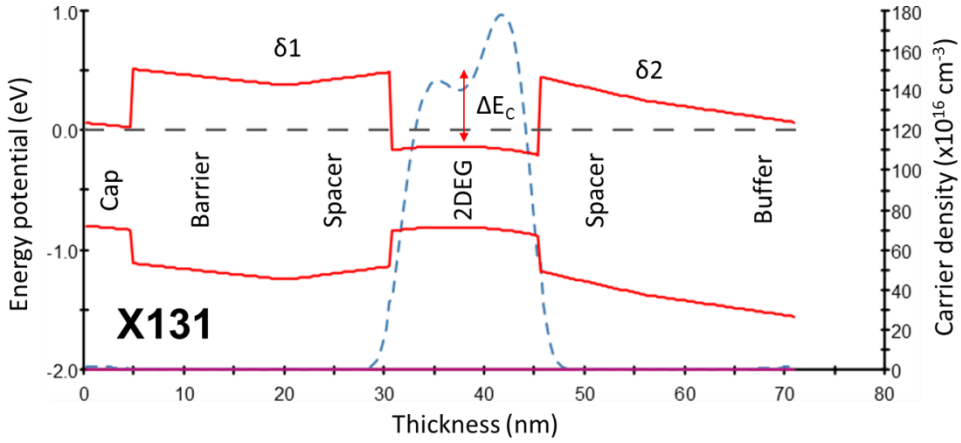


Figure 4.2 Energy band diagram of XMBE#131 structure (V_{GS} = 0, V_{DS} = 0)

The Hall Effect measurement data and energy band diagram using WinGreen® simulation are shown in Table 4.2 and Figure 4.2 respectively. Thanks to the enhanced carrier confinement by the double δ-doping scheme and the reduced ionised scattering by the thick spacer structures, the epitaxial structure exhibits a sheet resistance (R_{SH}) of 155 (Ω/□) and a corresponding carrier concentration (n_H) and mobility (μ_H) of $2.4x10^{12}$ (cm^{-2}) and 13896 (cm^2/Vs) at room temperature. These excellent figures are 30-100% better than similar GaAs and InP-based structures reported in [60] and [61]. It is expected that the device performance will be improved substantially due to the excellent material properties.

4.3.2 Device fabrication

Four samples (Ti-gate, 5nm Pd-gate, 10nm Pd-gate and 20nm Pd-gate), based on identical epilayer structures but with different gate metallisation schemes, were prepared for the study. The fabrication process was similar to the baseline process mentioned in section 3.3 and consisted of mainly five steps. The MESA step was defined by using a phosphoric based etchant followed by a selective side-wall etching to eliminate possible MESA side-wall leakage and thereby improving thermal stability during successive heat treatments.

Unlike the previous studies, a thinner Ohmic metallisation scheme was employed for cost saving on the valuable gold material. The alloyed Ohmic contact with AuGe/Au = 50/100nm was deposited by thermal evaporation, followed by a short annealing in a nitrogen (N_2) filled furnace for 90s at 280℃. The results of TLM measurement for individual sample are tabulated in Table 4.3. The extracted contact resistance (R_C) and sheet resistance (R_{SH}) are on average 0.17(Ω.mm) and 169(Ω/□). Those figures are slightly worse than the values obtained with the thick Ohmic scheme, but are still acceptable considering the massive reduction of gold cost (~80%).

Device name		**Ti-gate**	**5nm Pd-gate**	**10nm Pd-gate**	**20nm Pd-gate**
R_C (Ω.mm)		0.12	0.15	0.16	0.14
R_{SH} (Ω/□)	**TLM**	163	173	159	178
	Hall	155			

Table 4.3 TLM measurement data for the four samples based on XMBE#131 (R_{SH} from the Hall data attached for comparison)

The gate recess was defined by a highly selective (InGaAs: InAlAs = 120:1) Succinic etchant. The Schottky contact with different metallisation schemes, mainly Ti/Au (Ti-gate) and Pd/Ti/Au (Pd-gate) were formed for each sample by thermal evaporation, as detailed in Table 4.4. The process is completed with bond pad formation by evaporating of Ti/Au = 50/450 nm. The high temperature Silicon Nitride (Si_3N_4) passivation step was skipped to prevent the Pd diffusion prior to the heat treatment studies.

Device name	Ti-gate	5nm Pd-gate	10nm Pd-gate	20nm Pd-gate
Pd (nm)	N/A	5	10	20
Ti (nm)	50	50	50	50
Au (nm)	450	450	450	450

Table 4.4 Gate metallisation scheme for the four samples based on XMBE#131

All samples were first annealed in a N_2-rich furnace for 5 minutes at 200˚C to promote the adhesion between the bond pads and underlying contacts. Then, the Pd-gate samples were further annealed at different temperatures (200˚C, 230˚C and 250˚C) for different time durations (30 minutes, 95 minutes and 335 minutes) to ensure Pd diffusion into the InAlAs supply layer. As illustrated in Figure 4.3, the effective gate-channel separation will be reduced from the nominal "d" of the Ti-gate into "d'" of the Pd-gate during the heat treatment and the changes of device characteristics as a function of Pd metal thickness, annealing temperature and annealing time are presented in the next sections.

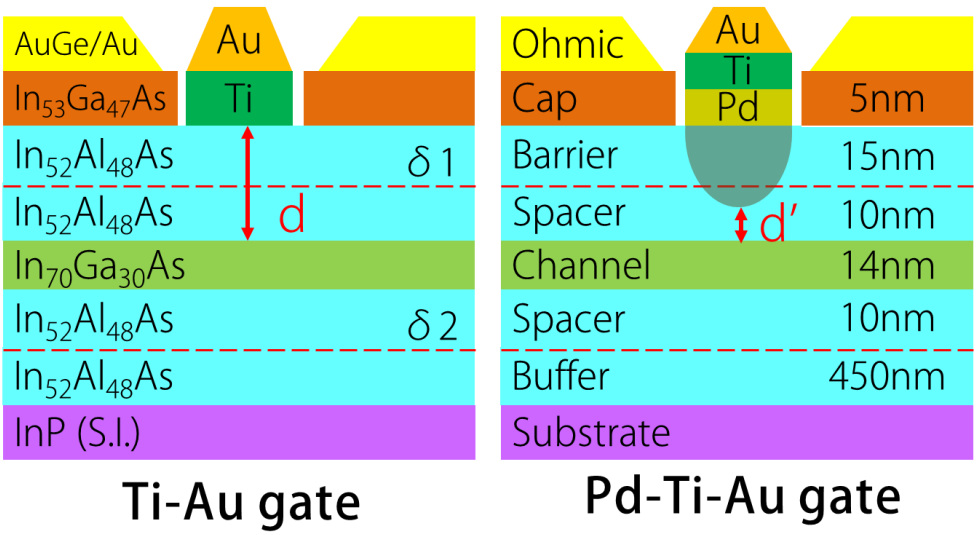


Figure 4.3 Change of effective gate-channel separation for the Ti-Au gate and Pd-Ti-Au gate devices after annealing

4.4 Results and discussions

DC and RF measurements were performed on 2x50 µm, double gate fingers device with 5 µm drain-source separation using a Cascade three fingers (G-S-G) prober, an HP 4142B DC parameter analyser and an Anritsu 37369A VNA. The RF calibration was done by LRRM standard. They were conducted at room temperature and all data presented were normalised with respect to the device gate width. The average results of the measured devices were presented. For simplicity, the device with Ti/Au metallisation scheme will be referred to as Ti-gate device and 5nm Pd-gate, 10nm Pd-gate and 20nm Pd-gate devices will denote those with Pd/Ti/Au schemes.

4.4.1 Effect of Palladium metal thickness

The aim of this study was to identify a suitable Palladium metal thickness for optimal device performance through the heat treatment study at fixed annealing temperature and time. All Pd-gate devices were annealed at 200˚C for 35 minutes, including the 5 minutes' bond pad heat treatment. Their Schottky diode characteristics are illustrated in Figure 4.4.

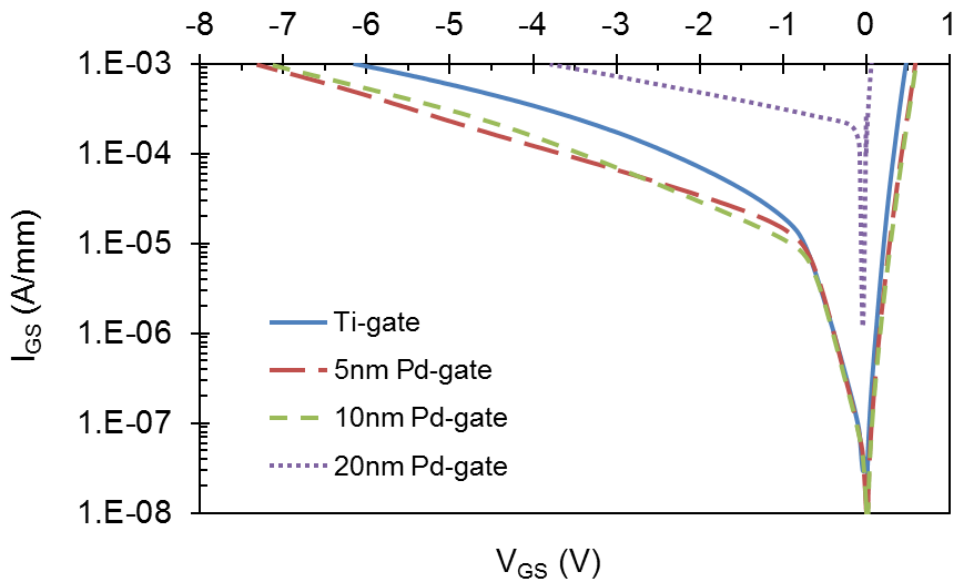


Figure 4.4 Schottky diode measurement for the Ti-gate (5minutes at 200˚C) and Pd-gate (35minutes at 200˚C) devices

The barrier height (Φ_B) and ideality factors (n) were extracted from the forward bias region. Ti-gate device shows Φ_B =0.56eV and n = 1.59 when compared with the 5nm and 10nm Pd-gate devices which have Φ_B =0.58eV and Φ_B =0.59eV and n = 1.58 and n = 1.57. The increased Φ_B and n for the Pd-gate devices are attributed to the break-down of native oxides on the InAlAs surface during the heat treatment [60]. Their off-state leakage is therefore suppressed with the higher Schottky barrier.

However, as the Pd metal thickness increased to 20nm, the device showed very poor Schottky behaviour, which has Φ_B =0.24eV and n = 2.83. The effective thickness of Pt after diffusion was reported to be up to ~1.4 times than that of the deposited metal thickness [55]. As Pd and Pt have close physical properties, it can be inferred that the effective diffusion thickness of 20nm Pd could be up to 28nm which is larger than the entire thickness of the barrier and spacer layer. The tunnelling current therefore increases rapidly with the excessive amount of Pd diffusion.

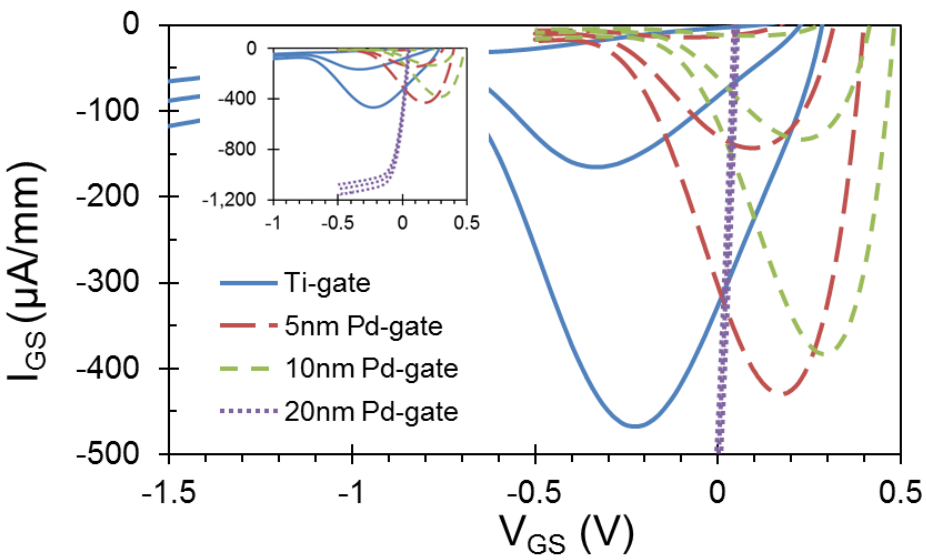


Figure 4.5 On-state leakage measurement for the Ti-gate (5minutes at 200˚C) and Pd-gate (35minutes at 200˚C) devices. (V_{DS} = 1 to 2V, step = 0.5V)

The effect of Pd diffusion is also observed on the on-state leakage measurement as depicted in Figure 4.5. With the increased thickness of Pd, the leakage due to impact ionization is gradually reduced because of the improved barrier height except that the

20nm Pd-gate device yields a much higher leakage current due to tunneling effects. This suggests that the thickness of Pd metal layer must be controlled to prevent the degradation of Schottky contact due to excessive gate-sinking.

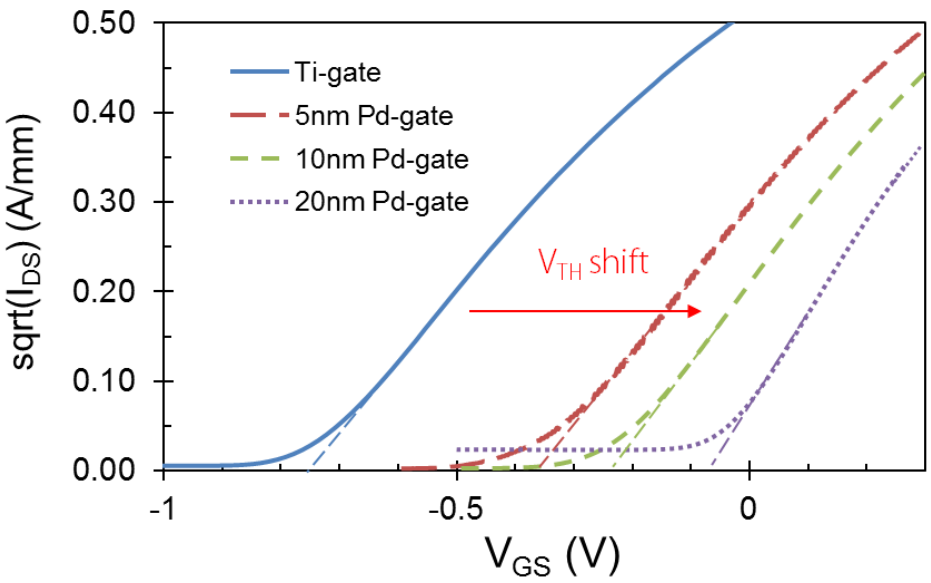


Figure 4.6 Square root of I_{DS} vs V_{GS} for the Ti-gate (5minutes at 200˚C) and Pd-gate (35minutes at 200˚C) devices. (V_{DS} = 1.0 V)

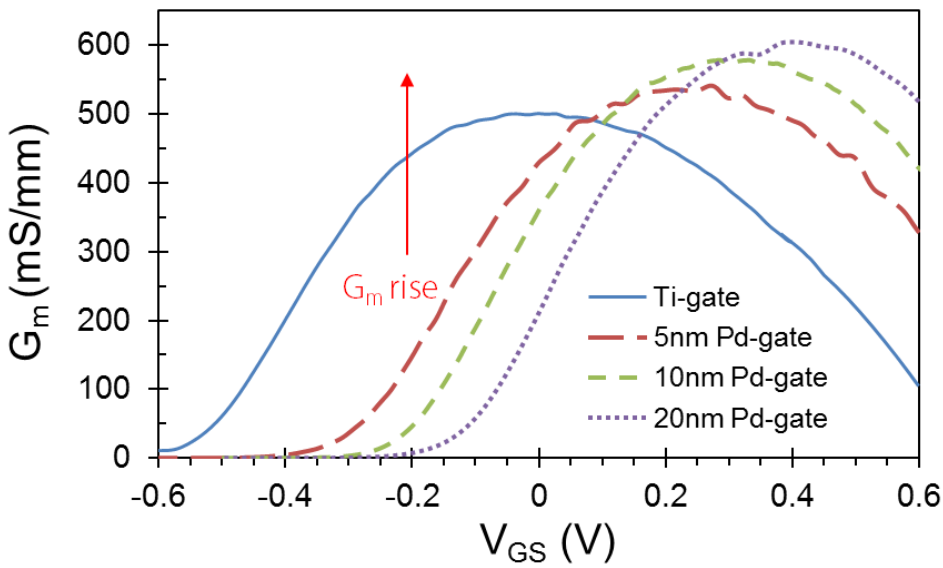


Figure 4.7 Gm vs V_{GS} for the Ti-gate (5 minutes at 200˚C) and Pd-gate (35minutes at 200˚C) devices. (V_{DS} = 1.0 V)

The change of threshold voltage (V_{TH}) and transconductance (Gm) are illustrated in Figure 4.6 and Figure 4.7 respectively. The reduction of effective gate-channel

separation (d) due to the Pd gate sinking initiates a positive V_{TH} shift and also improved Gm. By comparing the Ti-gate with the 5nm and 10nm Pd-gate devices, V_{TH} is shifted from -0.75V to -0.35V and -0.22V and peak Gm is increased from 500mS/mm to 540mS/mm and 580mS/mm respectively. These figures are pushed further to -0.08V and 610mS/mm on the 20nm Pd-gate device, which is essentially a quasi-enhancement-mode operation, but the channel current characteristic becomes worse with the increased Pd diffusion. For this reason, the discussion will be focused on the 5nm and 10nm Pd-gate structures.

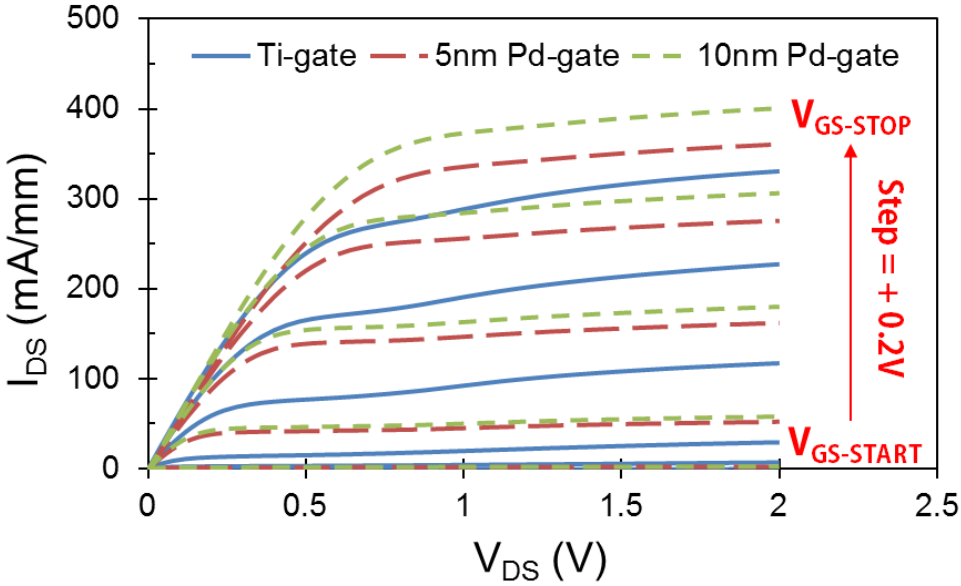


Figure 4.8 I_{DS} vs V_{DS} at different V_{GS} for the Ti-gate device (V_{GS} = -0.8 to 0.0V, step = + 0.2V), 5nm Pd-gate device (V_{GS} = –0.4 to + 0.4V, step = + 0.2V) and 10nm Pd-gate device (V_{GS} = –0.2 to +0.6V, step = + 0.2V)

The DC I-V characteristics for the Ti-gate, 5nm Pd-gate and 10nm Pd-gate devices are depicted in Figure 4.8. Considering the fact that each has different V_{TH}, the bias conditions were set from $V_{GS\text{-}START} = V_{TH}$ to $V_{GS\text{-}STOP} = V_{TH} + \Delta V$, where $\Delta V = 0.8V$ and step size = +0.2V, for meaningful making comparisons. With Pd gate sinking, the peak drain at V_{DS} = 2Vcurrent is increased from 330mA/mm to 360mA/mm and 400mA/mm, yielding 10 to 20% better current drivability in the Pd-gate devices. Moreover, the moderate kink-effect observed on the Ti-gate device is suppressed with Pd-gate metallisation. The improvement is attributed to the enhanced Schottky barrier

and the reduction of impact ionisation.

