Radio Frequency InGaAs MOSFETs

Garigapati, Navya Sri

2023

Radio Frequency InGaAs MOSFETs

NAVYA SRI GARIGAPATI
DEPARTMENT OF ELECTRICAL AND INFORMATION TECHNOLOGY
FACULTY OF ENGINEERING | LTH | LUND UNIVERSITY

RF

Radio Frequency InGaAs MOSFETs

Doctoral Thesis

Navya Sri Garigapati

LUND UNIVERSITY

Department of Electrical and
Information Technology

Lund, September 2023

Academic thesis for the degree of Doctor of Philosophy, which, by due permission of the Faculty of Engineering at Lund University, will be publicly defended on Wednesday, 27 September, 2023, at 9:15 a.m. in lecture hall E:1406, Department of Electrical and Information Technology, Ole Römers Väg 3, 223 63 Lund, Sweden. The thesis will be defended in English.

The Faculty Opponent will be Prof. David Moran, Electronic & Nanoscale Engineering, University of Glasgow, United Kingdom.

Organisation:
LUND UNIVERSITY
Department of Electrical and Information Technology
Ole Römers Väg 3
223 63 Lund
Sweden

Author:
Navya Sri Garigapati

Document Type:
DOCTORAL THESIS

Date of Issue:
September 2023

Sponsoring Organisation(s):
Swedish Foundation for Strategic Research (SSF)
Swedish Research Council (VR)
European Union (EU)
NanoLund

Title:
Radio Frequency InGaAs MOSFETs

Abstract:
III-V-based Indium gallium arsenide is a promising channel material for high-frequency applications due to its superior electron mobility property. In this thesis, InGaAs/InP heterostructure radio frequency MOSFETs are designed, fabricated, and characterized. Various spacer technologies, from high dielectric spacers to air spacers, are implemented to reduce parasitic capacitances, and f_T/f_{max} are evaluated. Three types of RF MOSFETs with different spacer technologies are fabricated in this work.

InP ∧-ridge spacers are integrated on InGaAs Nanowire MOSFET in an attempt to decrease parasitic capacitances; however, due to a high-dielectric constant of the spacers and smaller transistors transconductance, the f_T/f_{max} are limited to 75/100 GHz. InGaAs quantum well MOSFETs with a sacrificial amorphous silicon spacer are fabricated, and they have capacitances of a similar magnitude to other existing high-performing RF InGaAs FETs. An 80 nm InGaAs MOSFET has f_T/f_{max} = 243/147 GHz is demonstrated, and further optimization of the channel and layout would improve the performance. Next, InGaAs MOSFETs with nitride spacer are fabricated in a top-down approach, where the heterostructure is designed to reduce contact resistance and thus improve transconductance. In the first attempt, from the electrical characterization, it is concluded that the ON resistance of these MOSFETs is comparable to state-of-the-art HEMTs. Complete non-quasi-static small-signal modeling is performed on these transistors, and the discrepancy in the magnitude of f_{max} is discussed. InGaAs/InP 3D-nanosheet/nanowire FETs' high-frequency performance is studied by combining intrinsic analytical and extrinsic numerical models to estimate f_T/f_{max}. 3D vertical stacking results in smaller parasitic capacitances due to electric field perturbance because of screening.

An 8-band $\mathbf{k} \cdot \mathbf{p}$ model is implemented to calculate the electronic parameters of strained $In_xGa_{1-x}As$/InP heterostructure-based quantum wells and nanowires. Bandgap, conduction band energy levels, and their effective masses and non-parabolicity factors are studied for various indium compositions and channel dimensions. These calculated parameters are used to model the long channel quantum well InGaAs MOSFET at cryogenic temperatures, and the importance of band tails limiting the subthreshold slope is discussed.

Keywords:
InGaAs, Nanoelectronics, Radio Frequency

Classification System and/or Index Terms
e.g. Electronic Engineering, Nanotechnology

Language:
English

Supplementary Bibliographical Information:
–

ISBN (printed):
978-91-8039-780-3

Key title and ISSN:
Series of Licentiate and Doctoral Theses; 1654-790X, No. 162

ISBN (digital):
978-91-8039-779-7

Recipient's Notes:

Number of Pages:
174

Price: -

Security Classification:
Unclassified

Signature: G. Navya Sri

Date: 27 September 2023

Radio Frequency InGaAs MOSFETs

Doctoral Thesis

Navya Sri Garigapati

Department of Electrical and
Information Technology

Lund, September 2023

Navya Sri Garigapati
Department of Electrical and Information Technology
Lund University
Ole Römers Väg 3, 223 63 Lund, Sweden

Series of Licentiate and Doctoral Theses
ISSN 1654-790X, No. 162
ISBN 978-91-8039-780-3 (printed)
ISBN 978-91-8039-779-7 (digital)

This thesis is typeset using LaTeX 2_ε with the body text in Palatino and Goudy Initials, headings in Helvetica, text in figures in Arial, and the body text in Appendix B in Computer Modern.

Printed by Tryckeriet i E-huset, Lund University, Lund, Sweden.

To the teachers,
who educate with passion and patience.

Abstract

III-V-based Indium gallium arsenide is a promising channel material for high-frequency applications due to its superior electron mobility property. In this thesis, InGaAs/InP heterostructure radio frequency MOSFETs are designed, fabricated, and characterized. Various spacer technologies, from high dielectric spacers to air spacers, are implemented to reduce parasitic capacitances, and f_T/f_{max} are evaluated. Three types of RF MOSFETs with different spacer technologies are fabricated in this work.

InP $\wedge$-ridge spacers are integrated on InGaAs Nanowire MOSFET in an attempt to decrease parasitic capacitances; however, due to a high-dielectric constant of the spacers and smaller transistors transconductance, the f_T/f_{max} are limited to 75/100 GHz. InGaAs quantum well MOSFETs with a sacrificial amorphous silicon spacer are fabricated, and they have capacitances of a similar magnitude to other existing high-performing RF InGaAs FETs. An 80 nm InGaAs MOSFET has f_T/f_{max} = 243/147 GHz is demonstrated, and further optimization of the channel and layout would improve the performance. Next, InGaAs MOSFETs with nitride spacer are fabricated in a top-down approach, where the heterostructure is designed to reduce contact resistance and thus improve transconductance. In the first attempt, from the electrical characterization, it is concluded that the ON resistance of these MOSFETs is comparable to state-of-the-art HEMTs. Complete non-quasi-static small-signal modeling is performed on these transistors, and the discrepancy in the magnitude of f_{max} is discussed. InGaAs/InP 3D-nanosheet/nanowire FETs' high-frequency performance is studied by combining intrinsic analytical and extrinsic numerical models to estimate f_T/f_{max}. 3D vertical stacking results

in smaller parasitic capacitances due to electric field perturbance because of screening.

An 8-band $\mathbf{k} \cdot \mathbf{p}$ model is implemented to calculate the electronic parameters of strained $In_xGa_{1-x}As/InP$ heterostructure-based quantum wells and nanowires. Bandgap, conduction band energy levels, and their effective masses and non-parabolicity factors are studied for various indium compositions and channel dimensions. These calculated parameters are used to model the long channel quantum well InGaAs MOSFET at cryogenic temperatures, and the importance of band tails limiting the subthreshold slope is discussed.

Popular Science Summary

Imagine a world without smartphones, computers, or modern electronics. Our lives have been drastically transformed due to the great and powerful invention of the transistor. A *transistor* is a tiny device that controls the flow of electricity between two contacts and acts as a switch. It is like the traffic cop of the electronic world, directing the flow of electrical signals. Before transistors existed, electronic devices used bulky vacuum tubes like old-fashioned light bulbs with heated filaments. These tubes were bulky, expensive, and consumed a lot of power. However, in the late 1940s, the invention of the transistor changed everything.

The transistor is incredibly small, and this unique feature made it possible to create compact and lighter electronic devices for our daily lives. Without transistors, the size of a smartphone would be bigger than a brick. Not only did transistors make devices smaller, but they also made them more reliable and efficient. They allowed us to build computers that were faster, smarter, and capable of handling complex tasks. The birth of microprocessors, which are like the brains of computers, was made possible by transistors. These powerful chips can perform billions of calculations in a second, making our modern digital world a reality.