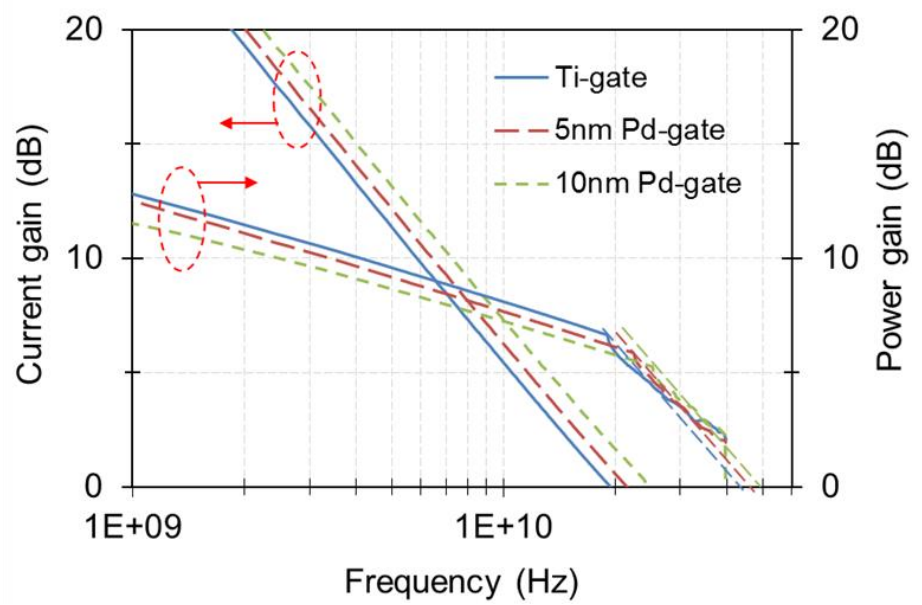


Figure 4.9 Frequency response of current and power gains for the Ti-gate and Pd-gate devices (V_{DS} = 1.0V) (Unity current gain and power gain are obtained by the 0dB crossing at x-axis)

Device summery – Gate width = 2x50 µm, drain-source spacing = 5 µm			
Parameters	**Ti-gate**	**5nm Pd-gate**	**10nm Pd-gate**
V_{DS} (V)	1.0	1.0	1.0
V_{GS} (mV)	0 ±5	+200 ±5	+300 ±5
Gm (mS/mm)	500 ±5	540 ±5	580 ±5
I_{DS} (mA)	28.5 ±1	25.3 ±1	19.8 ±1
f_T (GHz)	19.5 ±0.5	22 ±0.5	24.5 ±0.5
f_{max} (GHz)	42 ±0.5	46 ±0.5	49 ±0.5

Table 4.5 Bias conditions for the Ti-gate and Pd-gate devices

Microwave measurements on 2x50 µm devices with 5 µm drain-source separation were performed from 45MHz to 40GHz at V_{DS} = 1.0V and V_{GS} biasing at peak Gm. The measurement results and the corresponding bias conditions are shown in Figure 4.9 and Table 4.5 respectively. The measurement error is also attached. The cut-off frequencies (f_T) for the 5nm and 10nm Pd-gate devices are 22GHz and 24.5GHz and their corresponding and maximum frequency of oscillation (f_{max}) are 46GHz and 49GHz respectively. These figures are approximately 10 to 20% better than those of the Ti-gate device, which is a net result of the improved Gm and also the decreased intrinsic gate capacitances (C_{GS} and C_{GD}) [58].

4.4.2 Effect of heat treatment temperature and time

From the last study, the device with 10nm Pd metal is identified as the most balanced option between performance gain and gate leakage current. To probe further, a heat treatment study was performed on the 10nm Pd-gate device at 200˚C, 230˚C and 250˚C for 35minutes. Schottky measurements were performed before and after annealing. Figure 4.10 depicts the change of Schottky diode characteristics as a function of different annealing temperatures. At 200˚C and 230˚C, the device shows Φ_B = 0.58eV and 0.58eV, n = 1.58 and 1.57 which are very similar to the device before annealing with Φ_B = 0.57eV and n = 1.60, but the leakage current increases slightly due to the Pd-gate sinking. When the annealing temperature increases to 250˚C, the leakage due to tunnelling ramps up significantly with Φ_B = 0.35eV and n = 2.13. At higher temperature, the upper Ti and Au layers will start inter-diffusing with each other and therefore causes serious degradation of the Schottky contact [64]. Nevertheless, this issue would not affect normal device operation which is usually below 200˚C.

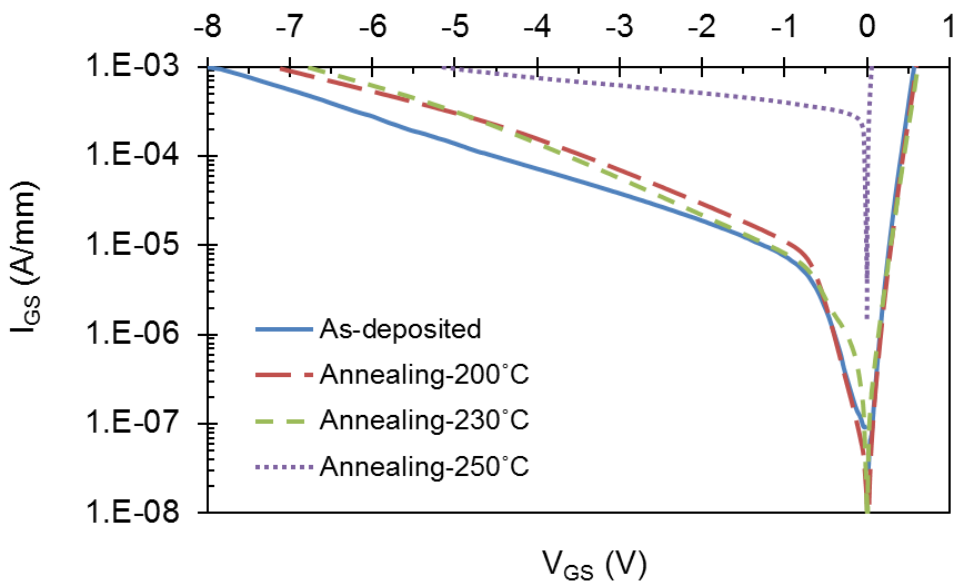


Figure 4.10 Change of Schottky diode characteristic as a function of annealing temperature for 10nm Pd-gate device (Annealing time = 35minutes)

The thermal stability of the 10nm Pd-gate device is accessed through the long heat

treatment at 200˚C and 230˚C for 340 minutes. DC and RF characterisations were conducted with the same configuration through the entire annealing stage, (before annealing, after 5 minutes, 35 minutes, 100 minutes and 340 minutes) to observe the change of device parameters during the annealing process.

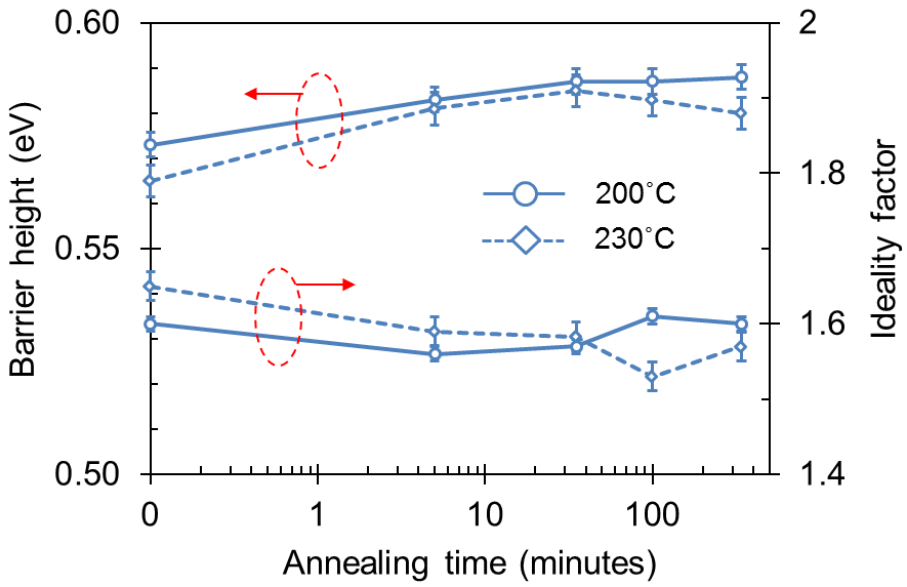


Figure 4.11 Change of barrier height and ideality factor as a function of annealing time in log scale for a 10nm Pd-gate device at 200˚C and 230˚C. (Error bar indicated)

Figure 4.11 depicts the change of Φ_B and n as a function of annealing time at different temperatures. In general, both parameters show no significant change during the heat treatment and the major improvement happens just after 5 minutes annealing at both temperatures. This observation is in agreement with the findings in [59]. The formation of Pd compounds during the annealing process acts as a diffusion barrier which stops further gate-sinking and also stabilises the Schottky behaviour.

The Gm and V_{TH} measurements also show similar trends as shown in Figure 4.12. The majority of V_{TH} shift and Gm enhancement occurs after 5 minutes heat treatment at both 200˚C and 230˚C and saturation occurs after 35 minutes annealing. After 340 minutes heat treatment, the device shows a V_{TH} shift of 0.2V and 0.25V with Gm improvement of 14% and 23% at 200˚C and 230˚C. This suggests that the magnitude of V_{TH} shift and increased Gm is proportional to the annealing temperatures, which

enables to tune the device parameters by changing the heat treatment conditions.

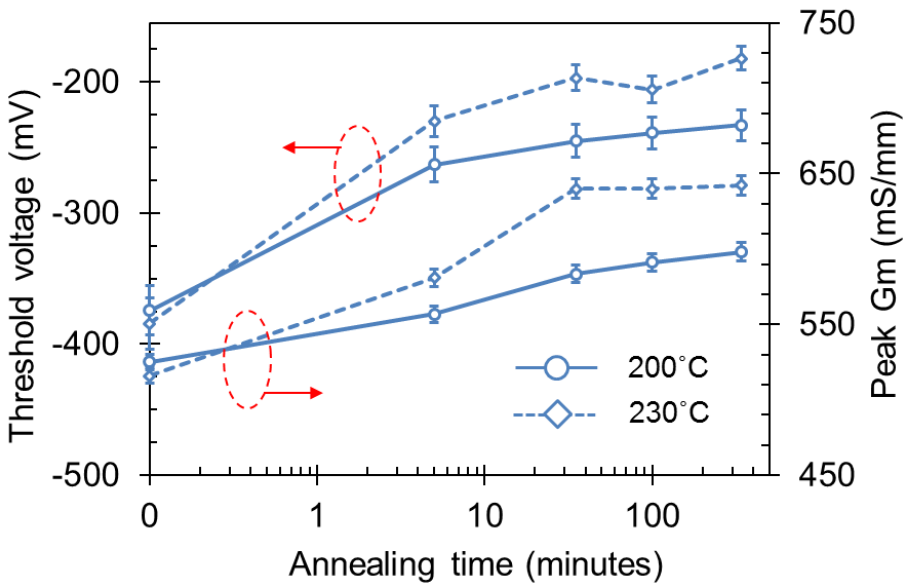


Figure 4.12 Change of transconductance and threshold voltage as a function of annealing time in log scale for a 10nm Pd-gate device at 200˚C and 230˚C (V_{DS} = 1.0V) (Error bar indicated)

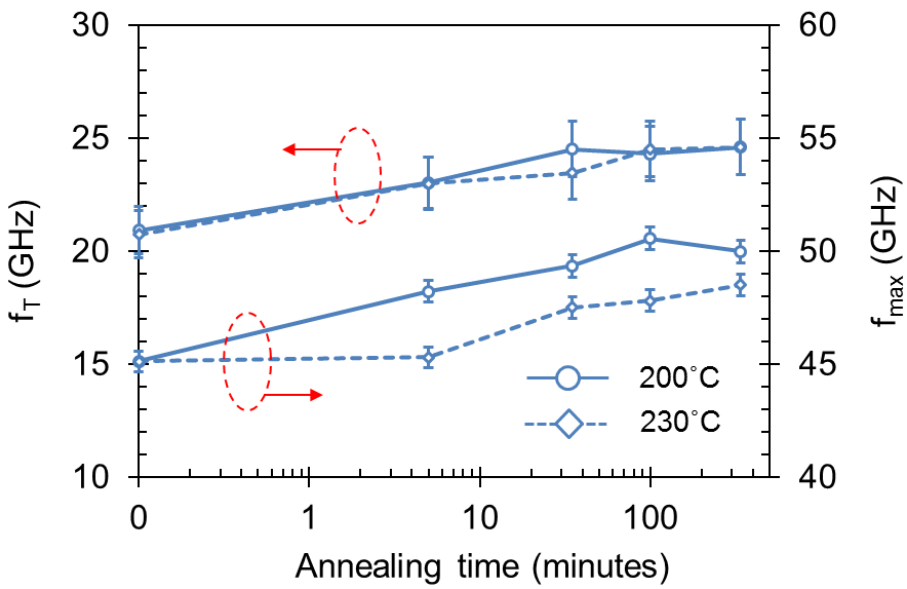


Figure 4.13 Change of unity current gain and power gain as a function of annealing time in log scale for a 10nm Pd-gate device at 200˚C and 230˚C (V_{DS} = 1.0V) (Error bar indicated)

The change of f_T and f_{max} in response to the annealing time at 200˚C and 230˚C are depicted in Figure 4.13. The saturation of these parameters is also observed after 35 minutes annealing. After the full heat treatment, f_T and f_{max} are 24.6GHz and 50GHz for 200˚C, and 25GHz and 48.5GHz for 230˚C. It turns out that the annealling

temperature makes no contribution to the absolute values and also the degree of improvement in f_T and f_{max}.

4.5 Summary

In this chapter, a comprehensive analysis was carried out on strained InGaAs-InAlAs pHEMT using Pd buried gates. Different parameters such as gate metallisation (Ti/Au and Pd/Ti/Au), Pd gate metal thicknesses (5nm, 10nm and 20nm), annealing temperatures (200℃ and 230℃), and annealing time duration (5 minutes to 340 minutes) have been investigated to examine the characteristics of Pd-gate devices.

Compared with conventional Ti-gate metallisation, the Pd-gate metallisation introduces a thermally stable shift of threshold voltage and retains a high quality Schottky contact. This new scheme shows promise with respect to enhancement of DC and RF characteristics such as increased Gm and I_{DS}, enhanced f_T and f_{max}, and elimination of kink effect.

By implementing a Pd-gate scheme using a cost effective thermal evaporation, the 10nm Pd-gate device exhibits improved Gm = 580mS/mm, high I_{DS} = 400mA/mm, and a f_T (f_{max}) of 24.5GHz (49GHz) after 35 minutes annealing at 200˚C. These Pd-gate devices demonstrate very stable DC and RF characteristics and no significant change were observed even after over 5 hours annealing at 230˚C.

5 DEVELOPMENT OF A NOVEL SOFT RELFOW PROCESS FOR THE FABRICATION OF SUBMICRON T-GATE STRAINED InGaAs-InAlAs pHEMTs

5.1 Introduction: The need for resist reflow

As stated by the well-known Moore's law [66], the number of transistor count in integrated circuits (IC) doubles every two years approximately, and this has been key in the production of cheaper, faster and more energy efficient consumer electronics. Over years, the continuous shrinkage of transistor gate dimension has introduced many technical challenges in the heart of IC manufacturing and especially lithography. With the advancements of the lithography techniques such as immersion lithography [67] and double patterning [68], microprocessors with transistor gate length of 22nm to 28nm are commercially available nowadays [69] while the next-generation 14nm process is also scheduled in the near future [70]. With the scaling down of devices, the cost of lithography machines (scanner and stepper) is approaching 10s to 20s and higher millions of US dollar (USD) price tag [71]-[72], which makes their use unique to the top semiconductor manufacturers like Intel Corporation and Taiwan Semiconductor Manufacturing Company, Limited (TSMC).

For academia, the acquisition of a toolset is often a compromise between cost and performance, due to their limited budget and cost of ownership. Table 5.1 outlines the strength and weakness of the three major lithography tools for academic research purposes. Projection system with high precision DUV stepper, capable of processing over a hundred wafers per hour and producing feature down to 100nm, is rated as the first choice for productivity and resolution. With a price tag just under a million USD, the cost of refurbished system is steep but remained a feasible option for institutes or companies with sufficient funding. However, the exorbitant phase shifting mask is a

major obstacle for long term usage of this lithography tool.

Lithography type	Projection	E-beam	Contact	
System name	Canon FPA 3000 EX5 wafer stepper [75]	Raith PIONEER E-beam system [76]	SUSS MicroTec MJB4 contact mask aligner [77]	
Light source	KrF Excimer Laser (248nm)	E-beam	i-line (365nm)	
System cost (USD)	~700,000 (Used)	~600,000 (New)	~140,000 (new)	
Mask specifications	6 inch, E-beam, PSM, 4x reduction, 0.1 µm	N/A	5 inch, E-beam, 0.5 µm	5 inch, laser, 1.0 µm
Mask cost (USD) [74]	20000 to 50000	0 (no mask)	~11000	~1000
Productivity (Wafer/hour)	~110	Very slow (< 1)	Operator dependent (< 20)	
Resolution (nm)	100	≤ 20	500	1000

Table 5.1 Comparison of different lithography options for research purpose (Price excluded installation, shipping charge and tax)

Because of its low cost, conventional contact lithography is preferred for the majority of research projects. A new contact mask aligner equipped with an i-line light source is available at a price of ~140,000 USD, which is less than a quarter of the price of a used stepper. The cost of photomask varies from ~1000 to ~10000 USD depending on the required resolution, which is relatively cheap compared with the stepper option. However, the only problem of the contact mask aligner is its inability to resolve features smaller than 500nm, which hinders its application below this limit. Electron beam (E-beam) lithography, with the advantages of excellent resolution and cost saving from the mask-less process, is widely employed for the fabrication of nanoscale devices, in particular HEMTs. Devices with sub-100nm gate length have shown f_T and f_{max} over 200GHz [78]-[80] and even higher with sub-50nm gate devices, approaching the half THz regime [81]-[88]. The major drawback is the low throughput and sometime poor repeatability due to the nature of E-beam [73].

It turns out that all three options suffer from different problems such as high cost, low performance and poor efficiency. A simple technique, called thermal reflow patterning,

has been widely employed in different studies to address the problem of fine line lithography [92]-[96], though this concept originated from micro lens fabrication for the optoelectronics field [89]-[91]. The principle of the technique, as illustrated in Figure 5.1, is to modify a predefined photoresist pattern by baking it above the corresponding photoresist glass transition temperature (T_g).

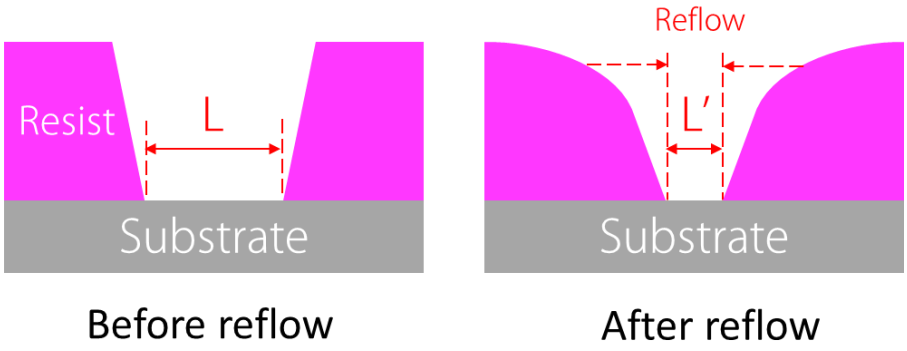


Figure 5.1 Schematic diagram before and after thermal reflow process

This process is also referred to as "hard reflow" due to the fact that the reflowed pattern is often hardened at such high temperature process. The final pattern length (L'), the initial pattern length (L) and the extended length (ΔL) are correlated through the following relation:

$$L' = L - \Delta L \qquad \textbf{Equation 5.1}$$

ΔL depends on the reflow temperature and time and L is defined by the lithography tools under use. The problem is that ΔL tends to be very limited (typically 0.1 to 0.4μm) even at high reflow temperatures in the works reported so far [92]-[96]. E-beam lithography must therefore be used to make L as small as possible, usually in the deca-nanometre regime [91]-[96], which implies that conventional thermal reflow loses its attraction. Furthermore, the absence of T-gate structures and the considerably extended undercut profile in the gate recess region would hinder the high frequency performance of devices [92].

In this work, we present a novel and economical approach for the fabrication of submicron T-gate structure by using a bi-layer soft reflow process with a solvent

vaporing technique at very low temperature. The advantage of soft reflow lies in its capability of fabricating nanoscale structures by using low-cost optical lithography with high throughput. By integrating the new soft reflow process with the Pd-gate metallisation scheme, the power of submicron device will be demonstrated through the fabrication of InGaAs-InAlAs pHEMT with a T-gate structure as small as 250nm, and which gives excellent and reproducible results.

5.2 Development of soft reflow process

5.2.1 Previous studies on the reflow technique

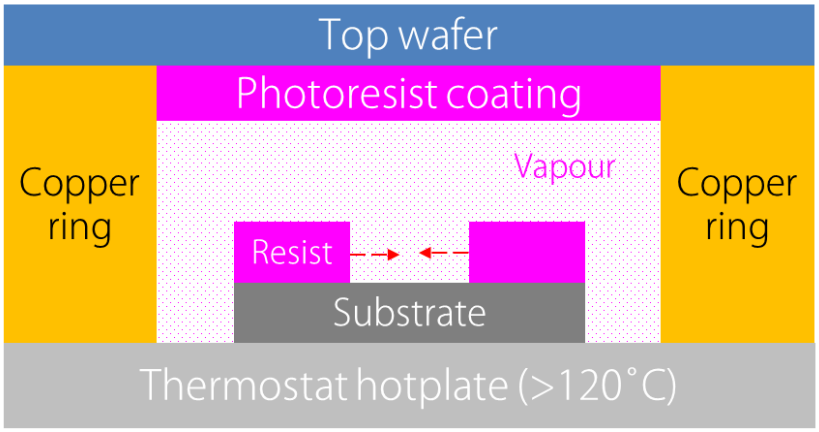


Figure 5.2 Schematic diagram of a previous reflow setup developed at Manchester

The idea of soft reflow was inspired by the works of previous co-workers, Dr. Sanae Boulay and Dr. Michael Exarchos. They combined the idea of thermal reflow and the principle of solvent absorption with a simple setup in an attempt to fabricate submicron structures. Figure 5.2 shows the details of the setup which consists of four components: a substrate with predefined photoresist pattern, a thermostat controlled hotplate, a copper ring and a wafer with photoresist coating. The hotplate acts as a direct heat source for the substrate to initiate the thermal reflow process. The top wafer, supported by the copper ring, is warmed up for generating solvent vapour in a closed atmosphere. This is then absorbed by the photoresist pattern and causes expansion of the photoresist profile. Therefore, the predefined pattern becomes

smaller due to the deformation effect of the photoresist melting and the solvent absorption.

However, this process was not successful for two reasons. Firstly, the top wafer must be coated individually to ensure a sufficient supply of vapour. The photoresist usage for large area coating is just too high in terms of cost. Secondly, the process is carried out at a rather high temperature (>120˚C), which is a potential hazard for the operator. The difficulty in removing the hard-baked photoresist is also a concern for efficient process. Therefore, a novel technique called soft reflow was developed to accommodate the requirement for efficient and reliable reflow process.

5.2.2 Principle of the new soft reflow technique

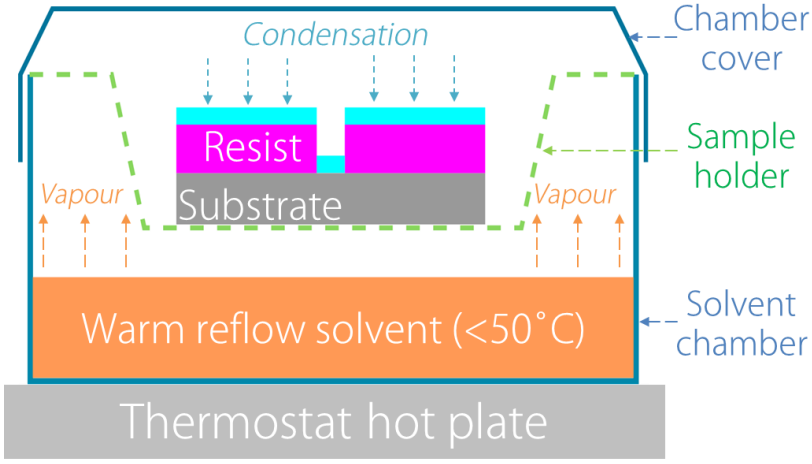


Figure 5.3 Schematic diagram of the solvent chamber for the soft reflow process

The principle of soft reflow process relies on the steaming effect in a specially designed solvent chamber as depicted in Figure 5.3. It consists of two major components including a solvent chamber and a metal sample holder. The chamber is partially filled up with a suitable solvent heated near room temperature (<50˚C) at atmospheric pressure (~1Bar). The warm solvent is employed as a vapour source, which populates the entire chamber. The choice of reflow solvent depends on its evaporation rate and also the characteristics of the photoresist. The sample, coated with photoresist, is placed on a metal holder floating above the solvent. It acts as a vapour collector, which

absorbs the condensed vapour from the solvent atmosphere. The key idea of this design is to create a controlled environment, which allows for generating and absorbing the vapour in a simple and efficient way.

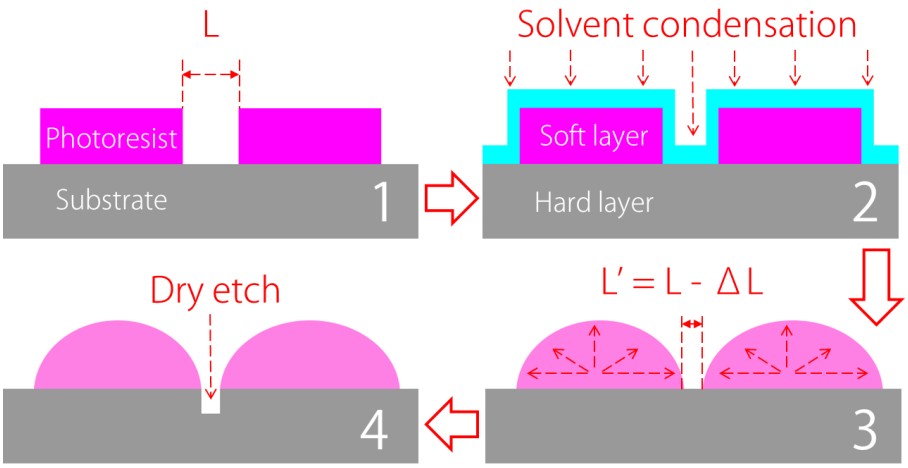


Figure 5.4 Principle of the soft reflow technique

The principle of soft reflow technique can be explained with a four-step process shown in Figure 5.4. The substrate, insoluble in the reflow solvent, is referred to as "hard layer" while the photoresist as "soft layer" by contrast. Firstly, a photoresist opening of length (L) is patterned on the substrate using conventional i-line optical lithography. The sample is then transferred into the reflow chamber and exposed to solvent vapour. A solvent coating then forms on the cool photoresist surface due to vapour condensation. To release the increasing stress from the solvent absorption, the photoresist increases its volume by isotopic expansion. Thus, the photoresist opening is reflowed from the initial length L into L' = L - ΔL. The amount of reflow (ΔL) is controlled by the reflow time. Finally, the pattern is transferred from the soft layer into the hard layer using dry etching.

The low temperature process (<50˚C) of the soft reflow enjoys several key advantages over the conventional hard reflow process. Firstly, unlike hard reflow; the low temperature operation is more process friendly as the removal of the reflowed pattern is easily performed using conventional photoresist strippers. Secondly, the hard reflow is

often restrained to devices on common semiconductor substrates (Silicon, Indium Phosphide and Gallium Arsenide, etc.) due to its high temperature operation. The proposed soft reflow could potentially be applied in other compatible substrates such as polymer based material for organic or thin film devices leading to potentially many new possible applications.

5.2.3 Controlling parameters of soft reflow process

The idea of soft reflow is very simple but the requirement for precise control over the reflow conditions remains a major challenge. In particular, the extended length (ΔL) and the movement of photoresist profile in the horizontal direction are highly sensitive to the changes of conditions. The study of those controlling parameters is the key to successful real-life application. The controlling parameters can be divided into three groups including solvent evaporation parameters, solvent absorption parameters and reflow time, as shown in Table 5.2. The study of solvent evaporation and absorption properties is very important for setting up a controllable process. Therefore, the following discussion will concentrate on the first two groups in this section.

Solvent evaporation parameters	**Solvent absorption parameters**	**Reflow time**
• Chamber pressure • Chamber temperature • Solvent properties	• Photoresist properties • Patterning area • Patterning thickness	

Table 5.2 List of soft reflow controlling parameters

Firstly, the solvent evaporation parameters, including the chamber pressure, the chamber temperature and the solvent material properties. They all affect the availability of solvent vapour generated in the chamber for a given photoresist and reflow time. The chamber pressure is maintained at atmospheric pressure (~1Bar) to avoid the use of vacuum pumping, making it a safer and, cheaper process with minimum maintenance.

Solvent	Acetone	Isopropyl Alcohol	Trichloroethylene	N-Methyl-2-pyrrolidinone
Molecular formula	CH3H6O	C3H8O	CH2HCl3	C5H9NO
Vapour pressure (mbar at 20˚C)	247	57	77.3	0.7
Evaporation rate (Butyl Acetate = 1.0)	5.6	1.7	0.69	0.03
Flash point (˚C)	-20	12	32	91

Table 5.3 List of common solvents for the soft reflow application (Data source: [97])

The suitable solvent candidate must be economical and accessible. Consequently, a list of common solvents was investigated as shown on Table 5.3. The two main criteria for the selection of solvent are controllability and safety. The range of solvent temperature should be between 25 ℃ to 100 ℃ to avoid the effect of room temperature variation and also the potential for photoresist hardening. Therefore, the candidate must have an appropriate evaporation rate to ensure a sufficient supply of vapour during the process.

The reflow study was carried out to observe the change of photoresist profile in response to different solvent environments. A 1 μm gate pattern was first formed on Si substrates. Soft reflow process was performed with different solvents at room temperature for a fixed duration of 5 minutes. The changes of photoresist pattern before and after the reflow process are illustrated in Figure 5.5 A to D. The closure of gate finger was observed on samples using Acetone and Isopropyl Alcohol (IPA). The high evaporation rate of these volatile solvents degrades controllability of the reflow process. By contrast, the reflow process is more controllable with N-Methyl-2-pyrrolidinone (NMP) and Trichloroethylene (Trike), which is evident from the clear reflow patterns.

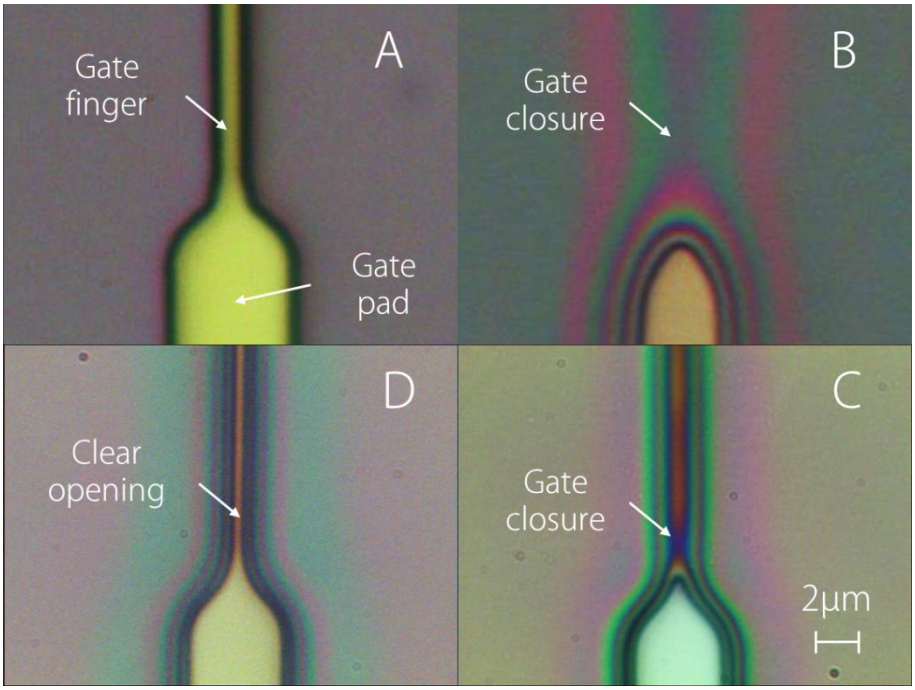


Figure 5.5 Change of 1μm gate pattern before reflow (A) and after 5 minutes in Acetone (B), IPA (C) and NMP (D). (S1805 resist coating on Si substrate) (Room temperature)

The flammable vapour generated during the reflow process is also a concern for safe operation. The candidate must have a high flash point to eliminate potential fire hazard. The flash points of Acetone and IPA are -20˚C and 12˚C respectively, which is lower than standard room temperature (25˚C), hence they are removed from the list. Trike is also eliminated as its low flashpoint (32˚C) leaves very small margins for process optimisation. NMP therefore is the most suitable solvent as it satisfies the requirement for safe operation and controllable process.

The solvent absorption parameters, including photoresist material properties, area and thickness of photoresist, are strongly photoresist dependent. They affect the solvent absorption ability of the photoresist profile for a given chamber condition and reflow time. In this study, a Shipley S1805 positive photoresist was chosen as the soft layer for its low cost and high availability. It has a nominal thickness of 500nm (4000rpm spinning), which helps to improve the accuracy of fine lithography and also reduces the formation of edge-bead issue, making it a suitable candidate for our 1μm patterning.

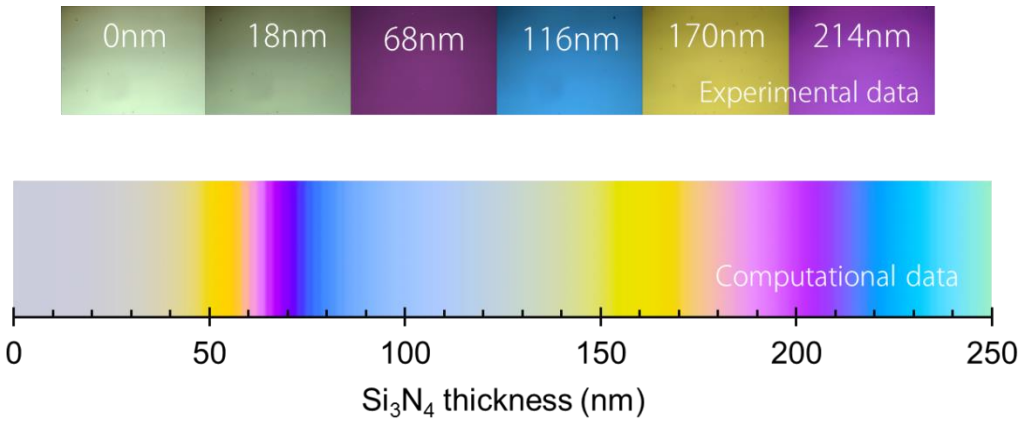


Figure 5.6 Silicon nitride colour vs film thickness on Silicon substrate. (Top: experimental data captured by digital microscope under fluorescent light at 90˚ incident angle) (Bottom: computational data generated by [98])

The characteristic of soft reflow process is mainly determined by the interface between the soft and hard layer. The later one will affect the quality of the transferred pattern. Silicon Nitride (Si_3N_4) was chosen as the hard layer for its unique colour property as shown in Figure 5.6. The Si_3N_4 colour could be varied from violet-blue (~200nm) to pale-brown (<30nm) under white light conditions due to the change of refractive index with film thickness, which is very useful for determining the film thickness. The high contrast between the common semiconductor substrates and the Si_3N_4 mask layer also helps to improve the measurement accuracy with digital microscopy.

The transfer of submicron pattern from the reflowed photoresist to the underlying Si_3N_4 relies on the dry etching process. The selectivity between soft and hard layers is a key factor in determining the quality of the resultant pattern. The dry etchant used was Tetrafluoromethane (CF_4) and was conducted on a Si substrate with S1805 photoresist and Si_3N_4 coating. The study was performed under conditions of low power and low pressure, in order to obtain anisotropic sidewall profile and minimise plasma damage.

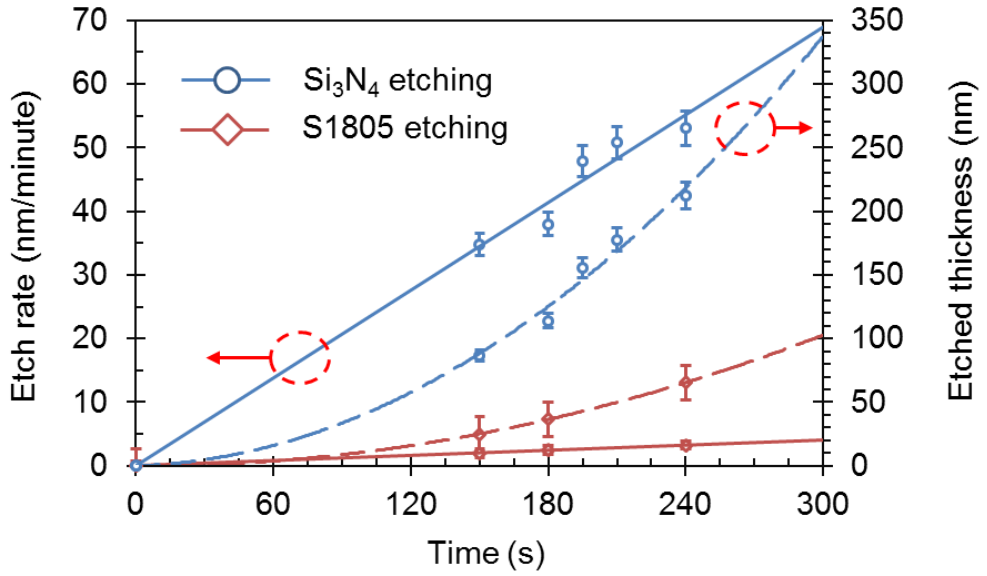


Figure 5.7 Characteristics of low power CF_4 dry etching on Si_3N_4 dielectric and S1805 photoresist on a Si substrate. (Gas: CF_4, pressure = 22 mTor, RF power = 20W, DC bias = 110V) (Solid line = etch rate, dot line = etched thickness) (Error bar indicated)

The change of etch rate and etched thickness as a function of etch time are depicted in Figure 5.7. The etch rate shows a linear relationship with the etching time for both materials, which indicates that the etch rate accelerates with longer etching process. The increment of etch rate on Si_3N_4 is significantly steeper than that on S1805 photoresist. The selectivity is defined by the ratio of etched thickness between Si_3N_4 and S1805 photoresist. In this case, the selectivity shows a constant ratio of 3:1 (Si_3N_4:S1805). The high selectivity between the proposed materials combination helps to minimise the problem of tapered sidewall profile during dry etch and enhances the resolution of transferred patterns.

5.2.4 Characteristics of soft reflow technique for submicron T-gate fabrication

5.2.4.1 Introduction

The most important parameter for the soft reflow technique is the reflow time because it controls the generation and absorption of the solvent vapour at a given chamber conditions. In this section, the relationship between the reflow time and the gate opening size is investigated with the controlling parameters previously obtained. The implementation of high performance mushroom gate structure relies on the optimisation of gate footprint. This is a milestone for the successful integration of soft reflow technique for the proposed submicron gate pHEMT fabrication.