In this work, high-frequency or radio-frequency transistors, the backbone of radio communication, are designed, fabricated, and characterized. These transistors are built using advanced materials and techniques to handle the demanding requirements of high-frequency applications. They are like super-charged versions of regular transistors, capable of operating at frequencies that range from tens of megahertz to several gigahertz. The increase in

maximum operating frequency is attributed to transistor size scaling, mainly the distance the carrier travels between two contact leads. However, over time, due to limited carrier mobility of the channel material, i.e., how fast the carriers move in the channel, and dominant external parasitics, i.e., the outer part of the transistor contributing significant unwanted capacitances and resistances, have started to limit the high-frequency performance.

Indium gallium arsenide (InGaAs) compound semiconductor material is a newly explored material possible to replace Silicon due to its carrier mobility $\times 10$ that of the latter, and thus, InGaAs high-frequency transistors can handle higher frequency signal with low power consumption. The next challenge is reducing the external parasitics, which dominate at smaller transistor sizes. The low-κ spacer technology is introduced to get smaller parasitic capacitances without degrading resistance. In this work, InGaAs metal oxide semiconductor field effect transistors (MOSFETs) are a type of transistor where the semiconductor channel is separated from the gate metal contact by high-dielectric oxide. Various spacers with a high dielectric constant to the least possible dielectric constant of 1, i.e., an air spacer, are implemented to reduce the external parasitics.

InGaAs nanowire MOSFETs, where the gate is covered on three sides of the channel, and two different quantum well MOSFETs are designed, fabricated, and characterized. The device's electrical characterization includes DC and high-frequency from 10 MHz to 67 GHz using a vector network analyzer. The importance of accurate mathematical modeling of the measured high-frequency data in estimating the f_T/f_{max} is highlighted. These advancements promise faster data processing, more efficient wireless networks, and exciting possibilities in emerging technologies like 5G, autonomous vehicles, and the Internet of Things (IoT).

Acknowledgements

I want to express my sincere gratitude and appreciation to all those who have supported and assisted me throughout my doctoral journey, without whom this thesis would not have been possible.

First and foremost, I am deeply indebted to my supervisor, *Prof. Erik Lind*, for his guidance, expertise, and unwavering support. Thank you for giving me this incredible opportunity and believing in me. Your invaluable insights, constructive feedback, and patience have been instrumental in shaping this research and my overall growth as a scholar. I am truly grateful for your mentorship and dedication. I thank my co-supervisors, *Johannes Svensson* and *Mattias Borg*, for their support and guidance. *Prof. Lars Erik Wernersson*, your vision and leadership qualities always inspire me. *Fredrik*, thank you for introducing me to the clean room; you taught me how to hold a tweezer to fabricate the transistor to characterization. Lars and Stefan, thank you for teaching and helping me with the RF measurement set-up. *Lasse* and *Patrik*, thank you for collaboration.

Furthermore, I would like to acknowledge wonderful colleagues, *Gautham*, *Heera*, *Anton*, *Saketh*, *Michael*, *Abinaya*, *Robin*, *Marcus*, *Hannes*, *Andre*, *Louise*, thank you for creating such a learning environment and your all inspiring work. I thank the administrative staff at the Electrical and Information Technology for their support, resource provision, and logistical assistance. I also sincerely thank the NanoLund lab staff for promptly assisting with the tools.

My family, Amma, Nanna, and Annayya, thank you for your unconditional love, understanding, and encouragement throughout this demanding academic journey. To my Srinu brother and Vadina, thank you so much for your support during the difficult times of my academic career, and I am always grateful to you both. To my remaining family, thank you for showering me

with love and care. I am always thankful for having an amazing and loving big family. I look forward to spending more time with you all.

Finally, I thank my dearest *Bangu*.

G. Naya Sri

Lund, Sept 2023

Contents

Preface

This thesis is the culmination of more than five years of work in the *Electromagnetism and Nanoelectronics* group at Lund University and presents detailed studies of very fascinating topics. The work was supervised by Prof. *Erik Lind*.

STRUCTURE OF THE THESIS

- **INTRODUCTION**
  The main body of the thesis consists of the publications appended in the back. The Introduction provides a broader and more comprehensive view than the very focussed publications and ties their work together. The Introduction is intended to be comprehensible for aspiring researchers with a Master's degree in physics or a related subject.

- **APPENDICES**

  A k·p Theory
  Appendix A provides the $In_xGa_{1-x}As$ and InP material parameters and Hamiltonian parameters used in the $\mathbf{k} \cdot \mathbf{p}$ model formalism. Also, $\mathbf{k} \cdot \mathbf{p}$ and strain matrices of bulk and confined structures are given.

  B Fabrication Details
  Appendix B provides detail process flow summary of InGaAs quantum well RF MOSFETs with nitride spacers.

- **PAPERS**
  The papers forming the main body of the thesis are reproduced in the back and listed in the following.

LIST OF PUBLICATIONS

The following papers form the main body of this thesis and the respective published or draft versions are appended in the back.

Paper I: N. S. Garigapati, F. Lindelöw, L. Södergren, and E. Lind, "Capacitance Scaling in $In_{0.71}Ga_{0.29}As/InP$ MOSFETs With Self-Aligned a:Si Spacers", *IEEE Transactions on Electron Devices*, vol. 68, pp. 3762-3767, Aug 2021, doi: 10.1109/TED.2021.3092299.

► *I fabricated the devices, carried out all measurements and analysis and wrote the paper.*

Paper II: F. Lindelöw, N. S. Garigapati, L. Södergren, M. Borg and E. Lind, "III-V nanowire MOSFETs with novel self-limiting ∧-ridge spacers for RF applications", *Semiconductor Science and Technology*, vol. 35, no. 6, pp. 065015, May. 2020, doi:10.1088/1361-6641/ab8398.

► *I fabricated the devices, participated in measurement, analysis and paper writing.*

Paper III: N. S. Garigapati, L. Södergren, P. Olausson, and E. Lind, "Strained $In_xGa_{(1-x)}As/InP$ near surface quantum wells and MOSFETs", *Applied Physics Letters*, vol. 120, no. 9, pp. 092105, Feb. 2022, doi:10.1063/5.0073918.

► *I performed all modeling, analysis and wrote the paper.*

Paper IV: N. S. Garigapati and E. Lind, "8-band k·p modeling of strained $In_xGa_{(1-x)}As/InP$ heterostructure nanowires", *Journal of Applied Physics*, vol. 133, no. 1, pp. 015701, Jan. 2023, doi:10.1063/5.0133229.

► *I performed complete modeling, analysis and wrote the paper.*

Paper V: N. S. Garigapati and E. Lind, "RF performance comparison of III-V InAs/InP Quantum Well, 3D-Nanosheet/Nanowire MOSFETs", *IEEE Transactions on Electron Devices*, Aug. 2023, TED-2023-08-2293-R (Under review)

► *I performed complete device design, modeling, analysis and wrote the paper.*

Paper VI: N. S. Garigapati and E. Lind, "High-Frequency small-signal modeling of CC Quantum Well InGaAs MOSFETs", (Manuscript)

► *I desinged, developed the process flow and fabricated the devices, carried out all measurements and data analysis and wrote paper.*

RELATED WORK

The following publications are not included in the thesis but summarised related work I contributed.

JOURNAL PAPERS

Paper vii: L. Södergren, N.S. Garigapati, Mattias Borg, and E. Lind, "Mobility of near surface MOVPE grown InGaAs/InP quantum wells", *Applied Physics Letters*, vol. 117, no. 1, pp. 013102, July. 2020, doi: 10.1063/5.0006530.

CONFERENCE CONTRIBUTIONS

Paper viii: N. S. Garigapati and E. Lind, "Self-aligned InGaAs composite channel MOSFET with f_T = 207 GHz", *Swedish Microwave days*, May. 2023.

Paper ix: N. S. Garigapati and E. Lind, "Fabrication of Self-aligned Quantum Well InGaAs MOSFETs for High-Frequency Applications", *EMRS Spring meeting*, May-Jun. 2023.