5.2.4.2 Experiment

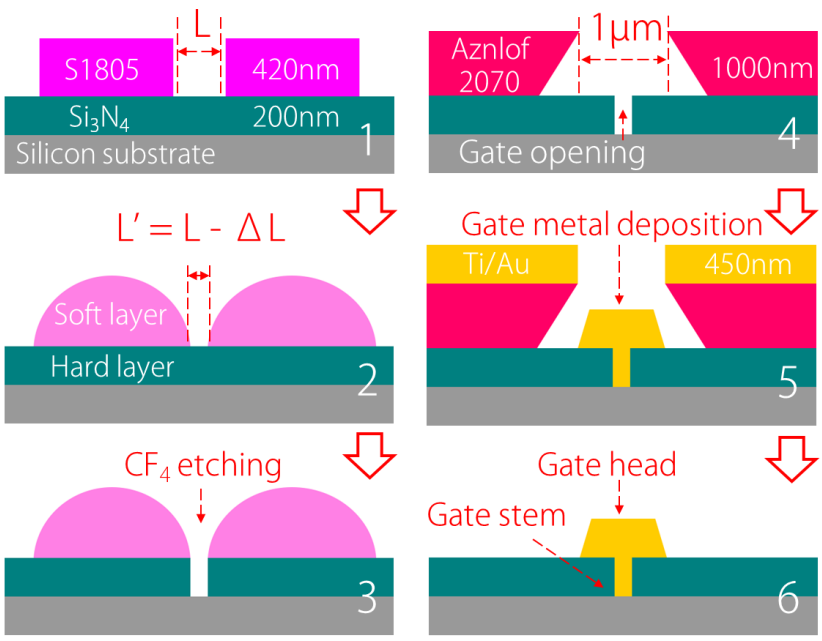


Figure 5.8 Submicron T-gate process flow using soft reflow technique

The feasibility of the soft reflow technique is evaluated through the process integration for the sub-micron T-gate fabrication, as depicted in Figure 5.8. A hard mask layer with 200 nm Si_3N_4 was initially formed on the Si substrate by plasma-enhanced

chemical vapor deposition (PECVD). Gate patterns with different sizes (1, 2 and 3 μm) were patterned on the Si_3N_4 using conventional i-line lithography with S1805 positive photoresist. The formation of submicron pattern was carried out in the reflow chamber at atmospheric pressure, with NMP solvent heating at low temperature (<50˚C). Subsequently, the formation of gate footprint was obtained by CF_4 plasma etching, which transfers the reflowed pattern from the photoresist to the Si_3N_4 mask. Upon the removal of S1805 soft layer by NMP, the 1 μm gate head was created by employing MicroChem Aznlof 2070 negative photoresist, which exhibits an undercut profile after development. The gate metal (Ti/Au = 50/450nm) was deposited by thermal evaporation followed by lift-off to form the T-gate metallisation.

5.2.4.3 Results and discussions

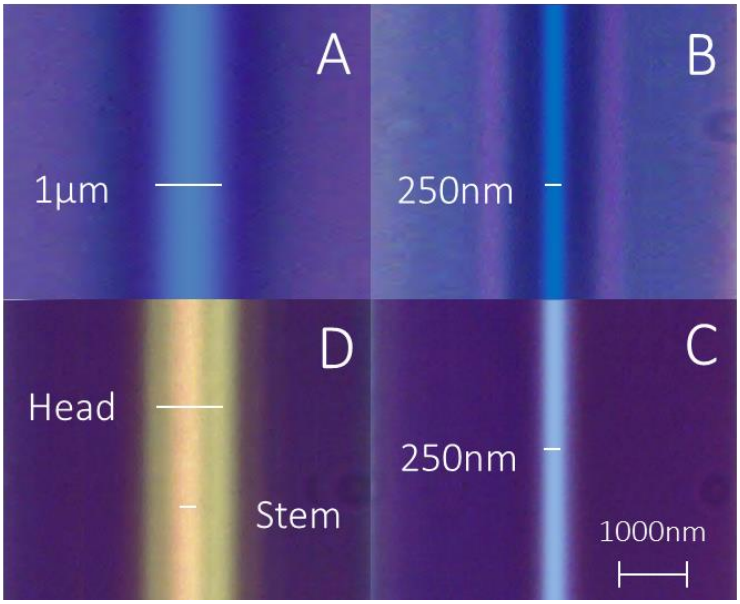


Figure 5.9 Digital microscope images of a 1μm gate opening after development (A), after 90s reflow (B), after CF4 etching (C) and after T-gate deposition (D).

Digital imaging and measurement were performed at different stages for monitoring purposes. The process flow of a 1 μm gate opening with 90s reflow time is shown in Figure 5.9A to 5.9D. The 1 μm gate opening is clearly illustrated in Figure 5.9A after development; the bold line pattern near the edge represents the slope profile of the positive photoresist. After the soft reflow process, the photoresist profile moves

inwards and the resulting opening size is reduced to about 250nm. The observation of the rainbow pattern around the edge is due to the expansion of the photoresist profile. As depicted in Figure 5.9C, the submicron gate footprint becomes more evident after the CF_4 pattern transfer onto the Si_3N_4 hard mask. Figure 5.9D illustrates the T-gate metallisation deposited on top of the original gate footprint. The grove pattern found in the middle of the gate is where the gate stem is located underneath the top gate head.

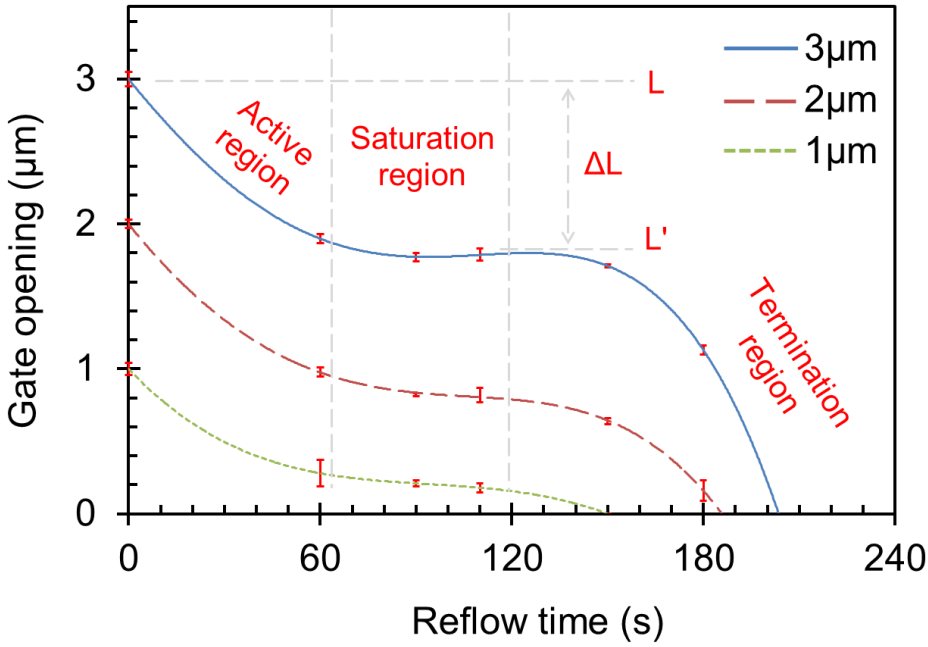


Figure 5.10 Relationship between gate opening size after dry etching and reflow time (Error bars indicated).

Figure 5.10 shows the changes of 1, 2 and 3-µm gate openings as a function of the reflow time. The individual gate opening sizes were measured after CF_4 etching step. The plot is categorised into three regions: the active region, the saturation region and the termination region. The active region is located within the first 60s, which shows rapid change in the gate opening size. From 60 to 120s, the curves reach the saturation region and the opening size becomes insensitive to the reflow time. Thereafter, the curves start falling again very quickly from 120s onwards. The closure of the gates eventually starts from 150s due to the excessive amount of reflow.

Notice that the three curves in Figure 5.10 share similar trends in all three regions, which highlights two important parameters: the width of the saturation region and the slope of the active region. For a given set of chamber and material parameters, the gate opening becomes smaller as the reflow time increases, but it might also lead to the closure of gate opening, which is undesirable for reliable device fabrication. In practice, there is a trade-off between feature size and process controllability. It is suggested that reflow time should be kept within the saturation region as labelled in Figure 5.10 for a wide process window. The width of the saturation region becomes a key factor for process controllability and reliability. As a result, the minimum size of the final gate opening (L') is limited by the magnitude of ΔL between the initial opening (L) and the saturation region. The slope of the active region, which restricts the amount of gate shrinkage in the saturation region, becomes another key factor for the process optimisation.

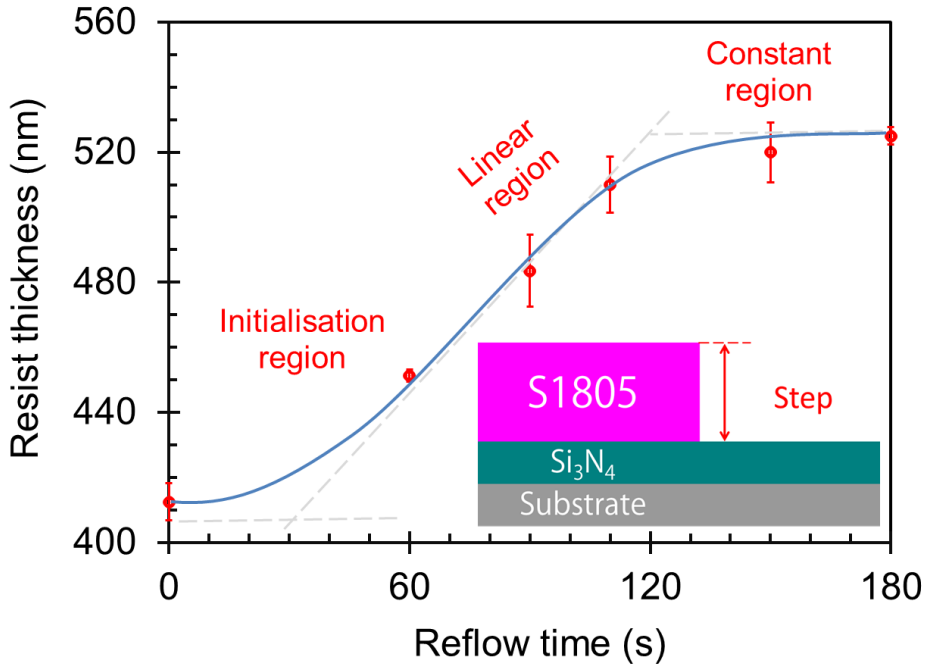


Figure 5.11 Relationship between S1805 photoresist thickness and reflow time (Error bars indicated).

Step measurements were performed on the photoresist profile using a Dektak 3 ST surface profiler. The change of photoresist thickness against reflow time is depicted in Figure 5.11, which shows a similar three regions model as in Figure 5.10. In the initialisation region, the change of thickness is very small and it increases proportionally to time in the linear region until the constant region, where the thickness becomes constant. The saturation of photoresist thickness indicates two possible mechanisms: initially, the photoresist tries to release the stress from the absorption of solvent by expanding in both horizontal and vertical directions. As the volume of photoresist reaches a certain critical point, which occurs at about 120s in our case, the gravitational force then help release the vertical stress and this increases the magnitude of horizontal expansion. As a result, the movement of photoresist profile increases tremendously after this point.

Concerning the mechanisms of the active and saturation region in Figure 5.10, it is necessary to clarify the correlation between Figure 5.10 and Figure 5.11. Figure 5.10 demonstrates the change of gate opening against reflow time, which corresponds to the change of photoresist profile in the horizontal direction. On the other hand, Figure 5.11 shows the corresponding change in the vertical direction, which implies the change of photoresist thickness. Each of these contains three different regions, which corresponds to different changes on the photoresist profile. In the first 60s, the photoresist profile expands mostly in the horizontal direction but not much in the vertical direction, which corresponds to the active and initialisation regions on the plots. From 60 to 120s, the photoresist profile grows thicker while the horizontal movement is insignificant, which corresponds to the saturation and linear regions on the plots. Starting from 120s onwards, the profile stops increasing in thickness but moves rapidly in the horizontal direction again, indicating the changes in the corresponding termination and constant region.

From this analysis, it is concluded that the horizontal expansion in the first 60s defines the slope of the active region and therefore the L' in the saturation region in Figure 5.10, while the next 60s defines the slope of the linear region in Figure 5.11 and more importantly the width of the saturation region in Figure 5.10. The key to achieving a smaller gate opening with a wider process window is to identify a suitable photoresist and solvent combination, which provides a maximum horizontal expansion in the active region and maximum vertical expansion in the linear region. A suggestion is to employ a thicker version of photoresist, such as Shipley S1813, which is capable of taking more solvent in a given reflow time and chamber conditions.

5.3 Fabrication of submicron T-gate pHEMT by soft reflow

5.3.1 Introduction

The fabrications and characterisations of submicron T-gate pHEMT are presented in this section, by combining the knowledge of submicron gate device described in chapter 3 with the reported soft reflow process. The benefit of reduced gate length will be revealed through the comparison between conventional 1 µm planar gate and reflowed submicron (350nm and 250nm) T-gate devices. Furthermore, the Pd-gate metallisation scheme from chapter 4 will be added to the simple soft reflow technique for optimal result.

5.3.2 Device fabrication

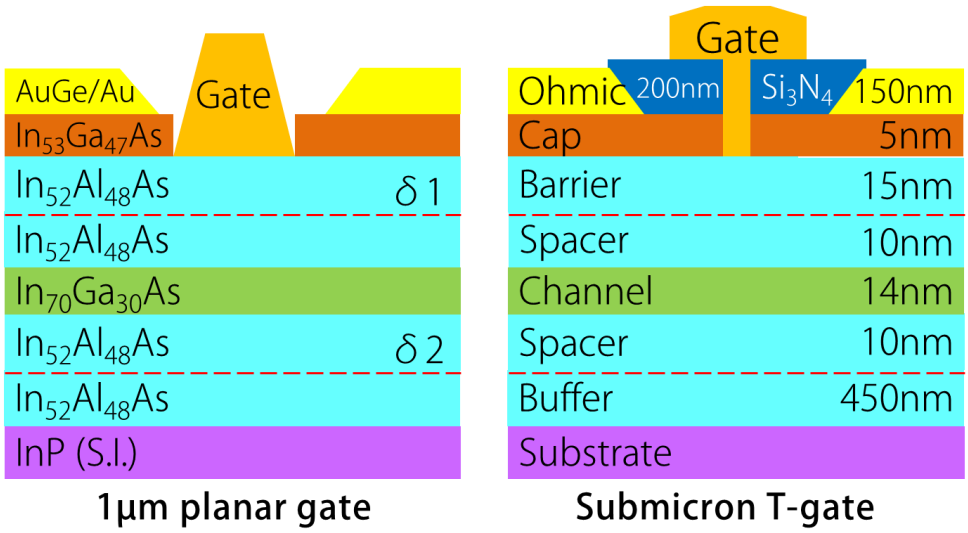


Figure 5.12 Schematic diagrams of 1µm planar gate pHEMT and submicron T-gate pHEMT.

In this study, four different devices (Ti-1 µm, Ti-350nm, Ti-250nm and Pd-250nm) were fabricated based on the same XMBE#131 epitaxial structures, as depicted in Figure 5.12, and which is identical to the one used in chapter 4. Its superior carrier transport properties is evident from the Hall measurements, with channel carrier concentration = 2.4×10^{12} cm^{-2} and mobility = ~13900 (cm^2/Vs) leading to a low sheet resistance (R_{SH}) = 155(Ω/□), which is the foundation for high speed device fabrication.

With a gate-channel distance (d) of 39nm, the shortest gate length (L_G) = 195nm can be achieved without significant short channel effect, by keeping a high aspect ratio (L_G/d) ≥ 5.

Device name		Ti-1μm	Ti-350nm	Ti-250nm	Pd-250nm
R_C (Ω.mm)		0.12	0.09	0.10	0.08
R_{SH} (Ω/□)	TLM	179	171	174	167
	Hall	155			

Table 5.4 TLM measurement data for the four different samples based on XMBE#131. (R_{SH} from the Hall data attached for comparison)

The device fabrication started with a conventional MESA isolation and a selective MESA side-wall etching as mentioned previously. Thin alloyed Ohmic contacts were formed by depositing AuGe/Au (50/100nm) on the $In_{0.53}Ga_{0.47}As$ cap layer followed by annealing in forming gas at 280˚C for 90s. As detailed on Table 5.4, their average contact resistance (R_C) and sheet resistance (R_{SH}) values, are 0.1 (Ω.mm) and 173 (Ω/□) respectively, at room temperature.

Device name	Ti-1μm	Ti-350nm	Ti-250nm	Pd-250nm
Gate geometry	Planar gate	T-gate	T-gate	T-gate
Gate length (nm)	1000	350	250	250
Gate metallisation	Ti/Au	Ti/Au	Ti/Au	Pd/Ti/Pd/Au
Metal thickness (nm)	50/450	50/450	50/450	10/50/10/450

Table 5.5 Details of gate metallisation for the four different samples.

The formation of Schottky contact is the most important step in the study. The difference between the 1μm planar gate and the submicron T-gate devices are illustrated in Figure 5.12 and the relevant parameters for the four samples are listed in Table 5.5. For the 1μm gate device (Ti-1μm), the gate opening was defined by i-line lithography with negative photoresist, followed by recess etching in a highly selective Succinic etchant (InGaAs: InAlAs = 120:1). A gate metal consisting of Ti/Au = 50/450nm was formed by thermal evaporation followed by lift off in hot NMP solvent.

For the submicron T-gate devices (Ti-350nm, Ti-250nm and Pd-250nm), a hard layer with Si_3N_4 = 200nm was first deposited on top of the cap layer. The 1μm gate opening,

as the soft layer, is then patterned on top of the Si_3N_4 by optical lithography. The gate opening was reduced from 1 µm to 350nm and 250nm by employing the soft reflow technique, as described in section 5.2. The submicron gate footprint was created by CF_4 pattern transfer from the reflowed photoresist to the Si_3N_4. The 1 µm gate head was patterned on top of the Si_3N_4 gate footprint using negative photoresist. Similarly, the gate recess profile was formed by using the highly selective Succinic etchant. For the Ti-350nm and Ti-250nm devices, the same gate metallisation scheme with Ti/Au = 50/400nm was employed as on the Ti-1 µm device.

For the Pd-250nm device, a multilayer metallisation scheme with Pd/Ti/Pd/Au was used for the formation of Schottky contact. The two Pd layers are introduced for different purposes. The middle one is designed to stop the inter-diffusion between Ti and Au, which was observed on chapter 4 while the bottom one is used as a diffusion source for the gate-sinking process. The device was further annealed in form gas at 200˚C for 30 minutes to initiate the Pd diffusion.

Finally, the devices fabrication was completed with the bonding pad deposition of Ti/Au = 50/400nm. Figure 5.13 shows a digital microscope and SEM pictures of the 250nm T-gate structure on the 2x50 µm device.

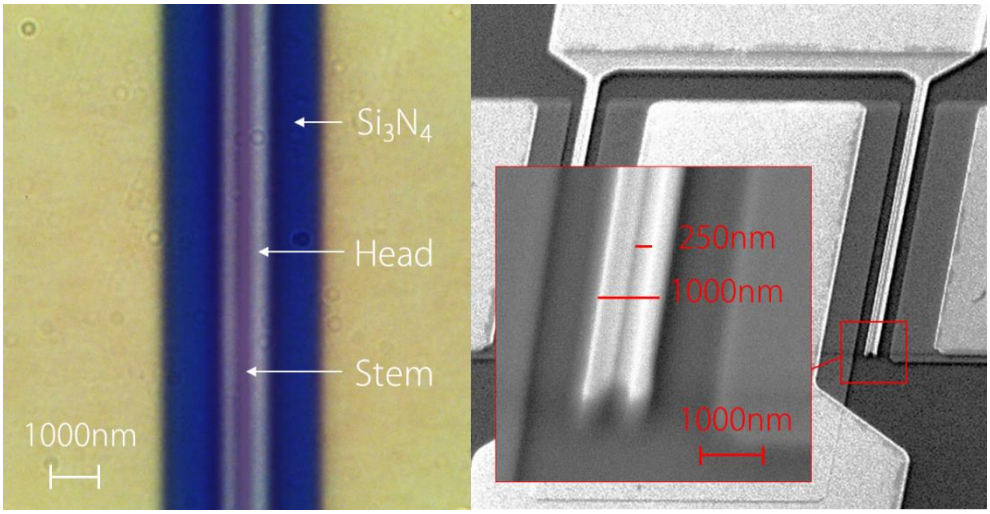
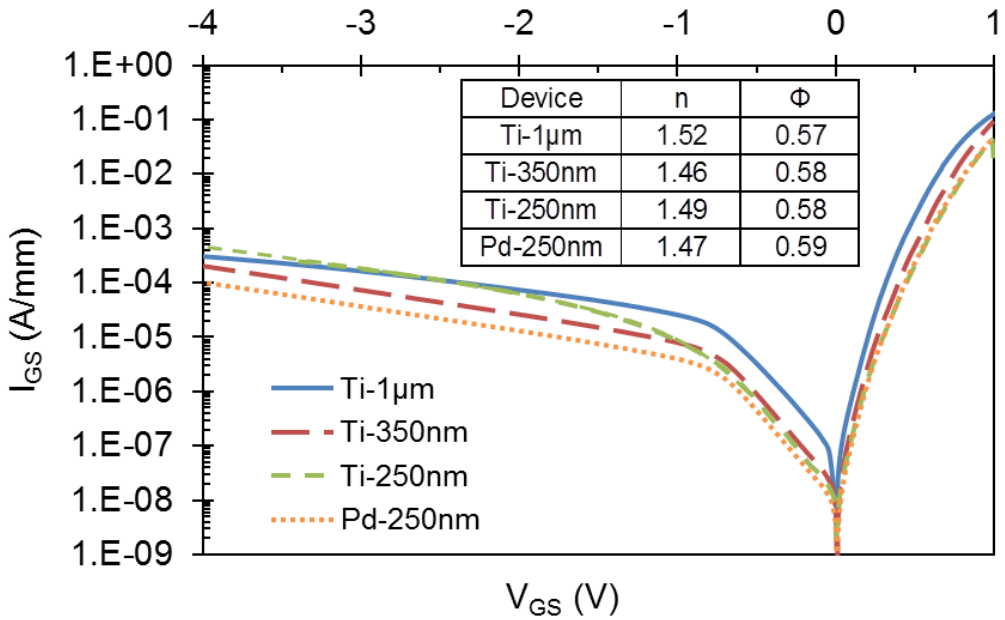


Figure 5.13 Digital microscope and SEM image of 250nm T-gate structure on the 2x50 µm device with 3 µm drain source spacing.

5.3.3 Results and discussions

DC and RF measurements were performed on 2x50 µm, double gate fingers device with 3 µm drain source separation at room temperature, using a Cascade three fingers (G-S-G) prober, an HP 4142B DC parameter analyser and an HP 8510C VNA. All data presented is normalised to the device gate width. The average results of the measured devices were presented.



Device	n	Φ
Ti-1µm	1.52	0.57
Ti-350nm	1.46	0.58
Ti-250nm	1.49	0.58
Pd-250nm	1.47	0.59

Figure 5.14 Schottky diode measurement of the 1 µm planar gate and submicron T-gate devices.

Schottky diode characteristics of the 1 µm planar gate and submicron T-gate devices are depicted in Figure 5.14. Their barrier height (Φ_B) and ideality factor (n) values, extracted from the forward bias region, are tabulated in the inset table. The submicron gate devices show similar characteristics to the 1 µm gate device, which implies that the shrinkage of gate length with the soft reflow technique does not alter the Schottky behaviour. However, the device with Pd metallisation scheme exhibits lower off-state leakage with better Φ_B and n, compared to the Ti-250nm device. The improvement shown using the Pd-gate device is attributed to the formation of Pd-compounds during the heat treatment process.

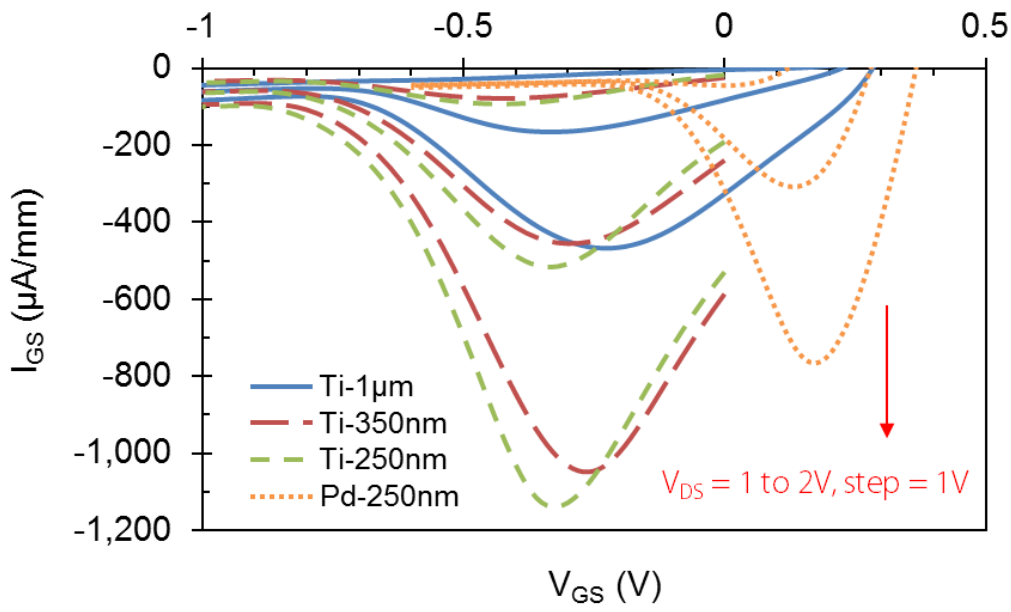


Figure 5.15 On-state leakage measurement of the 1µm planar gate and submicron T-gate devices. (V_{DS} = 1 to 2V, step = 0.5V)

Figure 5.15 shows the on-state leakage of different devices from V_{DS} = 1.0 to 2.0V. Compared with the 1µm gate device, the peak leakage current at V_{DS} = 2.0V is more than twice larger on the 350nm and 250nm device, increasing from -450µA/mm to -1050µA/mm and -1100µA/mm respectively. The massive reduction of L_G has increased the electric field in the active channel region, therefore promoting leakage due to impact on the submicron gate devices. By contrast though, the device with Pd-gate diffusion shows a reduction of 30% in its peak leakage current, compared with the ones with the same L_G but without a Pd metal layer. The integration of Pd metallisation scheme with the submicron gate structure shows effective suppression of the increased impact ionisation due to the higher effective barrier height effectively suppressing the movement of holes back into the gate region.

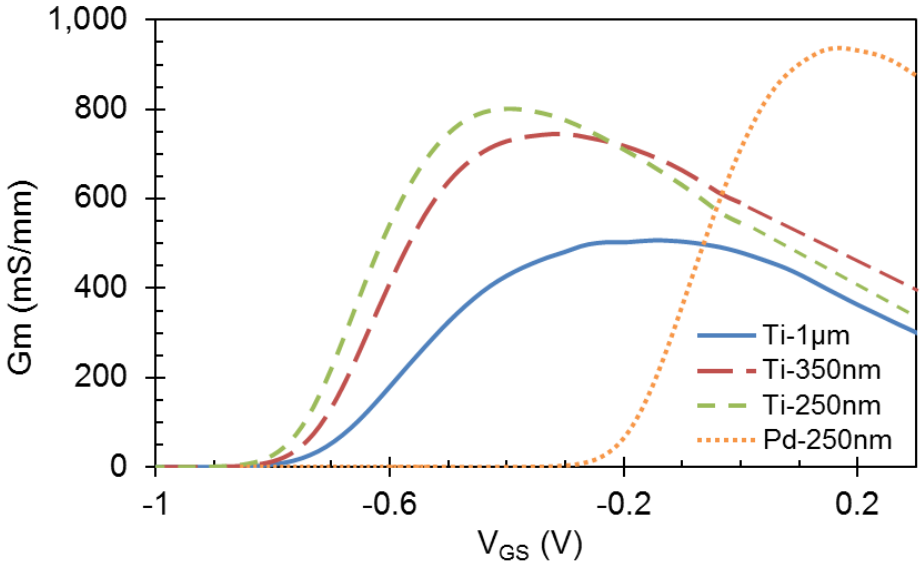


Figure 5.16 Gm vs V_{GS} of the 1 µm planar gate and submicron T-gate devices. (V_{DS} = 1.0 V)

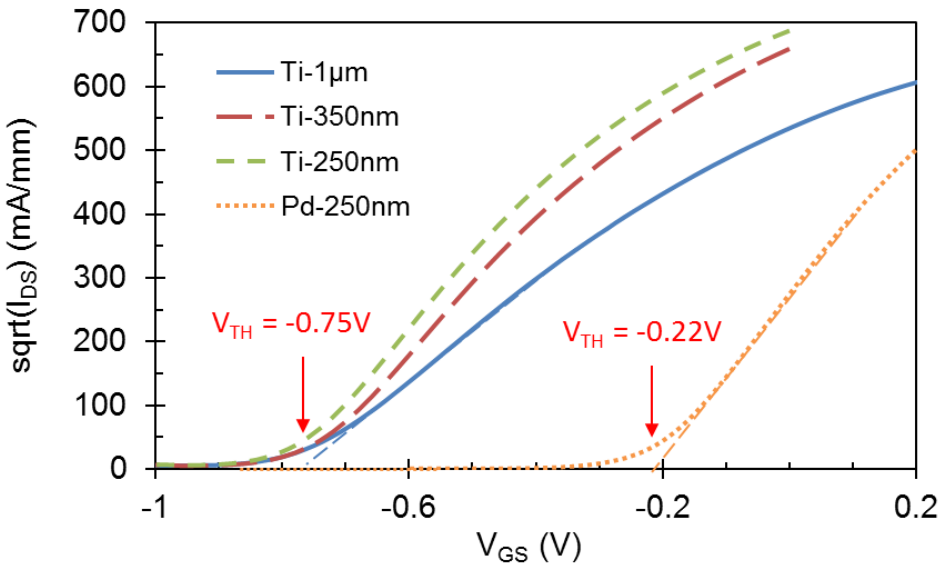


Figure 5.17 Square root of I_{DS} vs V_{GS} of the 1 µm planar gate and submicron T-gate devices. (V_{DS} = 1.0 V)

The change of transconductance (Gm) as a function of V_{GS} under biasing at V_{DS} = 1.0V for different devices is shown in Figure 5.16. By reducing the L_G from 1 µm to 350nm and 250nm, the peak Gm is enhanced by about 50 to 60%, increasing from 510mS/mm to 750mS/mm and 800mS/mm, which is attributed to the aggressive device scaling. With the additional Pd-gate sinking feature, Gm increases further to 940mS/mm due to the reduction of gate channel distance.

Figure 5.17 illustrates the change of drain-source current (I_{DS}) with different V_{GS} at V_{DS} = 1.0V for different devices. The current drivability is increased proportionally with the reduced L_G and more importantly, the submicron gate devices show no significant threshold voltage (V_{TH}) shift from the 1 µm gate device, which indicates the successful integration of the soft reflow process. For the Pd-gate device, the V_{TH} is shifted from -0.75V to -0.2V due to the Pd-gate sinking, which agrees well with the findings in chapter 4.

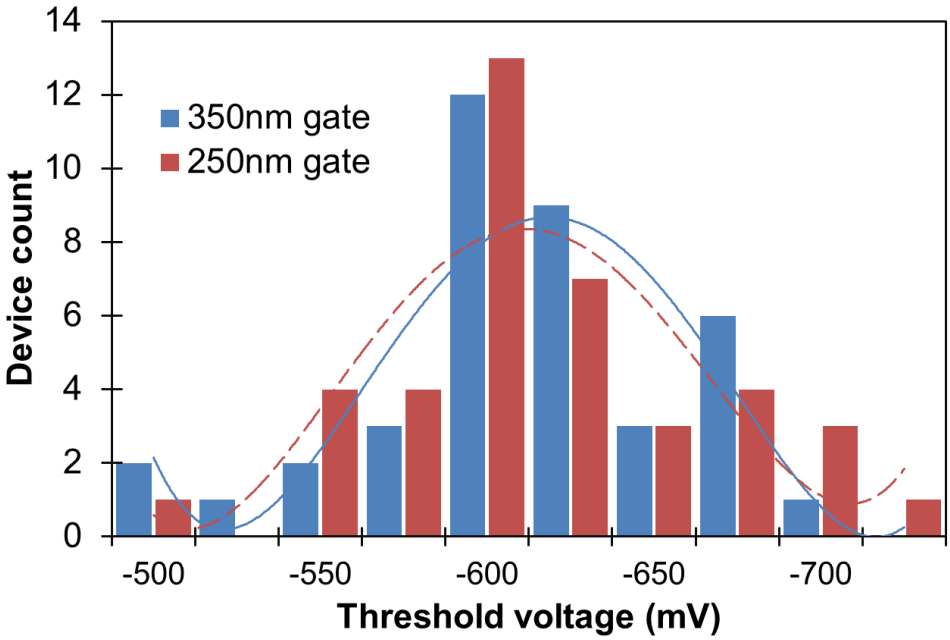


Figure 5.18 Threshold voltage distributions of the 350nm and 250nm T-gate devices. (V_{DS} = 1.0 V) (Total measured device for each type = 40)

Figure 5.18 shows the threshold voltage distribution of the devices with 350nm and 250nm T-gate structure respectively. Each type has a sample size of 40 devices measured on a wafer size of 15 x 15 mm^2. There are about 70% and 90% of devices showing a V_{TH} variation of 50mV and 100mV respectively. These figures are similar to those obtained in our conventional 1 µm gate process, attesting to the robust and highly uniform reflow process developed. The reliability of the soft reflow process is assured by the repeatable results and it demonstrates great potential for low cost mass production of submicron devices.

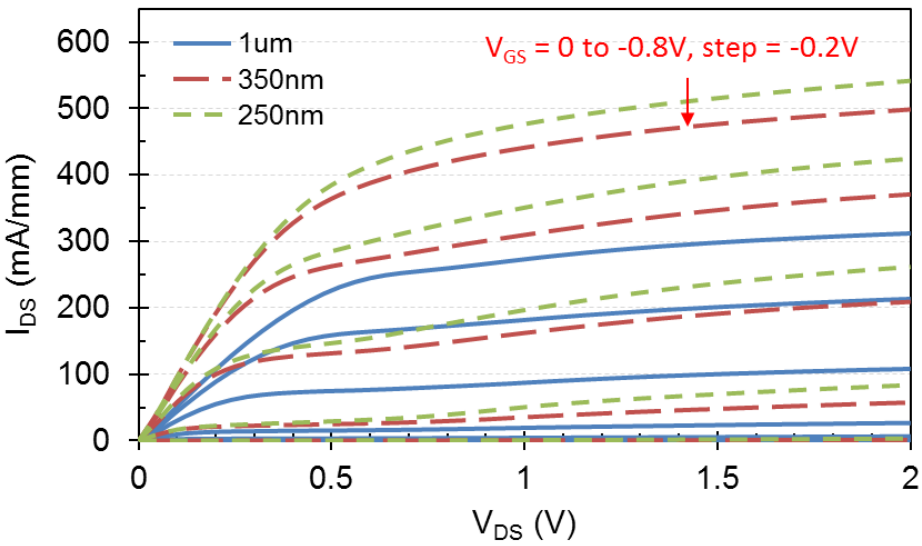


Figure 5.19 I_{DS} vs V_{DS} at different V_{GS} of the 1 µm planar gate device (V_{GS} = -0.8 to 0.0V, step = + 0.2V), 350nm T-gate device (V_{GS} = –0.8 to 0.0V, step = + 0.2V) and 250nm T-gate device (V_{GS} = –0.8 to 0.0V, step = + 0.2V).

A comparison between the output characteristics of the 1 µm planar gate, 350nm and 250nm T-gate devices is shown in Figure 5.19. Each device is biased under the same overdrive voltage (ΔV = +0.8V) from their corresponding V_{TH}. By reducing L_G from 1 µm to 350nm and 250nm, the maximum drain-source current ($I_{DS\ max}$) at V_{DS} = 2.0V is increased from 310mA/mm to 500mA/mm and 550mA/mm respectively. The current drivability of the submicron gate devices features 60% and 75% enhancement respectively, which is inversely proportional to the reduction of gate length as should be.

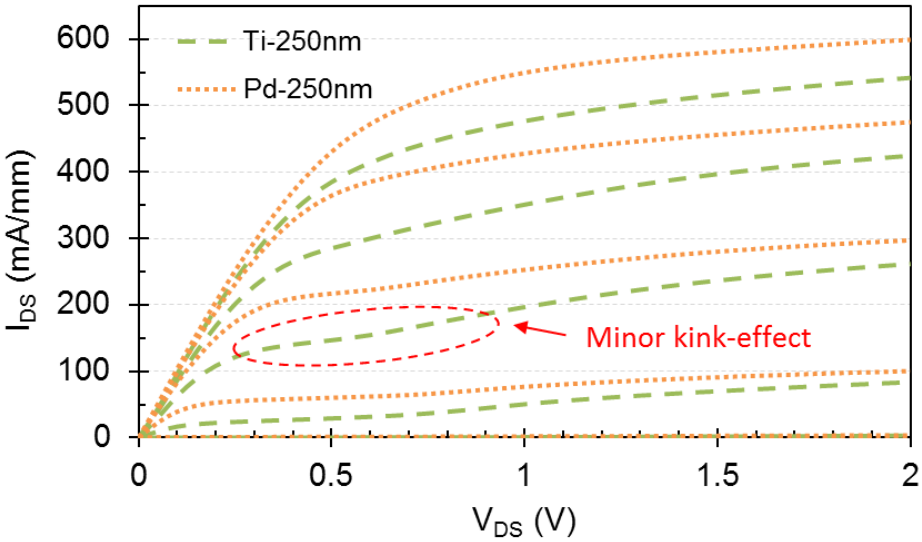


Figure 5.20 I_{DS} vs V_{DS} at different V_{GS} of the 250nm T-gate devices with Ti/Au (V_{GS} = -0.8 to 0.0V, step = + 0.2V) and Pd/Ti/Pd/Au (V_{GS} = -0.2 to +0.6V, step = + 0.2V) metallisation schemes.

The improvement of I_{DS} is also observed on the device with a Pd-buried gate structure, as depicted in Figure 5.20. The I_{DS} max of the 250nm Pd-gate device is increased by another 10%, from 550mA/mm to 600mA/mm, compared with the one with Ti/Au gate metallisation scheme. More importantly, the minor kink-effect observed on the submicron device was suppressed, which attributes to the increased barrier with the Pd-gate metallisation scheme.

The RF measurement was performed from 45MHz to 110GHz using the HP 8510C VNA with LRRM calibration at room temperature. Figure 5.21 and Figure 5.22 show the device current gain and power gain as a function of frequency. All devices were biased at their maximum Gm and their bias conditions are shown in Table 5.6. Their corresponding error tolerances are also labelled. For the devices with the Ti/Au metallisation scheme, the current gain cut-off frequency (F_T) of the 350nm and 250nm T-gate devices is 70GHz and 93GHz respectively, which are about 3 and 4.5 times that of the 1µm planar gate device. The improvement of F_T on the Pd-buried gate device is more significant, showing an excellent f_T = 108GHz.

Device summery – Size = 2x50µm, drain-source spacing = 3µm				
Device name	**Ti-1µm**	**Ti-350nm**	**Ti-250nm**	**Pd-250nm**
V_{DS} (V)	1.0	1.0	1.0	1.0
V_{GS} (mV)	-200 ±5	-400 ±5	-450 ±5	+100 ±5
Gm (mS/mm)	510 ±5	750 ±5	800 ±5	940 ±5
I_{DS} (mA)	18.5 ±1	15.5 ±1	14.5 ±1	15.8 ±1
f_T (GHz)	22 ±0.5	68 ±1	93 ±1	108 ±1
f_{max} (GHz)	42 ±0.5	112 ±1	123 ±1	131 ±1

Table 5.6 Bias conditions of the 1µm planar gate and submicron T-gate devices.