Acronyms and Symbols

ACRONYMS AND ABBREVIATIONS

AC	Alternating current
Al_2O_3	Aluminum oxide
ALD	Atomic layer deposition
BJT	Bipolar junction transistor
BOE	Buffered oxide etch
CB	Conduction band
CBO	Conduction band offset
CMOS	Complementary metal oxide semiconductor
DC	Direct current
DE	Digital etch
DG	Dummy gate
DIW	Deionized water
DOS	Density of states
DUT	Device under test
EBL	Electron beam lithography

EOT	Effective oxide thickness
FDSOI	Fully depleted silicon on insulator
FET	Field-effect transistor
GAA	Gate-all-around
GaAs	Gallium arsenide
H_2O_2	Hydrogen peroxide
H_3PO_4	Phosphoric acid
HCl	Hydrochloric acid
HEMT	High-electron-mobility transistor
HF	Hydrofluoric acid
HfO_2	Hafnium dioxide
hh	Heavy-hole
HSQ	Hydrogen silsequioxane
ICP	Inductively coupled plasma
II	Impact ionization
In	Indium
InAlAs	Inidum aluminum arsenide
InAs	Indium arsenide
InGaAs	Indium gallium arsenide
InP	Indium phosphide
IPA	Isopropyl alcohol
ITRS	International technology roadmap for semiconductors
lh	Light-hole
LNA	Low noise amplifier
MAG	Maximum available gain
MOCVD	Metal-organic chemical vapor deposition

MOSFET	Metal-oxide-semiconductor field-effect transistor
MOVPE	Metalorganic vapor-phase epitaxy
MSG	Maximum stable gain
NH_4S	Ammonium sulfide
NSFET	Nanosheet field effect transistor
NSQ	Non-quasi-static
NWFET	Nanowire field effect transistor
Pd	Palladium
PDA	Post deposition annealing
PECVD	Plasma enhanced chemical vapor deposition
PMMA	Polymethyl methacrylate
QCL	Quantum capacitance limit
QW	Quantum well
RF	Radio frequency
RIE	Reactive ion etching
S.I	Semi-insulating
S/D	Source/Drain
SEM	Scanning electron microscope
Si	Silicon
SiO_2	Silicon dioxide
Sn	Tin
SOI	Silicon on insulator
S-parameter	Scattering parameter
SO	Spin-orbit
SPL	Single pixel line

SS	Subthreshold swing
Ti	Titanium
TLM	Transfer length method
TMAH	Tetramethylammonium hydroxide
TMRR	Through-match-reflect-reflect
TRL	Through-reflect-line
VB	Valence band
VBO	Valence band offset
VLSI	Very large scale integration
VNA	Vector network analyzer
W	Tungsten
Y-parameter	Admittance parameter
Z-parameter	Impedance parameter
ZrO_2	Zirconium oxide

LATIN SYMBOLS

$\mathbf{C_{xx}}$	F, fF/μm	Trans capacitances in FET, where x = {S, G, D}, often normalised by total gate width.
$\mathbf{C_{it}}$	F, F cm^{-2}	Interface trap capacitance
$\boldsymbol{D_{it}}$	cm^{-2}eV^{-1}	Interface trap density
$\boldsymbol{E_F}$	eV	Fermi energy level
$\boldsymbol{E_g}$	eV	Bandgap
f	Hz	Frequency
f_T	Hz	Transition frequency
f_{max}	Hz	Maximum oscillation frequency
g_{me}	S	Extrinsic transconductance

$g_{\mathbf{de}}$	S	Extrinsic output conductance
$g_{\mathbf{mi}}$	S	Intrinsic transconductance
$g_{\mathbf{di}}$	S	Intrinsic output conductance
$\hbar$	Js	Reduced Planck constant and $\approx 1.055 \times 10^{-34}$ Js
$\mathbf{h}$	Js	Planck constant and $\approx 6.626 \times 10^{-34}$ Js
$\mathbf{h_{21}}$	dB	Current gain
$I_{\mathbf{D}}$	A, mA µm^{-1}	Drain Current, often normalised by the gate width
$I_{\mathbf{DS}}$	A, mA µm^{-1}	Drain-to-Source current, often normalised by the gate width
$I_{\mathbf{G}}$	A, mA µm^{-1}	Gate current, often normalised by the gate width
$I_{\mathbf{S}}$	A, mA µm^{-1}	Source current, often normalised by the gate width
$k_{\mathbf{B}}$	J/K	Boltzmann constant and $\approx 1.381 \times 10^{-23}$ J/K
k	m^{-1}	Wave vector
K		Stability factor
$L_{\mathbf{G}}$	m	Gate length
m_0		$\approx 9.109 \times 10^{-31}$ kg, Electron rest mass
m^*	m_0	Effective mass
N_d	cm^{-3}	Donor doping density
q	C	$\approx 1.602 \times 10^{-19}$ C, Elemental charge
T	K or °C	Temperature
U	dB	Mason's unilateral power gain
$V_{\mathbf{DS}}$	V	Drain-to-Source voltage
$V_{\mathbf{GD}}$	V	Gate-to-Drain voltage

V_{GS}	V	Gate-to-Source voltage
V_T	V	Threshold voltage
x		Indium composition

GREEK SYMBOLS

α	eV^{-1}	Nonparabolicity factor
ε_r		relative permitivitty
ε_0	F/m	Free space permitivitty and $\approx 8.85 \times 10^{-12}$ F/m
Γ		Center of the first Brillouin zone
λ_n	m	Natural length
λ_{MF}	m	Average mean free path
$\mu_{e/h}$	m^2/Vs	Electron/hole mobility
ω	$rads^{-1}$	Angular frequency
ω_0	$rads^{-1}$	Angular trap cut-off frequency
ϕ		Eigenstate
Ψ_s	eV	Surface potential
σ		Poisson's ratio

INTRODUCTION

1 Introduction

This chapter discusses the history of transistor invention and how transistor size scaling has shaped the electronics industry. Various radio frequency transistors are introduced, which are the fundamental components for today's mmWave applications like high-speed mobile networks, data transfer, and high-performance radar. The motivation for this thesis work and thesis outline are discussed at the end of this chapter.

1.1 TRANSISTOR HISTORY

Vacuum tubes were one of the earliest electronic devices, where the electric charge is transported between the filament and charged plate in a vacuum. They saw rapid development and applications in various fields, like signal amplification and modulation, computing, and radio broadcasting, to name a few [1]. However, they had several drawbacks because of their large size and significant power consumption. The invention of the 3-terminal device called *transistor* in the late 1940s marked a significant advancement in electronic technology. The first point-contact transistor was invented at Bell Labs by John Bardeen, Walter Brattin, and William Shockley 1948 [2] for which the three were awarded the Nobel prize later in 1956. The transistor is a 3-terminal device, where the third terminal controls the current flow between the other two terminals. Transistors were smaller, more reliable, and consumed less power than vacuum tubes. They later became the fundamental building blocks of modern electronic devices, including computers, smartphones, televisions, and countless other technologies. They have a vital impact on the progression of technology and the formation of our contemporary world.

The development of integrated circuits (ICs), which combine multiple transistors, resistors, and capacitors on a single semiconductor chip, revolutionized the miniaturization and increased functionality of electronic systems. In 1960s, metal-oxide-semiconductor field-effect transistors (MOSFETs) were introduced, offering improved performance, lower power consumption, and compatibility with integrated circuits. Complementary metal-oxide-semiconductor (CMOS) technology, which uses both p-type and n-type MOSFETs, emerged as the dominant design for integrated circuits due to its low power consumption and high noise immunity. Since then, transistors have continued evolving, becoming smaller, faster, and more efficient. CMOS technology has become more powerful by integrating them with real-world interaction. New exciting applications in radio-frequency, power, sensors, and photonics have emerged, and the performance demands of these types of transistors differ from logic devices. The International technology roadmap for semiconductors (ITRS) [3] has released a white paper discussing the performance requirements and naming it as *more than Moore* demands.

1.1.1 TRANSISTOR SCALING

Moore's law [4] has been a driving force for the electronic industry; by this law, the transistor density on a semiconductor chip doubles every two years. This law is realized by decreasing the device area by 50%, i.e., gate length and width are decreased by $\times 0.7$. Figure 1.1 illustrates the transistor geometries that kept the transistor scaling alive over the last 50 years and plausible future novel device geometries that would extend Moore's law. Dennard scaling [5] sets the transistor scaling based on constant field scaling, where the drive voltage, oxide thickness, and channel dimensions are scaled equally to improve electrical performance and maintain constant power density. Advancement in manufacturing tools, mainly lithography, has helped in decreasing the transistor's feature size, or in other words, process node, as per the industry standard. This process node number has been more or less the same as the gate length, which is the smallest feature size of MOSFET. Chip density has doubled in each generation by halving gate length, and more processors are made in the same chip area, which works faster due to minor RC delay. In the early 2000s, devices entered power constrained regime, where power density increased to 100 W/cm^2 [6], and more effort in the system-level design was advised. Also, further scaling the gate length below 30 nm, transistor behavior is lost due to strong short channel effects (SCE), and this has encouraged the researchers to look for new device architectures and channel materials.