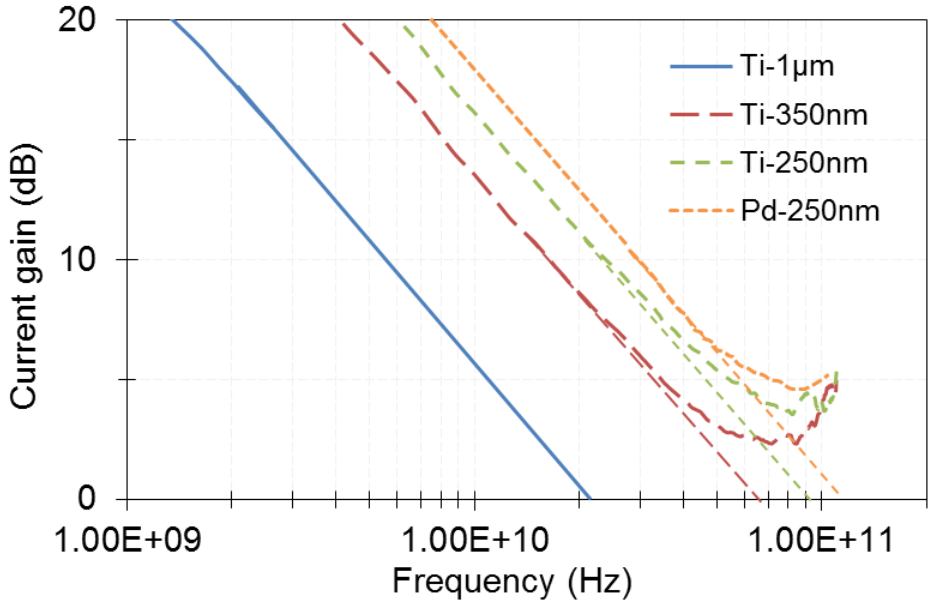


Figure 5.21 Frequency response of current gain of the 1µm planar gate and submicron T-gate devices. (V_{DS} = 1.0V) (Unity current gain is obtained by the 0dB crossing in x-axis)

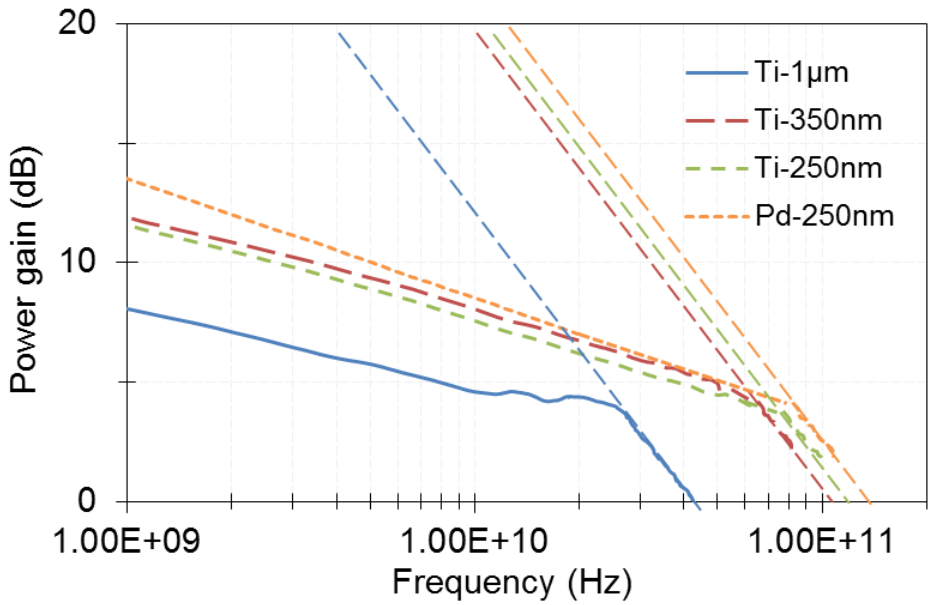


Figure 5.22 Frequency response of power gain of the 1μm planar gate and submicron T-gate devices. (V_{DS} = 1.0V) (Unity power gain is obtained by the 0dB crossing in x-axis)

The power gain cut-off frequency (F_{max}) on the submicron gate devices is also very impressive. The 350nm and 250nm gate devices have F_{max} = 112GHz and 123GHz respectively, and this is further improved to 131GHz on the Pd-buried gate device. Clearly, the implementation of T-gate structure on the submicron device has made a major contribution to suppressing the parasitic gate resistance (R_G), which is always a concern for nanoscale device fabrication and more importantly to the noise figure of low noise amplifiers employing these devices.

5.4 Comparison with other RF transistors

Reference		This work	[88]	[113]	[114]
Year		2013	2010	2010	2009
Technology		Ti-gate InP pHEMT	InAs pHEMT	HKMG RF CMOS	GaAs pHEMT
Substrate		InP	InP	Si	GaAs
L_G (μm)		0.25	0.03	0.03	0.25
Total gate width (μm)		100	100	30	200
V_{DS} (V)		1.0	0.5	1.0	3.0
I_{DS} (mA)		14.5	40	45.9	15
f_T (GHz)		93	644	395	55
f_{max} (GHz)		123	681	264	85[2]
NF_{min} (dB) (Sim./Meas.)	**2GHz**	0.13 Sim.[1]	N/A[3]	0.30 Meas.[2]	0.25 Meas.[2]
	10GHz	0.60 Sim.[1]	N/A[3]	0.70 Meas.[2]	1.00 Meas.[2]
Power (mW)		14.5	20	45.9	45.0

Table 5.7 Comparison between in-house fabricated InP pHEMT by reflow with other competitors in the commercial and research fields. (1. Simulated noise data from my co-worker Warsuzarina Mat Jubadi) (2. Extracted value from the original sources) (3. Parameter is not available)

In this section, a discussion will be carried out to compare the device performance between this work, the existing technologies and other advanced devices, as detailed on Table 5.7. The 250nm T-gate pHEMT with Pd/Ti/Pd/Au metallisation scheme is chosen from this work. The minimum noise figure, based on the same measured data was simulated by my co-worker, Warsuzarina Mat Jubadi.

pHEMT device based on the GaAs substrate is no doubt one of the most mature technology in RF transistors. The advantage of GaAs device is on the availability of large substrate (up to 150mm), which makes it relative economical for mass production and the packaged devices have been available on the market for several years. For example, the GaAs pHEMT with 0.25μm gate process is ready to order from TriQuint Semiconductor [114]. It provides a reasonable performance (f_T = 55GHz and f_{max} = 85GHz). However, its NF_{min} and power consumption are considerably worse than the InP option, which is mainly attributed to the poor

transport properties of GaAs based heterojuction.

The traditional Si industrial has shown a significant progress in competing with the III-V devices in the RF applications. The disadvantage of Si is gradually compromised by the aggressive scaling of transistor. The major semiconductor manufacturers such as Intel, has demonstrated Si-based RF CMOS with 30nm gate length [113]. Its f_T and f_{max} are much better than those in this work due to the reduction of L_G. However, the inherited issue of Si is still a major constraint on lowering NF_{min} and the power consumption. The benefit of RF CMOS is on its capability of providing SoC solution, which is why it becomes more popular for some lower frequencies applications such as mobile communication.

The shrinkage of gate length is even more beneficial on the InP pHEMT. By employing the 30nm gate structure and an InAs composite channel scheme, the InP device has shown a record high f_T = 644GHz and f_{max} = 681GHz from the Massachusetts Institute of Technology (MIT) [88]. It provides a good insight of InP device for the future development of high frequency RF electronics. However, the difficulty of fabricating nanoscale device on III-V is remained a major challenge due to the fact that the majority of the semiconductor industrials are based on Si. The significance of this work is to provide a well-balance solution in addressing the requirements for the low cost of infrastructure, highly efficient fabrication process and better device performance.

5.5 Summary

In this chapter, a novel soft reflow process based on the principle of solvent vaporing has been successfully developed. This new process provides a highly efficient and economical solution for the fabrication of submicron structures. Gate opening sizes down to 250nm are routinely and reproducibly achieved in this process. The opening size is controllable by changing the initial opening size, reflow time and reflow temperature. The saturation characteristic of the soft reflow provides a very controllable and repeatable process window for mass production.

The feasibility of the proposed soft reflow process has been demonstrated through the fabrication of a series of submicron T-gate devices. The fabricated devices show DC and RF characterises that are greatly improved with the shrinking of gate length. The devices are further improved through integration of the Pd-buried gate. This new soft reflow process shows great potential for high speed and low noise pHEMT fabrication and will be the heart of the next generation LNA designs at Manchester. This process can also be implemented on many other device applications such as Heterojunction Bipolar Transistors (HBT) and Resonant Tunnelling Diodes (RTD) and this indeed being pursued at Manchester.

6 CONCLUSIONS AND FURTHER WORK

6.1 Conclusions

The main idea of this research was to develop and fabricate ultra-high speed and ultra-low noise pHEMT for the next-generation SKA high frequency LNA designs. The work presented here shows great promise for future development of the SKA project and also other high frequency radio applications.

For years, the strained channel InGaAs-InAlAs pHEMT was the primary focus for SKA LNA developments due to its moderate performance with a simple 1 µm gate fabrication process. With the increasing demand for higher frequency and lower noise applications, the upper limit of 1 µm gate device became obvious as the drive towards ever lower noise figures is unabated both for SKA and microwave electronics in general.

The research work presented here started with a series of fabrication and characterisation of sub-100nm gate InGaAs/InAlAs/InP pHEMT using conventional E-beam lithography in collaboration with other research institutes. The negative impact of short-channel effect and the large parasitic elements was observed in the first generation devices with 120nm planar gate structure, which suffered from the issues of poor channel pinched-off characteristics, reduced transconductance (Gm) and degraded frequency responses (f_T and f_{max}). The suppression of short-channel effect was accomplished by employing thinner supply layer with wide band gap material, leading to the increased gate-channel aspect ratio ($L_G/d \geq 5$) and also preserved the low gate leakage. Furthermore, mushroom gate (T-gate) structure with reduced drain-source contact separation was introduced to eliminate the increased parasitic gate, source and drain resistance. With the optimisation of device geometry and epitaxial structure, the fabrication of 120nm T-gate device showed superior f_T = 163GHz and f_{max} = 182GHz, which illustrates the benefit of submicron gates while

still preserving the low leakage, high breakdown features of the Manchester InGaAs-InAlAs pHEMT process.

In addition to L_G reduction, the thermal stability of the Schottky gate is also an important concern for device operation under harsh environments. Therefore, investigation in the use of the refractory metal gate with Pd/Ti/Au metallisation scheme was undertaken. The characteristics of the developed Pd-buried gate has been exploited by exploring the relevant parameters such as Pd gate metal thicknesses, annealing temperatures and annealing time through a series of heat treatment studies. Compared with conventional Ti/Au metallisation, the Pd-gate scheme can offer a thermally stable shift of the threshold voltage and retains a high quality Schottky contact. The scheme also leads to improved DC and RF characteristics such as increased Gm and I_{DS}, enhanced f_T and fmax, and in particular a much reduced kink effect, which remains stable even after 5 hours annealing at 230˚C. The implementation of Pd-buried gate structure is effective in improving the device long-term reliability under high temperature operation, as well as reducing the cost of cooling high-speed circuitry and high-power circuitries such as PA and LNA.

The major achievement of this work is the development of a novel soft reflow technique presented in chapter 5. The key feature of this technology is its capability of resolving submicron patterns by taking advantage of conventional i-line optical lithography, and therefore providing an economical and effective solution of tackling the challenge of nano-scale device fabrication including complex 3D patters such as mushroom gates in pHEMTs. A comprehensive study was conducted to understand the mechanisms of soft reflow and the effect of relevant controlling parameters. The gate opening size is controllable by changing the initial opening size, reflow time and reflow temperature. A minimum gate opening of 250nm is achieved routinely and reproducibly due to its saturation characteristic, which provides a very controllable and repeatable process window for mass production.

The feasibility of the soft reflow process was proven through the fabrication of a series of submicron T-gate devices. Their DC and RF characterises are improved with the reduction of gate length. The devices with a 250nm T-gate structure showed excellent f_T = 93GHz and f_{max} = 123GHz at a very high yield (~90%). These figures were improved further with the integration of the Pd-buried gate, resulting in f_T = 108GHz and f_{max} = 131GHz, which is the highest values ever achieved with conventional optical lithography at this gate length. The fabrication of submicron T-gate pHEMT by implementing the soft reflow technique has helped to overcome the 1 μm gate process barrier, and resulted in noticeable improvement in the high frequency performances while still preserving all the desirable feature of the Manchester pHEMT materials, i.e. low leakage and high breakdown. This work will benefit the development of ultra-low noise LNA for SKA 2 and SKA3, and will a have an impact in the realisation of low cost submicron devices.

6.2 Further work

The fabrication of submicron gate devices, using with the new soft reflow technique, has laid the foundation for the next-generation LNA designs and also high frequency applications. The following 6 suggestions should permit the realisation of even more performing pHEMT devices:

1. The present devices' Ohmic contact showed a contact resistance (R_C) in between 0.09 to 0.20 (Ω/mm), which is significantly higher than the value reported in previous studies, typically with R_C <0.05 (Ω/mm) [88]. The implementation of heavily doped cap layer will help reducing R_C and hence the parasitics (R_S and R_D), resulting in better frequency responses.

2. The minimum drain-source separation (L_{SD}) was designed to be 3 μm in consideration of the difficulty of alignment between the Ohmic contacts and the gate fingers, which restricts the reduction of parasitics (R_S and R_D). With self-aligned Ohmic contact structure [86], L_{SD} could be reduced down to the

size of gate head (1 μm). The decreased L_{SD} will contribute to better DC and RF performances.

3. In current designs, the minimum achievable gate opening was restricted to 250nm to avoid the possible closure of gate finger with longer reflow time, which is closely related to the supply of solvent vapour. The reflow chamber should be redesigned with vapour regulating feature to enhance the process controllability down to the sub-100nm gate structures, aiming for higher f_T and better noise performance.

4. In the current reflow process, Si_3N_4 was chosen to be the hard layer for the formation of gate footprint. However, the Si_3N_4 deposition step is a very lengthy process. Other lower cost alternatives, such as Spin-on-Glass, are worthy of investigation to optimise the process flow [92].

5. The soft reflow technique can be used in a range of high frequency devices and indeed Resonant tunnelling devices with 1 x 0.35um contact areas and with impressive IV characteristics are currently being fabricated at Manchester. This process can further be implemented in HBT with even higher microwave performances.

6. A collection of submicron devices have been fabricated and characterised in this work by using the soft reflow technique. Device modelling and simulation should be carried out to evaluate the noise performance and also understand the properties of submicron devices, which will be beneficial for the development of LNA and other high speed circuits in the future.

REFERENCE

[1] C. Singer, “A short history of science to the nineteenth century” Dover Publications, pp. 217, 1997.

[2] Vanzella, E., et al. "The Great Observatories Origins Deep Survey." Astronomy and Astrophysics-Les Ulis 434.1 (2005): 53-66.

[3] R. N. Wilson, “Reflecting Telescope Optics I: Basic Design Theory and Its Historical Development”, Springer-Verlag, Berlin, pp. 302, 1996.

[4] W. Turner, A. J. Faulkner, B. W. Stappers, S. M. Ransom, R. L. Webster, R. P. Eatough, and M. Kramer, “High-level SKA signal processing description,” Signal processing CoDR document WP2-040.030, 2011.

[5] Kraus, John D. "Radio astronomy." New York: McGraw-Hill, 1966 1 (1966).

[6] Beck, Rainer. "Square kilometre array." Scholarpedia 5.3 (2010): 9321.

[7] P. E. Dewdney and T. J. Lazio, “The square kilometer array (SKA) radio telescope: Progress and technical directions,” Radio Science Bulletin, no. 326, p. 5, 2008.

[8] A. Faulkner, P. Alexander, A. Van-Ardenne, R. Bolton, J. Bregman, A. V. Es, M. Jones, D. Kant, S. Montebugnoli, and P. Picard, “Aperture Arrays for the SKA: the SKADS White Paper,” SKA2010, Mar, 2010.

[9] SKA telescope organisation, available at: http://www.skatelescope.org

[10] R. P. Millenaar and R. Bolton, “SKA Configurations Design: Status January 2010,” document WP3-050.020. 000-R-001, Rev, vol. 1, pp. 2010–2, 2010.

[11] Prof. M. Missous, “School of EEE: Prof Mohamed Missous - personal details”, available at: http://www.eee.manchester.ac.uk/aboutus/staff/show.html?ea=m.missous

[12] Angeliki Bouloukou, “Novel, High-breakdown, Low-Noise InGaAs-InAlAs Transistors for Radio Astronomy Applications”, PhD thesis, The University of Manchester, 2006.

[13] Ayman Sobih, “MMIC Broadband Low Noise Amplifiers for the Square Kilometre Array Radio Telescope”, PhD thesis, The University of Manchester, 2007.

[14] Shadzad Arshad, “Low-Power Wideband InP-Based Low Noise Amplifiers For The Sqaure Kilometre Array Radio Telescope”, PhD thesis, The University of Manchester, 2009.

[15] Muammar Mohamad Isa, “Low Noise Amplifiers Using Highly Strained InGaAs/InAlAs/InP pHEMT For Implementation In The Square Kilometre Array (SKA)”, PhD thesis, The University of Manchester, 2012.

[16] Norhawati Ahmad, “Modelling and Design of Low Noise Amplifiers Using Strained InGaAs/InALAs/InP pHEMT For the Square Kilometre Array (SKA) Application”, PhD thesis, The University of Manchester, 2012.

[17] A. G. Baca and C. I. H. Ashby, “Fabrication of GaAs devices,”, Institution of Electrical Engineers, 2005.

[18] F. Schwierz and J. J. Liou, “Semiconductor devices for RF applications: evolution and current status,” Microelectronics Reliability, vol. 41, no. 2, pp. 145–168, 2001.

[19] R. Pillarisetty, “Academic and industry research progress in germanium nanodevices,” Nature, vol. 479, no. 7373, pp. 324–328, 2011.

[20] R. L. Ross, S. P. Svensson, and P. Lugli, Pseudomorphic HEMT technology and applications: Springer, 1996.

[21] Feng, Milton, et al. "Device technologies for RF front-end circuits in next-generation wireless communications." Proceedings of the IEEE 92.2 (2004): 354-375.

[22] J. D. Cressler, “SiGe HBT technology: a new contender for Si-based RF and microwave circuit applications,” Microwave Theory and Techniques, IEEE Transactions on, vol. 46, no. 5, pp. 572–589, 1998.

[23] F. Schwierz and J. J. Liou, “RF transistors: Recent developments and roadmap toward terahertz applications,” Solid-State Electronics, vol. 51, no. 8, pp. 1079–1091, 2007.

[24] R. L. Anderson, “Germanium-Gallium Arsenide Heterojunctions [Letter to the Editor],” IBM Journal of Research and Development, vol. 4, no. 3, pp. 283–287, 1960.

[25] E. F. Schubert, “Delta doping of III–V compound semiconductors: Fundamentals and device applications,” Journal of Vacuum Science & Technology A: Vacuum, Surfaces, and Films, vol. 8, no. 3, pp. 2980–2996, 1990.

[26] S. Prasad, H. Schumacher, and A. Gopinath, High-speed Electronics and Optoelectronics: Devices and Circuits: Cambridge University Press, 2009.

[27] S. M. Sze and K. K. Ng, Physics of semiconductor devices: Wiley-interscience, 2006.

[28] S. Kayali, G. Ponchak, and R. Shaw, “GaAs MMIC reliability assurance guideline for space applications,” JPL Publication, pp. 96–25, 1996.

[29] I. Newsroom, Intel 22nm 3-D Tri-Gate Transistor Technology: Link: http://newsroom. intel. com/docs/DOC-2032.

[30] B. G. Streetman and S. Banerjee, Solid state electronic devices: Prentice-Hall Englewood Cliffs, NJ, 1995.

[31] V. Palankovski and R. Quay, Analysis and simulation of heterostructure devices: Springer Verlag, 2004.

[32] Ioffe Institute website, “New Semiconductor Materials. Characteristics and Properties”, Available at: http://www.ioffe.ru/SVA/NSM/Semicond/

[33] Y. Fuchs, O. Soppera, and K. Haupt, “Photopolymerization and photostructuring of molecularly

imprinted polymers for sensor applications–A review," Analytica chimica acta, 2011.

[34] Fukui, Hatsuaki. "Optimal noise figure of microwave GaAs MESFET's." Electron Devices, IEEE Transactions on 26.7 (1979): 1032-1037.

[35] Liechti, Charles A. "Microwave field-effect transistors--1976." Microwave Theory and Techniques, IEEE Transactions on 24.6 (1976): 279-300.

[36] Gonzalez, Guillermo. Microwave transistor amplifiers: analysis and design. Vol. 2. New Jersey: Prentice hall, 1997.