As a continuation of Moore's law, chip density can be increased by width folding even at more considerable gate lengths, as demonstrated in FinFET 22 nm technology [7], where the decoupling between the gate length and

process node has started. FinFETs have excellent electrostatics due to tri-gate architecture and a larger effective width for a small physical footprint. At smaller gate lengths, very high aspect ratio fins are required; for example, for a gate length of 6 nm, required fin height and width are 65 and 6 nm, respectively [8], and realizing taller fins have their process difficulties. Recently, gate-all-around (GAA) stacked nanosheet or nanowire FETs have been proposed, and IBM was able to realize a 2 nm technology node with nanosheet transistors [9]. Nanosheet transistors with vertical stacking have a larger effective width and drive current; the latter is not limited by the device footprint, unlike FinFETs. In addition to this, they also have excellent gate electrostatics due to the GAA gate structure. Nanowire transistors are similar to nanosheet transistors; however, drive current is limited due to smaller channel dimensions. Nanosheets and nanowires are slightly different; in the former, the width is more than $\times 2$ the thickness, whereas, in nanowires, the width and thickness are comparable. Even in GAA transistors, short channel effects emerge if the ratio between channel length and width is less than 5 [8]. Another GAA transistor geometry is the vertical nanowire transistors, where the device footprint and gate length are decoupled, and gate length is not limited by standard lithography but by the spacer thickness.

Vertical stacking technique has been used in flash memories [10] and has not been implemented for logic applications. Stacking different layers is one of the alternatives to increase the chip density and regain Moore's law. Another alternative approach for device improvement is changing the channel material with high mobility material, like strained Si or indium-rich InGaAs, where higher drive currents are realized with smaller drive voltages.

1.2 III-V CHANNEL MATERIAL

An alternate approach is to switch to larger mobility channel materials, which would deliver higher current density at smaller operating voltages. III-V materials have a wide range of applications due to their direct bandgap nature and superior high electron mobility, such as optoelectronic devices like lasers, LEDs, high-speed wireless networks, and high-frequency radar, to name a few [11]. III-V-based InGaAs as n-channel transistors have electron mobility higher than 10,000 cm^2/V-s in sub-100 nm regime [12]. InGaAs-based planar and nonplanar MOSFETs have exhibited record transconductance and > 3 $mS/\mu m$ [13–15]. Unlike Si, III-V-based InGaAs material does not have defect-free native oxide, thus limiting achieving the ideal minimum subthreshold swing (SS_{min}). A remarkable effort is put into the high permittivity gate oxides, which exhibited the same order of interface trap density as that of Si/SiO_2 [16,17].

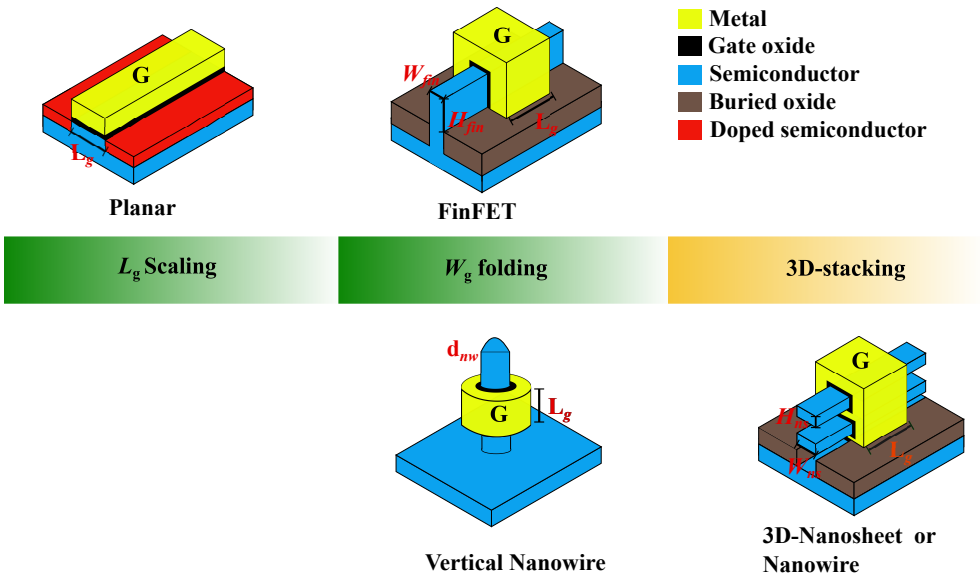

Figure 1.1: Transformation of transistor configurations caused by the scaling of transistor dimensions.

1.3 RADIO FREQUENCY TRANSISTORS

Radio frequency (RF) devices are the ones that operate in the frequency range of 3 kHz to 300 GHz. They have applications ranging from new-generation cellular networks (e.g., 5G and 6G) to long-range Wi-Fi networks and from global positioning systems to high-frequency radar and automotive sensing, to name a few [18, 19]. The development of radio frequency devices has come a long way from the Germanium bipolar transistor operating at a frequency of 1 GHz back in 1958 to III-V-based high-electron-mobility transistors (HEMTs) with oscillation frequency greater than 1 THz [20, 21]. The prominent figure of metrics of high-frequency performance is transition frequency (f_T), where the current gain of the device is 1, and maximum oscillation frequency (f_{max}), where the maximum available gain (MAG) of the device is unity or 0 dB. Another important small-signal and high-frequency metric of the transistor is the minimum noise figure (NF_{min}), which is the minimum noise of the transistor for the optimized source impedance. Output power (P_{out}) is the transistor's power handling capacity at high frequencies or the frequency of interest. The output power requirements are well diversified, ranging from a few mW in the case of Bluetooth to 100's of W in power base stations [8]. RF transistor performance requirements depend on the interested mmWave application; hence, transistor technology has become application-specific.

III-V-based HEMTs are one of the industry standards for radio frequency devices because of their superior electrical and high-frequency performance with $f_T > 650$ GHz [22, 23]. The epitaxial structure of HEMT involves a wide bandgap barrier layer with delta doping in contact with the narrow bandgap channel, leading to higher channel carrier mobility and, hence, transconductance. The wide bandgap barrier, which separates the gate contact from the channel, should be scaled down as the gate length scales to maintain the gate electrostatics, which resulted in significant gate-leakage under sub-100 nm gate length devices [24]. Another challenge is the dominant parasitic capacitance for gate length less than 30 nm [23] and larger source resistance [22], which requires optimization of epitaxial layers in the contact stack.

In the early 2000's, the RF performance of Si MOSFETs has been enhanced considerably. They have delivered f_T of 445 GHz with 32 nm gate length [25], given their smaller electron mobility, the larger density of states, and larger parasitic capacitance due to their closely spaced highly doped contacts. Fully depleted silicon on insulator (FDSOI), where the Si channel is placed on buried oxide, has also exhibited promising RF figure of metrics, with f_T and f_{max} of 375 and 290 GHz in 22 nm technology [26]. FDSOI FETs would have smaller parasitic capacitances than planar Si MOSFET due to buried oxide. Theoretical RF performance assessment of 28 nm UTBB FDSOI MOSFETs is presented in [27], and the effect of parasitic capacitances is discussed mainly.

Non-planar device technologies like tri-gate FinFETs have high intrinsic gain due to excellent gate electrostatics. Fin height and spacing between the fins are optimized for better high-frequency performance. Si 14 nm FinFET technology has delivered f_T over 250 GHz at lower gate overdrive voltage [28] and f_T = 314 GHz [29]. However, taller and thinner fins are demanded for sub-10 nm technology, where mobility is degraded due to surface scattering, and taller fins would increase the parasitic capacitance.

III-V-based lateral and vertical nanowires have drawn attention with their larger transconductance and excellent gate control over the channel. Lateral nanowire MOSFETs have reported f_{max} of 400 GHz [30] and vertical nanowire MOSFETs [31] have exhibited f_T and f_{max} both over 140 GHz, where optimization of the gate contact and spacer thickness could improve high-frequency performance.

Finally, we have III-V-based InGaAs MOSFETs, which manifested similar transconductance as that of HEMTs, and, however, poor RF metrics compared to the latter due to large parasitic capacitances. InGaAs quantum well MOSFET on InP has exhibited f_T/f_{max} = 357/410 GHz [32] and 20 nm InGaAs MOSFET on Si has exhibited record f_T/f_{max} = 370/310 GHz [33]. However, these FETs' parasitic capacitance is larger than HEMTs, and efficient spacer designs must be investigated to optimize and improve high-

frequency performance. III-V-based MOS-HEMTs, which inherit the HEMTs heterostructure design to provide high-mobility and high-κ gate oxide from MOSFETs to decrease gate leakage current, have exhibited excellent high-frequency metrics [34–36]. In [37], $In_{0.8}Ga_{0.2}As$ MOS-HEMT with gate length, L_g = 36 nm has exhibited $f_{max} > 1$ THz.

1.4 MOTIVATION

From the earlier discussion on FET-based transistors for high-frequency applications, it can be concluded that parasitic capacitance and resistances are limiting factors, and the trade-off between these two parameters made the problem even more challenging. A low-κ dielectrics (ideally, air spacer) spacer introduction into the device design without a substantial increase in resistance is the main goal in designing the RF FETs. In this work, InGaAs/InP heterostructure-based quantum confined 1D nanowires and 2D quantum well FETs are designed, fabricated, characterized, and modeled. MOSFETs are fabricated from the bottom-up approach, where the channel and highly doped contacts are selectively regrown, and in the top-down approach, where the epitaxial layers are etched down to define the device. Various spacers like delta-doped InP $\wedge$-ridge spacers, a:Si, and nitride spacers are used to reduce parasitic capacitances. The main goal is to realize a spacer technology to achieve parasitics similar to HEMTs and improve the high-frequency metrics, mainly f_T/f_{max}.

1.5 THESIS OUTLINE

The introduction chapter provides a brief transistor evolution history and challenges to regain Moore's law. Various radio frequency transistors overview and their challenges are discussed in brief, and ended the discussion with the motivation for the thesis work.

In Chapter 2, an introduction to 8-band $\mathbf{k} \cdot \mathbf{p}$ theory is given, starting from the stationary Schrödinger equation and $\mathbf{k} \cdot \mathbf{p}$ matrix formation. Spin-orbit and strain matrices for bulk semiconductors are defined. Envelope function approximation explains how the bulk Hamiltonian matrices are modified for confined structures like $In_xGa_{1-x}As$/InP 2D quantum well and 1D nanowires. At this chapter's end, the discussion focuses on implementing $\mathbf{k} \cdot \mathbf{p}$ in the COMSOL Multiphysics FEM solver. Papers III and IV summarise the results on electronic parameters behavior of strained heterostructure quantum well and nanowire structures, respectively.

In Chapter 3, MOSFET DC and the high-frequency small-signal model are discussed in detail. The electrical performance of InGaAs/InP heterostructure-based planar, tri-gate nanowire, and gate-all-around FETs are modeled using the top-of-the-barrier model in combination with the Landauer model. FET high-frequency small-signal model, including impact ionization, band-to-band tunneling, and high-κ gate oxide loss, is presented, and extrinsic and intrinsic parameter extraction is explained. Various high-frequency gains, like unity power gain, maximum available gain, and maximum stable gain, are discussed in detail at the end of this chapter. The discussed transistor model is used to design the 3D-vertically stacked horizontal nanosheet/nanowire MOSFETs, and their RF metrics are estimated. Paper V summarises the results on InAs/InP 3D nanosheet/nanowire FETs RF performance.

In Chapter 4, the main fabrication steps in realizing the RF MOSFET are discussed, and various RF MOSFET process designs implemented in this work are presented in brief.

Chapter 5 discusses the electronic parameter behavior of $In_xGa_{1-x}As$/InP quantum well and nanowire and long channel quantum well MOSFET I-V and C-V modeling. This is followed by the discussion on the high-frequency characterization of various RF MOSFETs designed and implemented in this thesis work. Paper I, II, and VI details the various RF MOSFETs realized in this work.

A summary of the thesis work, followed by the potential future continuation of the presented work, is given in Chapter 6.

Appendix A presents 8-band $\mathbf{k} \cdot \mathbf{p}$ and strain Hamiltonian matrices of bulk and confined structures. Also, $In_xGa_{1-x}As$ and InP material parameters and Hamiltonian parameters expressions are provided.

In Appendix B, fabrication details of the InGaAs quantum well MOSFET with nitride spacer are given in more detail.

2

$\mathbf{k} \cdot \mathbf{p}$ Theory

A comprehensive understanding of the confined channel electronic parameters, like conduction band subband energies, effective masses, and non-parabolicity factors, is necessary for precise transistor modeling. Additionally, the employed theoretical modeling helps to provide insights into the experimental data. This chapter presents a detailed 8-band $\mathbf{k} \cdot \mathbf{p}$ model to calculate the strained heterostructure electronic parameters. The implementation of the developed model in the COMSOL Multiphysics FEM solver and results, including insights into the inhomogeneous strain distribution in the 1D nanowire, is presented in the end. This $\mathbf{k} \cdot \mathbf{p}$ model is inspired from [38], and more details can be found there.

2.1 SCHRÖDINGER EQUATION

In the realm of nanometer-scale semiconductors, quantum mechanics takes precedence as it governs the evolution of quantum states within the system, delineated by the single particle real space time-dependent Schrödinger equation,

$$\left(-\frac{\hbar^2}{2m}\nabla^2 + V(r,t)\right)\psi(r,t) = i\hbar\frac{\partial\psi(r,t)}{\partial t} \tag{2.1}$$

Where m is the particle mass, $\hbar$ is the reduced Planck constant, r is the position vector, t is time, $V(r,t)$ is the potential energy. $\psi(r,t)$ corresponds to the wave function of system. The term in the bracket in the above equation (2.1) is known as the Hamiltonian of the system, which constitutes the system's kinetic and potential energy. Solving this equation is not straightforward and

involves assumptions valid in crystalline semiconductors. First, stationary crystal potential is assumed, resulting in time and position-dependent solutions, i.e., $\psi(r,t) = \phi(r)\zeta(t)$. The time-dependent Schrödinger equation becomes,

$$\frac{1}{\phi(r)}\left(-\frac{\hbar^2}{2m}\nabla^2 + V(r)\right)\phi(r) = \frac{1}{\zeta(t)}i\hbar\frac{\partial\zeta(t)}{\partial t} \tag{2.2}$$

This solution is divided into two eigenvalue problems with the same eigenvalue, E. The two equations are as follows,

$$\left(-\frac{\hbar^2}{2m}\nabla^2 + V(r)\right)\phi(r) = E\phi(r) \tag{2.3}$$

$$i\hbar\frac{\partial\zeta(t)}{\partial t} = E\zeta(t) \tag{2.4}$$

The equation (2.3) is famously known as the stationary Schrödinger equation and also the so-called eigenvalue problem for the Hamiltonian, where $\phi(r)$ is the eigenstate with the eigenvalue corresponding to energy, E. The equation (2.4) describes the temporal evolution of the state, and the solution is harmonic and E dependent. This work mainly focuses on solving the stationary Schrödinger equation for bulk and confined structures. Hamiltonian incorporates every electron's energy and their interactions, making it complex and requiring additional assumptions for solutions. The adiabatic approximation allows us to consider a stationary ion core, and thus, electrons in the outer shell experience a stationary potential due to the core. All electrons are assumed to experience the same average potential defined by mean-field theory. In equation (2.3), the potential energy term, $V(r)$ contains the lattice periodic potential due to the crystalline nature of semiconductors, $V_{cr}(r)$ and also any external applied electric field, $V_{ex}(r)$.