[37] Mohamad Isa, M., et al. "Fabrication of novel high frequency and high breakdown InAlAs-InGaAs pHEMTs." Advanced Semiconductor Devices & Microsystems (ASDAM), 2010 8th International Conference on. IEEE, 2010.

[38] Rai-Choudhury, P. Handbook of Microlithography, Micromachining and Microfabrication: Handbook of Microlithography. Vol. 1. Iet, 1997.

[39] La Fontaine, Bruno. "Lasers and moore's law." SPIE Professional (2010): 20.

[40] Alonso, Marcelo, and Edward J. Finn. "Physics Addison-Wesley." Harlow, UK (1992).

[41] Moolji, A. A., S. R. Bahl, and J. A. Del Alamo. "Impact ionization in InAlAs/InGaAs HFET's." Electron Device Letters, IEEE 15.8 (1994): 313-315.

[42] Faqir, Mustapha. Analysis of the physical mechanisms limiting performance and reliability of GaN based HEMTs. Diss. University of "Modena, 2009.

[43] Guerra, Diego, et al. "Aspect ratio impact on RF and DC performance of state-of-the-art short-channel GaN and InGaAs HEMTs." Electron Device Letters, IEEE 31.11 (2010): 1217-1219.

[44] Matsuzaki, Hideaki, et al. "Suppression of short-channel effect in pseudomorphic In0. 25Al0. 75P/In0. 75Ga0. 25As high electron mobility transistors." IEICE Electronics Express 1.11 (2004): 292-297.

[45] Zhao, Han, et al. "High performance In 0.7 Ga 0.3 As metal-oxide-semiconductor transistors with mobility≫ 4400." Applied Physics Letters 94.19 (2009): 193502-193502.

[46] Saguatti, Davide, et al. "Improvement of breakdown and DC-to-pulse dispersion properties in field-plated InGaAs-InAlAs pHEMTs." Compound Semiconductor Week (CSW/IPRM), 2011 and 23rd International Conference on Indium Phosphide and Related Materials. IEEE, 2011.

[47] Chau, Robert, Suman Datta, and Amlan Majumdar. "Opportunities and challenges of III-V nanoelectronics for future high-speed, low-power logic applications." Compound Semiconductor Integrated Circuit Symposium, 2005. CSIC'05. IEEE. IEEE, 2005.

[48] Kim, Dae-Hyun, et al. "Performance evaluation of 50 nm In/sub 0.7/Ga/sub 0.3/As HEMTs for beyond-CMOS logic applications." Electron Devices Meeting, 2005. IEDM Technical Digest. IEEE International. IEEE, 2005.

[49] Pavlidis, Dimitris. "HBT vs. PHEMT vs. MESFET: What's best and why." International Conference on Compound Semiconductor Manufacturing Technology. 1999.

[50] Wei, John S., et al. "An Enhancement Mode PHEMT for Single Supply Power Amplifiers." (1998).

[51] Okamoto, Y., K. Matsunaga, and M. Kuzuhara. "Enhancement-mode buried gate InGaP/AlGaAs/InGaAs heterojunction FETs fabricated by selective wet etching." Electronics Letters 31.25 (1995): 2216-2218.

[52] Suemitsu, Tetsuya, et al. "High-performance 0.1-μm gate enhancement-mode InAlAs/InGaAs HEMT's using two-step recessed gate technology." Electron Devices, IEEE Transactions on 46.6 (1999): 1074-1080.

[53] Ao, Jin-Ping, et al. "InP-based enhancement-mode pseudomorphic HEMT with strained In/sub 0.45/Al/sub 0.55/As barrier and In/sub 0.75/Ga/sub 0.25/As channel layers." Electron Device Letters, IEEE 21.5 (2000): 200-202.

[54] N. Harada, S. Kuroda, T. Katakami, K. Hikosaka, T. Mimura, and M. Abe, "Pt-based gate enhancement-mode InAlAs/InGaAs HEMTs for large-scale integration," Indium Phosphide and Related Materials, 1991, Third International Conference, pp. 377–380, 1991.

[55] K. J. Chen, T. Enoki, K. Maezawa, K. Arai, and M. Yamamoto, "High-performance InP-based enhancement-mode HEMTs using non-alloyed ohmic contacts and Pt-based buried-gate technologies: Electron Devices, IEEE Transactions on," Electron Devices, IEEE Transactions on DOI - 10.1109/16.481725, vol. 43, no. 2, pp. 252–257, 1996.

[56] A. Mahajan, M. Arafa, P. Fay, C. Caneau, and I. Adesida, "0.3-μm gate-length enhancement mode InAlAs/InGaAs/InP high-electron mobility transistor," Electron Device Letters, IEEE, vol. 18, no. 6, pp. 284–286, 1997.

[57] Sun, Yanning, et al. "Enhancement-Mode Buried-Channel." IEEE electron device letters 28.6 (2007): 473.

[58] C.I. Kuo, H.T. Hsu, E.Y. Chang, C.Y. Chang, Y. Miyamoto, S. Datta, M. Radosavljevic, G.W. Huang, and C.T. Lee, Eds, RF and logic performance improvement of In0. 7Ga0. 3As/InAs/In0. 7Ga0. 3As composite-channel HEMT using gate-sinking technology: Institute of Electrical and Electronics Engineers, 2008.

[59] Chuang, H. F., et al. "Thermal annealing of Pd/InAlAs Schottky contacts for transistor buried-gate technologies." Journal of applied physics 83.1 (1998): 366-371.

[60] Chiu, Hsien-Chin, et al. "High thermal stability AlGaAs/InGaAs enhancement-mode pHEMT using palladium-gate technology." Microelectronics Reliability 50.6 (2010): 847-850.

[61] T. Ishikura, Z. Cui, T. Matsuda, and K. Yoh, "Thermal Stability of Pd Gate in Pseudomorphic InGaAs Heterostructures," Japanese Journal of Applied Physics, vol. 50, no. 6, pp. 06GF19, 2011.

[62] E. F. Chor, W. K. Chong, and C. H. Heng, "Alternative (Pd, Ti, Au) contacts to (Pt, Ti, Au) contacts for In0. 53Ga0. 47As," Journal of applied physics, vol. 84, no. 5, pp. 2977–2979, 1998.

[63] Bouloukou, A., et al. "Novel high breakdown InGaAs/InAlAs pHEMTs for radio astronomy applications." 4th ESA Workshop on Millimeter Wave Technology and (7th MIN Millimiter-Wave Int. Symp.), Finland. 2006.

[64] Suemitsu, Tetsuya. "Recent achievements in the reliability of InP-based HEMTs." Thin solid films 515.10 (2007): 4378-4383.

[65] Bahl, Sandeep R., and I. A. Del Alamo. "Elimination of mesa-sidewall gate leakage in InAlAs/InGaAs heterostructures by selective sidewall recessing." Electron Device Letters, IEEE 13.4 (1992): 195-197.

[66] Schaller, Robert R. "Moore's law: past, present and future." Spectrum, IEEE 34.6 (1997): 52-59.

[67] Owa, Soichi, and Hiroyuki Nagasaka. "Immersion Lithography; its potential performance and issues." Proc. of SPIE Vol. Vol. 5040. 2003.

[68] Finders, Jo, et al. "Double patterning for 32nm and below: an update." Advanced Lithography. International Society for Optics and Photonics, 2008.

[69] Newsroom, Intel. "Intel 22nm 3-D Tri-Gate Transistor Technology."

[70] Courtland, R. "3-D transistors for all." Spectrum, IEEE 50.1 (2013): 11-12.

[71] Levinson, Harry J. "Principles of lithography." Bellingham: SPIE, 2005.

[72] Muzio, Ed. "Optical lithography cost of ownership (COO)–final report for LITG501." International SEMATECH, Technology Transfer\ 00104014A-TR (2000).

[73] Rai-Choudhury, P. Handbook of Microlithography, Micromachining and Microfabrication: Handbook of Microlithography. Vol. 1. Iet, 1997.

[74] Private communication (03/2013): Gordon Hughes from Compugraphics International Ltd, “Cost of ownership of different photomasks”, Email: customer-support-uk@cgi.co.uk

[75] Private communication (02/2013): Jaydeep Patel from Global Capital Equipment Market, “Cost of entry stepper lithography system for research purposes”, Email: jd@GCEmarket.com

[76] Private communication (03/2013): Chris Royston from E M Systems Support Ltd, “Cost of E-beam lithography system for research purposes”, Email: chris@emsys.co.uk

[77] Private communication (01/2013): Ian Smith from SUSS MicroTec Lithography GmbH, “Cost of contact lithography system for research purposes”, Email: ian.smith@suss.com

[78] Moran, D. A. J., et al. "Self-aligned 0.12/spl mu/m T-gate In/sub. 53/Ga/sub. 47/As/In/sub. 52/Al/sub. 48/As HEMT technology utilising a non-annealed ohmic contact strategy." European Solid-State Device Research, 2003. ESSDERC'03. 33rd Conference on. IEEE, 2003.

[79] Kim, Dae-Hyun, et al. "High performance of 0.15μm quasi enhancement-mode (E-mode) In< sub> 0.4</sub> GaAs/In< sub> 0.4</sub> AlAs metamorphic HEMTs on GaAs substrate using new triple-gate technology." Indium Phosphide and Related Materials, 2004. 16th IPRM. 2004

International Conference on. IEEE, 2004.

[80] Jie, H. U. A. N. G., et al. "InAlAs/InGaAs Pseudomorphic High Eelectron Mobility Transistors Grown by Molecular Beam Epitaxy on the InP Substrate." Chinese Physics Letters 27.11 (2010): 118502.

[81] Kuo, Chien-I., Heng-Tung Hsu, and Edward Yi Chang. "InAs Channel-Based Quantum Well Transistors for High-Speed and Low-Voltage Digital Applications." Electrochemical and Solid-State Letters 11.7 (2008): H193-H196.

[82] Shinohara, K., et al. "Ultrahigh-speed pseudomorphic InGaAs/InAlAs HEMTs with 400-GHz cutoff frequency." Electron Device Letters, IEEE 22.11 (2001): 507-509.

[83] Yamashita, Yoshimi, et al. "Ultra-short 25-nm-gate lattice-matched InAlAs/InGaAs HEMTs within the range of 400 GHz cutoff frequency." Electron Device Letters, IEEE 22.8 (2001): 367-369.

[84] Yamashita, Yoshimi, et al. "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 23.10 (2002): 573-575.

[85] Shinohara, K., et al. "547-GHz f< sub> t</sub> In< sub> 0.7</sub> Ga< sub> 0.3</sub> As-In< sub> 0.52</sub> Al< sub> 0.48</sub> As HEMTs with reduced source and drain resistance." Electron Device Letters, IEEE 25.5 (2004): 241-243.

[86] Moran, David AJ, et al. "50-nm self-aligned and “standard” T-gate InP pHEMT comparison: the influence of parasitics on performance at the 50-nm node." Electron Devices, IEEE Transactions on 53.12 (2006): 2920-2925.

[87] Lee, Kang-Sung, et al. "35-nm zigzag T-gate In0: 52Al0: 48As/In0: 53Ga0: 47As metamorphic GaAs HEMTs with a ultrahigh fmax of 520 GHz." IEEE Electron Device Lett 28.8 (2007): 672-675.

[88] Kim, Dae-Hyun, and Jesús A. Del Alamo. "30-nm InAs PHEMTs with fT= 644 GHz and fmax= 681 GHz." IEEE Electron Device Letters 31.8 (2010): 806.

[89] King, C. R., L. Y. Lin, and M. C. Wu. "Out-of-plane refractive microlens fabricated by surface micromachining." Photonics Technology Letters, IEEE 8.10 (1996): 1349-1351.

[90] Toshiyoshi, Hiroshi, et al. "A Surface Micromachined Optical Scanner Array Using Photoresist Lenses Fabricated by a Thermal Reflow Process." Journal of Lightwave Technology 21.7 (2003): 1700.

[91] Pan, C. T., and C. Y. Su. "Study of micro-lens array by reflow process." Journal of modern optics 55.17 (2008): 2843-2856.

[92] Tan, S. W., et al. "Sub-0.5-μm gate doped-channel field-effect transistors with HEMT-like channel using thermally reflowed photoresist and spin-on glass." Semiconductor science and technology 19.2 (2004): 167.

[93] Meng, C. C., G. R. Liao, and S. S. Lu. "Formation of submicron T-gate by rapid thermally reflowed resist with metal transfer layer." Electronics Letters 37.16 (2001): 1045-1046.

[94] Lien, Y. C., et al. "Low-noise metamorphic HEMTs with reflowed 0.1-μm T-gate." Electron Device Letters, IEEE 25.6 (2004): 348-350.

[95] Downey, Brian P., et al. "Thermally reflowed ZEP 520A for gate length reduction and profile rounding in T-gate fabrication." Journal of Vacuum Science & Technology B: Microelectronics and Nanometer Structures 30.5 (2012): 051603-051603.

[96] Chu, Kuei-Yi, et al. "Characteristics of InGaP/InGaAs pseudomorphic high electron mobility transistors with triple delta-doped sheets." Semiconductors 46.2 (2012): 203-207.

[97] Material Safety Data Sheets from Fisher Scientific UK Ltd, Available at: http://www.fisher.co.uk/index.php/en/msds-search

[98] Brigham Young University, “Silicon Dioxide/Nitride Color vs. Film Thickness and Viewing Angle Calculator”. Available at: http://www.cleanroom.byu.edu/color_chart.phtml

[99] Berger, H. H. "Models for contacts to planar devices." Solid-State Electronics 15.2 (1972): 145-158.

[100] Murrmann, Helmuth, and Dietrich Widmann. "Current crowding on metal contacts to planar devices." Electron Devices, IEEE Transactions on 16.12 (1969): 1022-1024.

[101] Reeves, G. K., and H. B. Harrison. "Obtaining the specific contact resistance from transmission line model measurements." Electron Device Letters, IEEE 3.5 (1982): 111-113.

[102] Schroder, Dieter K. Semiconductor material and device characterization. Wiley-interscience, 2006.

[103] Sischka, Franz. "The IC-CAP characterization and modelling handbook." Agilent Technologies, Dept EEsof, Munich (2002).

[104] Mishra, U., & Singh, J. (2007). Semiconductor device physics and design. Springer.

[105] Officical website (2013), “Agilent Technologies - Integrated Circuit Characterization and Analysis Program 2010”.
Available :http://www.home.agilent.com/en/pd-1889249/ic-cap-201008?nid=-34268.942202.00&cc=GB&lc=eng

[106] Official website (2013), “Cascade Microtech - Wincal. High-performance RF calibration software” Available: http://www.cmicro.com/products/probes/wincal-xe

[107] Microtech, Cascade. "A Guide to Better Vector Network Analyser Calibrations for Probe-Tip Measurement." Technical Brief, TECHBRIEF4-0694 (1994).

[108] Datasheet (2013), “Agilent - B1500A Semiconductor Device Analyzer”. Available: http://cp.literature.agilent.com/litweb/pdf/5989-2785EN.pdf

[109] Datasheet (2013), “Karl Suss - PM5 Probe System Datasheet”. Available:http://www.ptb-sales.com/manuals/karlsuss/PM5_DC_TD.PDF

[110] Datasheet (2013), “Anritsu - 37369A Vector Network Analyzer Datasheet”. Available: http://www.teknetelectronics.com/DataSheet/ANRITSU/ANRIT_373XXASeriesN23910.pdf

[111] Datasheet (2013), “HP - Agilent 8510C Network Analyzer Datasheet”. Available:http://cp.literature.agilent.com/litweb/pdf/5091-8484E.pdf

[112] Packard, Hewlett. "HP 4142B Modular DC source/Monitor Practical Applications—High Speed DC Characterization of Semiconductor Devices from Sub pA to 1A." (1987): 1-4.

[113] VanDerVoorn, P., et al. "A 32nm low power RF CMOS SOC technology featuring high-k/metal gate." VLSI Technology (VLSIT), 2010 Symposium on. IEEE, 2010.

[114] Datasheet (2013), “TriQuint Semiconductor - 0.25-µm mmW pHEMT 3MI”. Available: http://www.triquint.com/products/d/25-mmw-phemt-3mi

APPENDIX A1: COVENTIONAL 1 µm FLAT GATE pHEMT PROCESS FLOW

Structure	No.	Step name	Chemicals	Equipment	Temp.	Time	Comments
MESA	1.1	Sample cleaning	NMP(1165)/Acetone/IPA	USB	RT	5m each	Finally dry the sample with N_2
	1.2	Preheat	N/A	Hotplate	120 ℃	~1m	After baking, cool down on the heat sink for 60s
	1.3	Resist Coating	Photoresist: S1805	Spinner	RT	30s	Program-4 (Acc=2000, RPM=4000, Duration=30s)
	1.4	Prebake	N/A	Hotplate	115 ℃	1m	After hotplate, heat sinks for 1 min at least.
	1.5	Exposure	N/A	MA4	RT	20s	Intensity=0.9mW/cm^2 for i-line
	1.6	Develop S1805	MIF 319	Beaker	RT	1m	Rinse with water and dry with N_2.
	1.7	Post-bake	N/A	Oven	120 ℃	30m	To harden the remaining resist before etch.
	1.8	Ortho-phosphoric etch	H_3PO_4:H_2O_2:H_2O (3:1:50) *H_2O_2 High grade	Beaker	RT	60s	Ortho-phosphoric etch rate: ~100nm/min InGaAs/InAlAs type pHEMT, and, ~80nm/min GaAs/InAlAs type pHEMT. It doesn't etch InP. Mesa height = ~100nm.
	1.9	Side-wall etch	Succinic acid powder (10g), H_2O (50ml), NH_3 (~10ml to pH of 5.5), H_2O_2 (5 ml)	pH meter	RT	10m	Succinic etch rate: InGaAs = 240A/min InAlAs = 2A/min Etch 50A InGaAs layer InGaAs/InAlAs selectivity = 120:1 Allow the solution to settle for 5m before etching. Add H_2O_2 only after pH of 5.5 is achieved by the intermittent addition of ammonia.

Ohmic	2.1	Sample cleaning	Acetone/IPA	USB	RT	5m each	Dry with N_2
	2.2	Prebake	N/A	Hotplate	>100 ℃	1m	To drive off moisture from the sample. After baking, cool down on the heat sink for 60s.
	2.3	Resist coating	Photoresist: AZnLOF 2070 (2um grade)	Spinner	RT	1m	Program-6 (Acc=2000, RPM=3000, Duration=30s)
	2.4	Edge bead removal	Solvent: AZ EBR	Spinner	RT	1m	Apply solvent on each corner, the spin again with Program-6
	2.5	Soft-bake	N/A	Hotplate	110 ℃	1m	After baking, cool down on the heat sink for 60s
	2.6	Exposure	N/A	MA4	RT	5.5s	Intensity=0.9mW/cm^2 for i-line
	2.7	PEB (Post Exposure Bake)	N/A	Hotplate	110 ℃	1m	After baking, cool down on the heat sink for 60s
	2.8	Develop AZnLOF 2070	MIF 326	Beaker	RT	1m	Because of the negative resist, the exposed region polymerises and unexposed region gets dissolved in the developer. Rinse with water and dry with N_2.
	2.9	Load evaporator with crude material	AuGe/Au 70mg/12cm	JNR	RT	N/A	Crude Au (and also boats in the same beaker if needed) pre-cleaned in Trike/Acetone/IPA (5m each). Then etched in H_2O:HCL (15ml:15ml) for 2 min. Finally cleaned with water and dried with N_2.
	2.10	Residue removal by plasma etching	O_2	Plasma kit	RT	20s	P:60mTorr, Power= 25W(100 a.u.), O_2 flow: 50 sccm
	2.11	De-oxidation	HCl:H_2O (1:1)	N/A	RT	30s	15:15ml Put the sample for 30sec and then rinse with water and dries with N_2. Load the sample very fast in the evaporator so that to delay HCL evaporation.