Translation symmetry invariance in the crystalline semiconductors leads to the Bloch theorem, which provides the solution to the stationary Schrödinger equation. There exist many solutions for equation (2.3), and they are identified by index, n, and are given by

$$\phi_n(r) = \frac{1}{\sqrt{\Omega_v}}e^{i\mathbf{k}\cdot r}u_{nk}(r) \tag{2.5}$$

Where $\phi_n(r)$ is the so-called Bloch function of the nth state with a wave vector $\mathbf{k}$ in reciprocal space and normalized to the unit cell volume, Ω_v, and $u_{nk}(r)$, is the lattice-periodic function having the same periodicity as the lattice potential. So, the solution of the system is periodic and enveloped by a plane

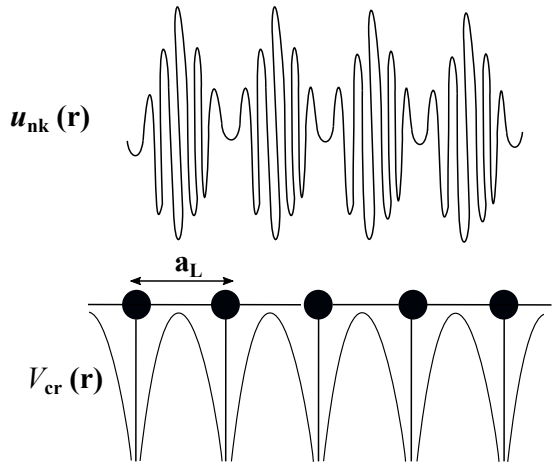

Figure 2.1: A 1D crystalline semiconductor with lattice constant, a_L, and illustration of its periodic crystal potential and the Bloch wave function.

wave with wave vector **k**. The energies that are calculated for every **k** in the reciprocal space form bands, which is the band structure $E_n(k)$ of the system, which is periodic; hence the discussion is limited to the primitive cell in the reciprocal space, or the first Brillouin zone (BZ) [39]. Figure 2.1 illustrates the periodic crystal potential and block wave function of a 1D semiconductor with lattice constant, a_L.

2.1.1 ENVELOPE FUNCTION APPROXIMATION (EFA)

In confined structures, such as quantum wells and nanowires, the translational symmetry breaks, and the wave function is not a plane wave anymore in the confined direction. In a case where the structure is confined in r direction and is free in z direction, the Bloch wavefunction is scaled by slowly varying envelop function $\xi_{nk,m}(r)$. The modified wave function, in general, is given by [40],

$$\phi_{nk}(r,z) = \sum_m e^{i\mathbf{k}\cdot z} u_m(r) \xi_{nk,m}(r) \tag{2.6}$$

The crystal momentum, $\hbar\mathbf{k}$, is no longer continuous in the confined direction and is quantized due to boundary conditions. Thus, the bands are split in a confined direction and form what are known as subbands. The crystal momentum is replaced with its momentum operator, $\hat{p} = -i\hbar\nabla_r$, where ∇_r is defined in the confined r direction. The wave vector becomes an operator [41],

$\hat{\mathbf{k}_\mathbf{r}}$

$$k_r \rightarrow \hat{\mathbf{k}_\mathbf{r}} = -i\frac{\partial}{\partial r} \tag{2.7}$$

The Hamiltonian problem after applying the envelop function approximation becomes,

$$H\xi(r) = E\xi(r) \tag{2.8}$$

Assuming that the lattice constant is smaller than the confined dimension, the spatial probabilistic distribution of the states in the confined direction is given by $|\xi_{n,k}(r)|^2 = \xi_{n,k}(r).\xi^*_{n,k}(r)$. From here on, we will discuss the wave function referring to the envelope function.

2.2 K · P FORMALISM

Solving for the $E_n(k)$ involves many assumptions and approximations, and there are proposed models from simple, effective mass approximation to fully atomistic models. It is always a compromise in choosing the model in terms of complexity, accuracy, and computational cost. This theory was first developed by [42,43]. $\mathbf{k} \cdot \mathbf{p}$ theory is a continuum model with the advantages of including additional symmetry-breaking perturbations such as strain and spin-orbit coupling (SO) and empirically adjusting the parameters to obtain accurate and spurious-free solutions. This theory follows the envelope function approximation to extend the model for nanostructures. It also has a few drawbacks, like the expected decrease in the model's accuracy far from the high symmetry point in the reciprocal space and the complexity of the Hamiltonian for nanostructures. Nevertheless, it has been implemented in many numerical tools to calculate band structure [44–47]. In [48], the band structure analytical expressions are developed using symmetry and experimental data to get high-precision electronic parameter estimation around the $\mathbf{\Gamma}$-point.

The model formalism starts with substituting the Bloch wave function equation (2.5) into the stationary Schrödinger equation (2.3) and introducing the momentum operator. The stationary Schrödinger equation becomes,

$$\left(\frac{\hat{p}^2}{2m_0} + V(r) + \frac{\hbar}{m_0}\mathbf{k} \cdot \hat{\mathbf{p}}\right) u_{nk}(r) = \left(E_n(k) - \frac{\hbar^2\mathbf{k}^2}{2m_0}\right) u_{nk}(r) \tag{2.9}$$

Where the unperturbed Hamiltonian, H_0, is

$$H_0 = \frac{\hat{p}^2}{2m_0} + V(r) \tag{2.10}$$

and the Hamiltonian of the perturbaiton, H_1, is

$$H_1 = \frac{\hbar}{m_0} \mathbf{k} \cdot \hat{\mathbf{p}} \tag{2.11}$$

$\mathbf{k} \cdot \mathbf{p}$ term in the Hamiltonian is why this theory is named $\mathbf{k} \cdot \mathbf{p}$ theory. Using perturbation theory, equation (2.9) can be solved easily. According to this theory, the solution to the perturbed system close to the $\mathbf{k}_0$ is calculated with high accuracy with the known solution of the unperturbed system at $\mathbf{k}_0$. Assuming non-degenerate condition, applying second-order perturbation theory, and performing additional simplifications, the eigenvalue of the total perturbed system becomes,

$$E_n(k) = E_{n0} + \frac{\hbar^2 \mathbf{k}^2}{2m_0} + \sum_{m \neq n} \frac{|\mathbf{k} \cdot p_{mn}|^2}{E_{m0} - E_{n0}} \tag{2.12}$$

Where the momentum matrix element

$$p_{mn} = \langle u_{m0} | \hat{p} | u_{n0} \rangle \tag{2.13}$$

The next step is to choose the basis of the Hamiltonian, which decides how many bands should be considered in modeling the band structure around the symmetry point. In III-V direct narrow bandgap semiconductors, coupling between the remote conduction and valence bands is strong. Thus, one CB and three VBs are considered in calculating the eigenvalues of the total Hamiltonian, and the interaction with the remaining states is neglected. The conduction band is composed by s-type states, $|S\rangle$ and valence band has 3 p-type states, $|p_x\rangle$, $|p_y\rangle$, and $|p_z\rangle$. The spin is ignored for now. The complete derivation of the unperturbed $\mathbf{k} \cdot \mathbf{p}$ Hamiltonian of bulk semiconductor considering four bands can be found in [38], and this Hamiltonian matrix, $H_{kp,4}$ is given in Appendix A. The diagonal elements account for the coupling of states with themselves, and the off-diagonal elements account for the coupling among them. E_c is the conduction band energy, and $E_v^{'}$ is the valence band energy without spin-orbit (SO) coupling. k_x, k_y and k_z are the wave vectors in x, y and z direction. Table A.2 in Appendix A summarises Hamiltonian parameter expression in detail. A_c is related to the curvature of the band structure with remote states coupling included. The interband momentum parameter, P, contains the coupling between the s and 3 p-type states. The three more parameters L, M, N define the valence band states evolution from the unperturbed coupled states. Finally, considering the spin of the states will extend the 4×4 Hamiltonian matrix to 8×8 matrix,

$$H_{kp,8} = \begin{bmatrix} H_{kp,4} & 0 \\ 0 & H_{kp,4} \end{bmatrix} \tag{2.14}$$

2.2.1 SPIN-ORBIT COUPLING (SO)

The spin-orbit interaction arises due to the spin and angular moment coupling, which contributes to the system's total energy. The narrow bandgap semiconductors exhibit strong spin-orbit coupling, which must be considered while calculating the band structure at the conduction band minima. The Hamiltonian for spin-orbit coupling is taken from [38].

$$H_{SO} = \frac{\Delta_{so}}{3} \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -i & 0 & 0 & 0 & 0 & 1 \\ 0 & i & 0 & 0 & 0 & 0 & 0 & -i \\ 0 & 0 & 0 & 0 & 0 & -1 & i & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 & 0 & 0 & i & 0 \\ 0 & 0 & 0 & -i & 0 & -i & 0 & 0 \\ 0 & 1 & i & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \tag{2.15}$$

Δ_{so} is the spin-orbit splitting parameter, an intrinsic material property. The SO perturbation lifts the 6-fold degeneracy at the valence band maxima into 4-fold degeneracy and 2-fold degeneracy separated by Δ_{so}.

2.2.2 STRAIN

The lattice mismatch between the grown epitaxial layer and substrate results in deformation in the grown layer and leads to strain. The linear relationship between strain (ϵ) and stress (σ) is given by Hooke's law [49], and the proportionality constant is called the elastic stiffness tensor. The elastic stiffness tensor for cubic semiconductors is given by,

$$D = \begin{bmatrix} C_{11} & C_{12} & C_{12} & 0 & 0 & 0 \\ C_{12} & C_{11} & C_{12} & 0 & 0 & 0 \\ C_{12} & C_{12} & C_{11} & 0 & 0 & 0 \\ 0 & 0 & 0 & C_{44} & 0 & 0 \\ 0 & 0 & 0 & 0 & C_{44} & 0 \\ 0 & 0 & 0 & 0 & 0 & C_{44} \end{bmatrix} \tag{2.16}$$

The C_{ij} is the stiffness coefficient. Due to the translation symmetry nature of the crystalline semiconductor, most of the elements in D matrix are zero, and only three nonzero independent coefficients (C_{11}, C_{12}, C_{44}) exist. Bardeen introduced the deformation potential theory, which was later developed by

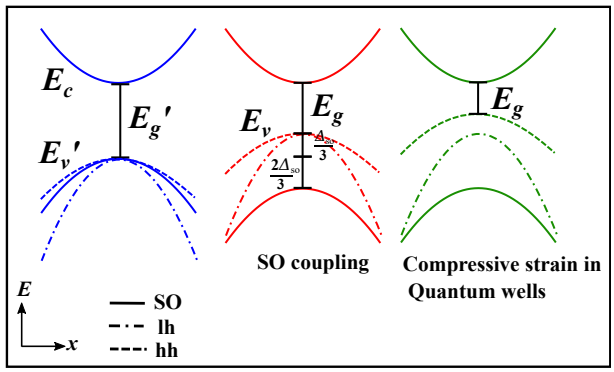

Figure 2.2: From left to right: a simple direct bandgap semiconductor band structure and the SO coupling reduces the 6-fold VB degeneracy to 4-fold degeneracy by pushing the spin-orbit (SO) band down in energy. The light-hole (lh) and heavy-hole (hh) degeneracy are removed when strain is included. For example, in compressively strained quantum wells, hh is pushed higher in energy than lh, and the opposite in the case of tensile strain.

Herring [50, 51], and the same is used to define the strain interaction matrix. The coupling between CB and VB contains a shear strain component, which is quite small and can be neglected and considered only the term containing wave vector. The strain matrix for 3D bulk material, $H_{st,4}$, and parameter expressions are given in Appendix A. The strain coefficients are (l, m, n) and a_c is the deformation potential. Here, as well, including the spin of the states results in 8×8 strain matrix,

$$H_{st} = \begin{bmatrix} H_{st,4} & 0 \\ 0 & H_{st,4} \end{bmatrix} \tag{2.17}$$

The total **k** · **p** Hamiltonian matrices, including the spin-orbit coupling and strain for 3D bulk, are given in Appendix A. Figure 2.2 illustrates the effect of strain and SO coupling on the valence bands and bandgap. The behavior of various valence bands with strain becomes very complex, and the main focus here is to discuss conduction band properties. The $E_v = E_v^{'} + \frac{\Delta_{so}}{3}$, where $E_v = E_c - E_g$ and E_g is the semiconductor bandgap.

The strain matrix elements are all the same for quantum well and nanowires except for strain tensor elements. In quantum wells, the strain in the epitaxial layer is homogeneous, and tensor element calculation is straightforward. In quantum wells, $\epsilon_{21} = 0$, $\epsilon_{12} = 0$, $\epsilon_{11} = \epsilon_{22} = f$, $\epsilon_{33} = -\frac{C_{12}}{C_{11}} f$, where f is the lattice misfit between the epitaxial layer and substrate. The lattice mismatch is defined as $f = \frac{a_L - a_S}{a_S}$, where a_L and a_S are lattice constants of the epitaxial

layer and substrate. In heterostructure nanowires, calculating the tensor elements becomes intricate because of inhomogeneous strain distribution and complex geometry.

2.2.3 HAMILTONIAN OF CONFINED STRUCTURES

Now that the Hamiltonian of the bulk semiconductors is discussed, this subsection presents the transformation of this Hamiltonian for confined nanostructures. Following the EFA, the wave vector, k_r in the confined direction, is replaced by its respective momentum operator, $\hat{\mathbf{k}}_\mathbf{r}$. This transformation results in Hamiltonian with the first and second-order operator terms. The second-order operator ordering concerns the model's numerical stability, mainly spurious solutions. The Burt-Foreman formalism [52] is widely used in the appropriate operator ordering, which maintains the hermiticity of the Hamiltonian. For example, the s-type states coupling term is rewritten as,

$$A_c k_r^2 = k_r A_c k_r \tag{2.18}$$

In the same way, Burt-Foreman summarization is used in ordering the p-type coupling term. The parameter N is divided into N_+ and N_-, obtained from the remaining two parameters, M and L. For example, the term

$$N k_x k_y = k_x N_+ k_y + k_y N_- k_x \tag{2.19}$$

Appendix A discusses the modified Hamiltonian matrices for quantum wells and nanowires in more detail. The Hamiltonian transforms into the partial differential equation system in quantum-confined structures after replacing the wave vector with the momentum operator.

ELIMINATING SPURIOUS SOLUTIONS

In numerical modeling, it is necessary to eliminate spurious solutions. They account for the odd behavior of CB and VB energy levels and forbidden energy levels in the bandgap. A few ways of eliminating the spurious solutions are putting constraints on input parameters to the model or simply discarding the unwanted solutions. It was argued in [38] that the wrong selection of parameters is the reason for spurious solutions and not the choice of mathematical modeling. It is proposed that the main motive is to preserve the ellipticity of the $\mathbf{k} \cdot \mathbf{p}$ operators to get real eigenvalues. In [53], the operator's ellipticity is assured if the following parameter constraints are satisfied.

$$A_c > 0, M - N_- < 0, M + N_- < 0, L - N_- < 0, L + 2N_+ < 0 \tag{2.20}$$

E_P, the optical energy parameter is generally used to fit the experimental data, and it can be varied to satisfy the constraints described in equation (2.20). The momentum interband matrix element P is set by E_P, and the relation is given by

$$E_P = \frac{2m_0}{\hbar^2} P^2 \tag{2.21}$$

The renormalization of other parameters is necessary to produce accurate, effective mass even at reduced E_P values. The E_P value of $In_xGa_{1-x}As$ used in this work is fitted with indium composition, x while satisfying conditions given in equation (2.20):

$$E_p = 17x + 23(1-x) - 4.97x(1-x) \tag{2.22}$$

The remaining $In_xGa_{1-x}As$ and InP material parameter expressions are summarized in Appendix A in Table A.1.

2.2.4 IMPLEMENTATION IN COMSOL MULTIPHYSICS

The heterostructure nanowire is assumed to be infinite in length, which converts the 3D nanowire structure into a more straightforward 2D problem and quantum well into a pure 1D problem. The geometry of these structures is illustrated in Figure 2.3, and boundary conditions used to solve for strain and eigenvalue calculations are highlighted. We use the COMSOL Multiphysics FEM solver and especially the solid mechanic's module to solve for deformation in the heterostructure nanowire. For simplicity, the model uses the linear elastic model and fixed and free boundary conditions to evaluate the strain matrix elements in the case of 1D heterostructure nanowire.