	2.12	Evaporation	AuGe/Au 50nm/100nm	JNR	N/A	N/A	Load ~70mg AuGe and 12cm Au
	2.13	Lift-off	NMP (1165)		85 ℃	>30m	Can be left overnight at RT
	2.14	Clean	Water			3m	Dry with N_2
	2.15	Annealing	N/A	Furnace	280 ℃	90s	N_2 flow: 150 (N_2 at 30 is idle), Rc = ~0.1Ω/mm
	2.16	TLM measurements	N/A	IC-CAP	RT	N/A	File: Constant current (I = 1mA)
Gate	3.1	Sample cleaning	Acetone/IPA	USB (Level 1)	RT	5m each	Finally dry the sample with N_2
	3.2	Prebake	N/A	Hotplate	>100 ℃	1m	To drive off moisture from the sample. After baking, cool down on the heat sink for 60s.
	3.3	Resist Coating	Photoresist: AZnLOF 2070 (1μm grade)	Spinner	RT	30s	Program-6 (Acc=2000, RPM=3000, Duration=30s)
	3.4	Edge bead removal	Solvent: AZ EBR	Spinner	RT	1m	Apply solvent on each corner, the spin again with Program-6
	3.5	Soft-bake	N/A	Hotplate	110 ℃	1m	After baking, cool down on the heat sink for 60s
	3.6	Exposure	N/A	MA4		12s	Intensity=0.9mW/cm^2 for i-line. Wedge Error =1.
	3.7	PEB (Post Exposure Bake)	N/A	Hotplate	110 ℃	1m	After baking, cool down on the heat sink for 60s
	3.8	Develop AZ nLOF 2070	MIF 326	Beaker	RT	5m	Develop in agitator. Rinse with DI water for 30s
	3.9	Load evaporator with crude material	Ti/Au/Au 1.5cm/15cm/15cm	EDWINA	RT	N/A	Crude Au and Ti (and also boats in the same beaker if needed) pre-cleaned in Trike/Acetone/IPA (5m each). Then etched in H2O:HCL (15ml:15ml) for 2 min. Finally cleaned with water and dried with N2. Outgases the boats if needed.

	3.10	Plasma residue removal	O_2	Plasma kit	RT	20s	P:60mT, Power= 25W (100 a.u.), O_2 flow: 50 sccm
	3.11	Gate Recess	Succinic acid powder (10g), H2O (50ml), NH3 (~10ml to pH of 5.5), H2O2 (5 ml)	pH meter	RT	5m	Succinic etch rate: InGaAs = 240A/min InAlAs = 2A/min Etch 50A InGaAs layer InGaAs/InAlAs selectivity = 120:1 Allow the solution to settle for 5m before etching. Add H_2O_2 only after pH of 5.5 is achieved by the intermittent addition of ammonia.
	3.12	Evaporation	Ti/Au 50nm/450nm	EDWINA	N/A	N/A	* Load 1.5cm Ti (or 2cm but bent it first), and, 15cm of Au in each of the two adjacent boats. ** Ti has high barrier height, and not allows Au to diffuse.
	3.13	Lift-off	NMP	USB	85 ℃	~ 30m	Rinse with water and dry with N_2.
	3.14	Sample cleaning	Water	Beaker	RT	3m	Dry with N_2.
	3.11	Diode Test	N/A	IC-CAP	RT	N/A	
Bonding pad	4.1	Sample cleaning	Acetone-IPA	USB	RT	5m each	Finally dry the sample with N_2.
	4.2	Prebake	N/A	Hotplate	>100 ℃	1m	To drive off moisture from the sample. After baking, cool down on the heat sink for 60s.
	4.3	Resist coating	Photoresist: AZnLOF 2070 (2um grade)	Spinner	RT	30s	Program-6(Acc=2000, RPM=3000, Duration=30s)
	4.4	Soft-bake	N/A	Hotplate	110 ℃	1m	After baking, cool down on the heat sink for 60s
	4.5	Exposure	N/A	MA4	RT	5.5s	Intensity=0.9mW/cm^2 for i-line
	4.6	PEB (Post Exposure Bake)	N/A	Hotplate	110 ℃	1m	After baking, cool down on the heat sink for 60s

	4.7	Develop AZ nLOF 2070	MIF 326	Beaker	RT	1m	Because of negative resist, exposed region polymerises and unexposed region gets dissolved in the developer.
	4.8	Load evaporator with crude material	Ti/Au/Au = 1.5cm/15cm/15cm	EDWINA	N/A	N/A	Crude Au and Ti (and also boats in the same beaker if needed) pre-cleaned in Trike/Acetone/IPA (5m each). Then etched in H_2O: HCL (15ml:15ml) for 2 min. Finally cleaned with water and dried with N_2.
	4.9	Residue removal by plasma etching	O_2	Plasma kit	RT	20s	P:60mT, Power= 25W (100 a.u.), O_2 flow: 50 sccm
	4.10	Evaporation	Ti/Au = 50nm/450um	EDWINA	N/A	N/A	Load 1.5cm (or 2cm but bent it first) Ti, and, 15cm of Au in each of the two adjacent boats.
	4.11	Lift-off	NMP	USB	80 ℃	~ 30m	Rinse with water.
	4.12	Clean	Water	Beaker	RT	3m	Dry with N_2.
	4.12	Diode Test	N/A	IC-CAP	RT	N/A	
Dielectric bridge support	5.1	Sample cleaning	Acetone-IPA	USB	RT	5m each	Finally dry the sample with N_2.
	5.2	Prebake	N/A	Hotplate	>100 ℃	1m	To drive off mositure from the sample. After baking, cool down on the heat sink for 60s.
	5.3	Resist Coating	Photoresist: SF11	Spinner	RT	45s	Program-13 (Acc=10000, rpm=500,time=5s), (Acc=10000, rpm=4000,time=45s), (Acc=10000, rpm=7000,time=5s)
	5.4	Soft-bake	N/A	Hotplate	190 ℃	5m	After baking, cool down on the heat sink for 60s
	5.5	Resist Coating	Photoresist: S1805	Beaker	RT	30s	Program-4(Acc=2000, rpm=4000,time=30s)
	5.6	Soft-bake	N/A	Hotplate	115 ℃	1m	After baking, cool down on the heat sink for 60s

	5.7	Exposure	N/A	MA4	RT	18s	Intensity=0.9mW/cm^2 for i-line.
	5.8	Develop S1805	MIF 319	Beaker	RT	1m	
	5.9	DUV (Deep UV)	N/A	UV EPROM eraser	RT	15m	Deep Ultraviolet Flooding is done to make sure SF11 reacts to XP101A
	5.10	Develop SF11	XP101A	Beaker	RT	3m	Visually see the colour change to bright grey of the exposed area and rinse the sample with DI water.
	5.11	Sample cleaning	Acetone-IPA	USB	RT	5m each	Finally dry the sample with N_2.
	5.12	Reflow	N/A	Hot plate	200 ℃	10m	Reflow rate = 0.17um/min
Metal bridge	6.1	Resist Coating	Photoresist: S1813	Spinner	RT	30s	Program-4(Acc=2000, rpm=4000,time=30s)
	6.2	Soft-bake	N/A	Hotplate	115 ℃	1m	After baking, cool down on the heat sink for 60s
	6.3	Exposure	N/A	MA4	RT	40s	Intensity=6mW/cm^2. Dielectric bridge mask.
	6.4	Develop S1813	Microdev:H_2O = 1:1	Beaker	RT	1min 30s	
	6.5	Evaporation	Au = 500 nm	EDDY JNR	N/A	N/A	Load 15cm Au
	6.6	Lift-off	Acetone+IPA only!	NO HEAT	RT	~ 30m	Rinse with water

APPENDIX A2: NOVEL SUBMICRON T-GATE pHEMT PROCESS USING SOFT REFLOW

Structure	No.	Step name	Chemicals	Equipment	Temp.	Time	Comments
MESA	1.1	Sample cleaning	NMP(1165)/Acetone/IPA	USB	RT	5m each	Finally dry the sample with N_2
	1.2	Preheat	N/A	Hotplate	120 ℃	~1m	After baking, cool down on the heat sink for 60s
	1.3	Resist Coating	Photoresist: S1805	Spinner	RT	30s	Program-4 (Acc=2000, RPM=4000, Duration=30s)
	1.4	Soft-bake	N/A	Hotplate	115 ℃	1m	After hotplate, heat sinks for 1 min at least.
	1.5	Exposure	N/A	MA4	RT	20s	Intensity=0.9mW/cm^2 for i-line
	1.6	Develop S1805	MIF 319	Agitator& beaker	RT	1m	Rinse with water and dry with N_2.
	1.7	Post-bake	N/A	Oven	120 ℃	30m	To harden the remaining resist before etch.
	1.8	Ortho-phosphoric etch	H_3PO_4:H_2O_2:H_2O (3:1:50) *H_2O_2 High grade	Agitator& beaker	RT	60s	Ortho-phosphoric etch rate: ~100nm/min InGaAs/InAlAs type pHEMT, and, ~80nm/min GaAs/InAlAs type pHEMT. It doesn't etch InP. Mesa height = ~100nm.
	1.9	Side-wall etch	Succinic acid powder (10g), H_2O (50ml), NH_3 (~10ml to pH of 5.5), H_2O_2 (5 ml)	pH meter	RT	10m	Succinic etch rate: InGaAs = 240A/min InAlAs = 2A/min Etch 50A InGaAs layer InGaAs/InAlAs selectivity = 120:1 Allow the solution to settle for 5m before etching. Add H_2O_2 only after pH of 5.5 is achieved by the intermittent addition of ammonia.

Ohmic	2.1	Sample cleaning	Acetone/IPA	USB (Level 1)	RT	5m each	Dry with N_2
	2.2	Prebake	N/A	Hotplate	>100 ℃	1m	To drive off moisture from the sample. After baking, cool down on the heat sink for 60s.
	2.3	Resist coating	Photoresist: AZnLOF 2070 (2um grade)	Spinner	RT	1m	Program-6 (Acc=2000, RPM=3000, Duration=30s)
	2.4	Edge bead removal	Solvent: AZ EBR	Spinner	RT	1m	Apply solvent on each corner, the spin again with Program-6
	2.5	Soft-bake	N/A	Hotplate	110 ℃	1m	After baking, cool down on the heat sink for 60s
	2.6	Exposure	N/A	MA4	RT	5.5s	Intensity=0.9mW/cm^2 for i-line
	2.7	PEB (Post Exposure Bake)	N/A	Hotplate	110 ℃	1m	After baking, cool down on the heat sink for 60s
	2.8	Develop AZnLOF 2070	MIF 326	Agitator& beaker	RT	1m	Because of the negative resist, the exposed region polymerises and unexposed region gets dissolved in the developer. Rinse with water and dry with N_2.
	2.9	Load evaporator with crude material	AuGe/Au 70mg/12cm	JNR	RT	N/A	Crude Au (and also boats in the same beaker if needed) pre-cleaned in Trike/Acetone/IPA (5m each). Then etched in H_2O:HCL (15ml:15ml) for 2 min. Finally cleaned with water and dried with N_2.
	2.10	Residue removal by plasma etching	O_2	Plasma kit	RT	20s	Pressure: 60mTorr, Power= 25W(100 a.u.), O_2 flow: 50 sccm
	2.11	De-oxidation	HCl:H_2O (1:1)	N/A	RT	30s	15:15ml Put the sample for 30sec and then rinse with water and dries with N_2. Load the sample very fast in the evaporator so that to delay HCL evaporation.

	2.12	Evaporation	AuGe/Au 50nm/100nm	JNR	N/A	N/A	Load ~70mg AuGe and 12cm Au
	2.13	Lift-off	hot NMP + DI water	Hotplate	85 ℃	>30m	Can be left overnight at RT
	2.14	Annealing	N/A	Furnace	280 ℃	90s	N_2 flow: 150 (N_2 at 30 is idle), Rc = ~0.1Ω/mm
	2.15	TLM measurements	N/A	IC-CAP	RT	N/A	File: Constant current (I = 1mA)
Gate foot print	3.1	Sample cleaning	hot NMP + DI water	Agitator& beaker	85 ℃	15m	Immersing in hot NMP first, then clean with DI water. Avoid USB!
	3.2	Prebake	N/A	Hotplate	>100 ℃	1m	To drive off moisture from the sample. After baking, cool down on the heat sink for 60s.
	3.3	Silicon Nitride deposition	Si3N4 = 200nm	PECVD	250 ℃	30s	Deposition parameters: SiH4 = 6sccm, NH3 = 20sccm, N2 = 55 sccm, temperature = 250 ℃, pressure = 300 mTorr, power = 150W, rate = ~10nm/min
	3.4	Sample cleaning	hot NMP + DI water	Agitator& beaker	85 ℃	15m	Immersing in hot NMP first, then clean with DI water. Avoid USB!
	3.5	Resist Coating	Photoresist: S1805	Spinner	RT	30s	Program-4 (Acc=2000, RPM=4000, Duration=30s)
	3.6	Soft-bake	N/A	Hotplate	115 ℃	1m	After hotplate, heat sinks for 1 min at least.
	3.7	Gate foot print exposure	N/A	MA4	RT	20s	Definition of the 1um gate footprint
	3.8	Develop S1805	MIF 319	Beaker	RT	1m	Rinse with water and dry with N_2.
	3.9	Reflow	NMP	Reflow chamber	<40 ℃	Variable	Soft reflow process, reflow time varies with size

	3.10	Pattern transfer	CF_4	Plasma kit	RT	6m	Pressure: 20mTorr, Power = 20W(90 a.u.), CF_4 flow: 40 sccm Etch time varies with Si_3N_4 thickness
	3.11	Residue removal by plasma etching	O_2	Plasma kit	RT	20s	Pressure: 60mTorr, Power= 20W(90 a.u.), O_2 flow: 50 sccm
Gate head	4.1	Sample cleaning	hot NMP + DI water	Agitator& beaker	85 ℃	15m	Immersing in hot NMP first, then clean with DI water. Avoid USB!
	4.2	Prebake	N/A	Hotplate	>100 ℃	1m	To drive off moisture from the sample. After baking, cool down on the heat sink for 60s.
	4.3	Resist Coating	Photoresist: AZnLOF 2070 (1μm grade)	Spinner	RT	30s	Program-6 (Acc=2000, RPM=3000, Duration=30s)
	4.4	Edge bead removal	Solvent: AZ EBR	Spinner	RT	1m	Apply solvent on each corner, the spin again with Program-6
	4.5	Soft-bake	N/A	Hotplate	110 ℃	1m	After baking, cool down on the heat sink for 60s
	4.6	Exposure	N/A	MA4	RT	12s	Definition of 1um gate head. Intensity=0.9mW/cm^2 for i-line. Wedge Error =1.
	4.7	PEB (Post Exposure Bake)	N/A	Hotplate	110 ℃	1m	After baking, cool down on the heat sink for 60s
	4.8	Develop AZ nLOF 2070	MIF 326	Agitator& beaker	RT	5m	Develop in agitator. Rinse with DI water for 30s
	4.9	Load evaporator with crude material	Ti/Au/Au 1.5cm/15cm/15cm	EDWINA	RT	N/A	Crude Au and Ti (and also boats in the same beaker if needed) pre-cleaned in Trike/Acetone/IPA (5m each). Then etched in H2O:HCL (15ml:15ml) for 2 min. Finally cleaned with water and dried with N_2. Outgases the boats if needed.
	4.10	Residue removal by plasma	O_2	Plasma kit	RT	20s	P:60mT, Power= 25W (100 a.u.), O_2 flow: 50 sccm

		etching					
	4.11	Gate Recess	Succinic acid powder (10g), H2O (50ml), NH3 (~10ml to pH of 5.5), H2O2 (5 ml)	pH meter	RT	5m	Succinic etch rate: InGaAs = 240A/min InAlAs = 2A/min Etch 50A InGaAs layer InGaAs/InAlAs selectivity = 120:1 Allow the solution to settle for 5m before etching. Add H_2O_2 only after pH of 5.5 is achieved by the intermittent addition of ammonia.
	4.12	Evaporation	Ti/Au 50nm/450nm	EDWINA	N/A	N/A	* Load 1.5cm Ti (or 2cm but bent it first), and, 15cm of Au in each of the two adjacent boats.
	4.13	Lift-off	hot NMP + DI water	Agitator& beaker	85 ℃	~ 30m	Rinse with DI water and dry with N_2. Avoid USB!
	4.14	Diode Test	N/A	IC-CAP	RT	N/A	
Bonding pad	5.1	Sample cleaning	hot NMP + DI water	Agitator& beaker	85 ℃	15m	Immersing in hot NMP first, then clean with DI water. Avoid USB!
	5.2	Prebake	N/A	Hotplate	>100 ℃	1m	To drive off moisture from the sample. After baking, cool down on the heat sink for 60s.
	5.3	Resist coating	Photoresist: AZnLOF 2070 (2um grade)	Spinner	RT	30s	Program-6 (Acc=2000, RPM=3000, Duration=30s)
	5.4	Soft-bake	N/A	Hotplate	110 ℃	1m	After baking, cool down on the heat sink for 60s
	5.5	Exposure	N/A	MA4	RT	5.5s	Intensity=0.9mW/cm^2 for i-line
	5.6	PEB (Post Exposure Bake)	N/A	Hotplate	110 ℃	1m	After baking, cool down on the heat sink for 60s
	5.7	Develop AZ nLOF 2070	MIF 326	Agitator&	RT	1m	Because of negative resist, exposed region

				beaker			polymerises and unexposed region gets dissolved in the developer.
	5.8	Load evaporator with crude material	Ti/Au/Au = 1.5cm/15cm/15cm	EDWINA	N/A	N/A	Crude Au and Ti (and also boats in the same beaker if needed) pre-cleaned in Trike/Acetone/IPA (5m each). Then etched in H_2O:HCL (15ml:15ml) for 2 min. Finally cleaned with water and dried with N_2.
	5.9	Plasma residue removal	O_2	Plasma kit	RT	20s	P:60mT, Power= 25W (100 a.u.), O_2 flow: 50 sccm
	5.10	Evaporation	Ti/Au = 50nm/450um	EDWINA	N/A	N/A	Load 1.5cm (or 2cm but bent it first) Ti, and, 15cm of Au in each of the two adjacent boats.
	5.11	Lift-off	hot NMP + DI water	Agitator& beaker	85 ℃	~ 30m	Rinse with DI water and dry with N_2. Avoid USB!
	5.12	Diode Test	N/A	IC-CAP	RT	N/A	
Dielectric bridge support	6.1	Sample cleaning	hot NMP + DI water	Agitator& beaker	85 ℃	15m	Immersing in hot NMP first, then clean with DI water. Avoid USB!
	6.2	Prebake	N/A	Hotplate	>100 ℃	1m	To drive off moisture from the sample. After baking, cool down on the heat sink for 60s.
	6.3	Resist Coating	Photoresist: SF11	Spinner	RT	45s	Program-13 (Acc=10000, rpm=500,time=5s), (Acc=10000, rpm=4000,time=45s), (Acc=10000, rpm=7000,time=5s)
	6.4	Soft-bake	N/A	Hotplate	190 ℃	5m	After baking, cool down on the heat sink for 60s
	6.5	Resist Coating	Photoresist: S1805	Agitator& beaker	RT	30s	Program-4 (Acc=2000, rpm=4000,time=30s)
	6.6	Soft-bake	N/A	Hotplate	115 ℃	1m	After baking, cool down on the heat sink for 60s

	6.7	Exposure	N/A	MA4	RT	18s	Intensity=0.9mW/cm^2 for i-line.
	6.8	Develop S1805	MIF 319	Beaker	RT	1m	
	6.9	DUV (Deep UV)	N/A	UV EPROM eraser	RT	15m	Deep Ultraviolet Flooding is done to make sure SF11 reacts to XP101A
	6.10	Develop SF11	XP101A	Agitator& beaker	RT	3m	Visually see the colour change to bright grey of the exposed area and rinse the sample with DI water.
	6.11	Sample cleaning	Acetone & IPA	Agitator& beaker	RT	5m each	Finally dry the sample with N_2. Avoid USB! Don't use NMP!
	6.12	Reflow	N/A	Hot plate	200 ℃	10m	Reflow rate = 0.17um/min
Metal bridge	7.1	Resist Coating	Photoresist: S1813	Spinner	RT	30s	Program-4 (Acc=2000, rpm=4000,time=30s)
	7.2	Soft-bake	N/A	Hotplate	115 ℃	1m	After baking, cool down on the heat sink for 60s
	7.3	Exposure	N/A	MA4	RT	40s	Intensity=6mW/cm^2. Dielectric bridge mask.
	7.4	Develop S1813	Microdev:H_2O = 1:1	Agitator& beaker	RT	1min 30s	
	7.5	Evaporation	Au = 500 nm	EDDY JNR	N/A	N/A	Load 15cm Au
	7.6	Lift-off	Acetone & IPA	Agitator& beaker	RT	~ 30m	Rinse with water. No heat! Avoid USB! Don't use NMP!