Next, we add the Eigenvalue study to the model in COMSOL Multiphysics and provide the energy value around which the eigenvalues (or subband energies either in CB or VB) are to be found. We can parameterize the wave vector in an unconfined direction to get *E-K* relation at the **Γ**-point from which the electronic parameters are extracted from this relation. Plotting the spatial probability distribution of the eigenstates of the CB and VB and strain tensor elements distribution in different parts of the geometry is possible. Figure 2.4 (a) plots the conduction band eigenstate distribution of InAs/InP heterostructure quantum well with thickness, H = 5 and 10 nm. It is observed that wave function leakage into the InP is larger in thin quantum wells compared to thicker quantum wells. Similarly, In Figure 2.4 (b), conduction band eigenstates of InAs/InP heterostructure nanowires are plotted for various nanowire dimensions.

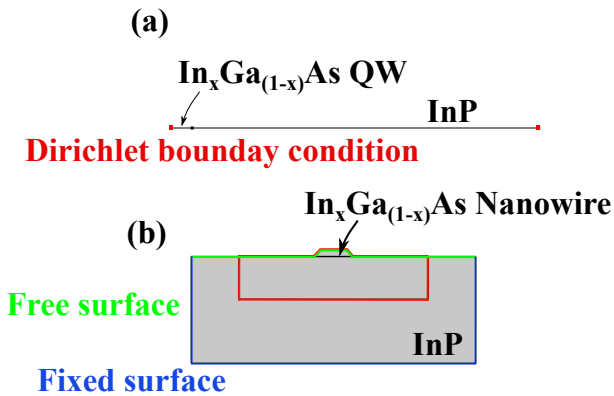

Figure 2.3: The geometry of $In_xGa_{1-x}As/InP$ heterostructure (a) quantum well and (b) nanowire with implemented Dirichlet boundary conditions for calculating eigenvalues are highlighted. Fixed and free boundary conditions are used to calculate the strain tensor elements of the heterostructure nanowire.

STRAIN DISTRIBUTION IN 2D AND 1D $IN_XGA_{1-X}AS$/INP HETEROSTRUCTURE

The main elastic strain tensor components ϵ_{11}, ϵ_{22}, and ϵ_{33} are defined along the X, Y, and Z directions. Strain tensor elements are homogenous in 2D quantum well and are independent of the quantum well thickness, whereas, in heterostructure nanowires, the strain distribution becomes size-dependent. These strain tensors are plotted for InAs/InP nanowires with different sizes in Figure 2.5. The ϵ_{33} is uniform throughout the nanowire irrespective of nanowire size due to infinite nanowire length; in this case, the nanowire is compressed in the axial direction. The magnitude of ϵ_{33} is -0.03, equal to the lattice mismatch at the interface. Further, the inplane elastic strain tensor elements are maximum at the interface and gradually decrease towards the surface of the nanowire due to free surfaces. The lattice mismatch between InP and InAs leads to compression in the X direction, i.e., ϵ_{11} is negative at the interface, and expansion in the Y direction, i.e., ϵ_{22} is positive at the interface. ϵ_{11} and ϵ_{22} in wider nanowires become approximately uniform around the X = 0 axis, and strain relaxes at the top corners of the nanowire. This behavior is reflected in the E_{c1} probability distribution of a nanowire with H = 13 nm and W_B = 80 nm. It is illustrated in Figure 2.4, and the probability distribution is split into two peaks away from the center of the nanowire instead of a single peak located in the center.

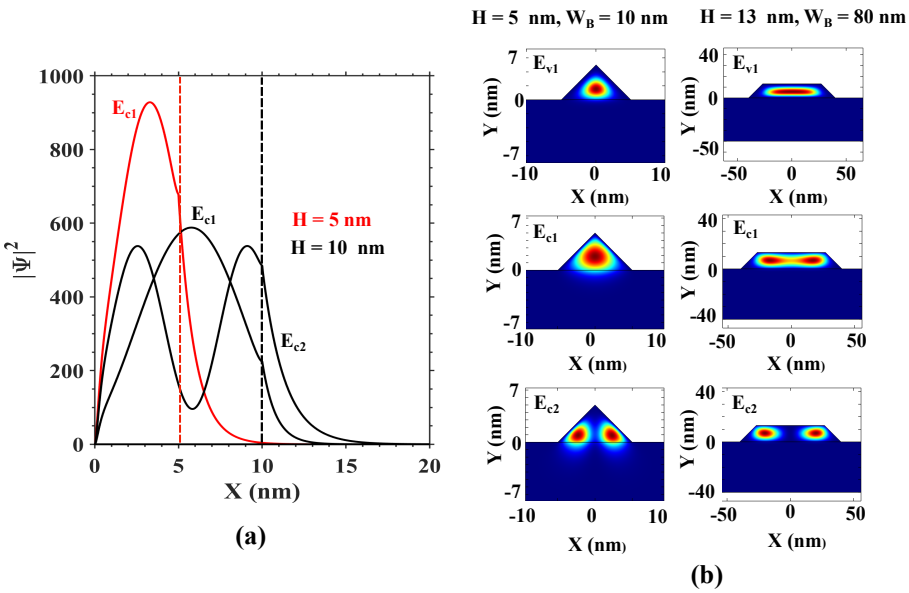

Figure 2.4: (a) InAs/InP heterostructure quantum well with thickness, H = 5 and 10 nm. The available eigenstates probability distributions of the conduction band are plotted as a function of position. (b) The conduction band and valence band eigenmodes of the InAs/InP H = 5 nm, W_B = 10 nm (left) and H = 13 nm, W_B = 80 nm (right).

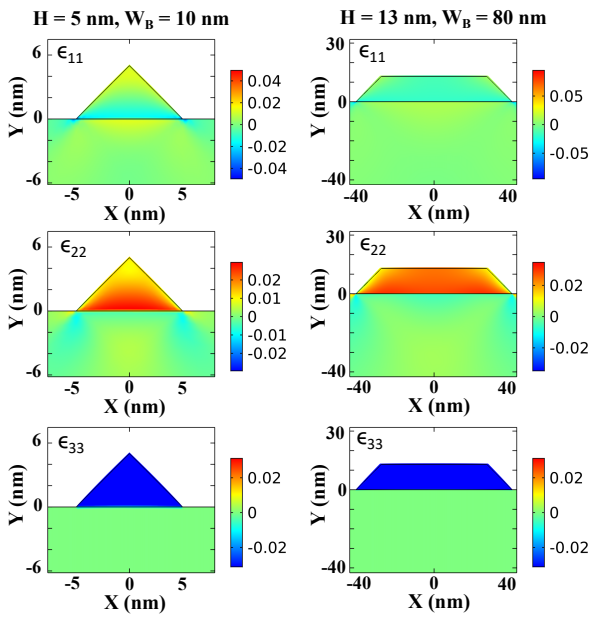

Figure 2.5: The main elastic strain tensor elements, ϵ_{11}, ϵ_{22} and ϵ_{33} are plotted for InAs/InP heterostructure nanowire strain distribution for various heights and widths.

3

MOS Theory

In this chapter, MOSFET DC and small-signal high-frequency models are presented. III-V InGaAs/InP based quantum well, tri-gate, and GAA FETs electrical performance are discussed in the first half of the chapter. An extended high-frequency small-signal model is presented in the second half of the chapter.

3.1 MOSFET BACKGROUND

A MOSFET is a 3-terminal device, where the input gate to source voltage, V_{GS} controls the output drain to source current I_{DS} at a given drain to source voltage V_{DS}. General transfer and output characteristics of a MOSFET are plotted in Figure 3.1(a) and (b), respectively. From the transfer characteristics, the transistor's DC performance metrics such as threshold voltage, (V_T), minimum subthreshold swing (SS_{min}), OFF-current (I_{OFF}) and ON-current (I_{ON}) and transconductance (g_m) are extracted. From the output characteristics, ON resistance (R_{ON}) and output conductance (g_d) are derived. The threshold voltage is the minimum V_{GS} required for the drain current to increase exponentially; in other words, the V_T point separates the off and on-state of the device. It is an important metric for logic devices and their scaling to maintain power efficiency. The gate metal work function, gate oxide bandgap, semiconductor surface potential, and oxide capacitance set V_T. The SS_{min} is the V_{GS} required to increase the drain current by one order (or $\times 10$) at given V_{DS}. In subthreshold region, $I_{DS} \propto e^{\left(\frac{V_{GS}-V_T}{k_B T}\right)}$, where k_B and T are Boltzmann constant and temperature, respectively. The ideal $SS_{min} = ln(10)\frac{k_B T}{q}$ and

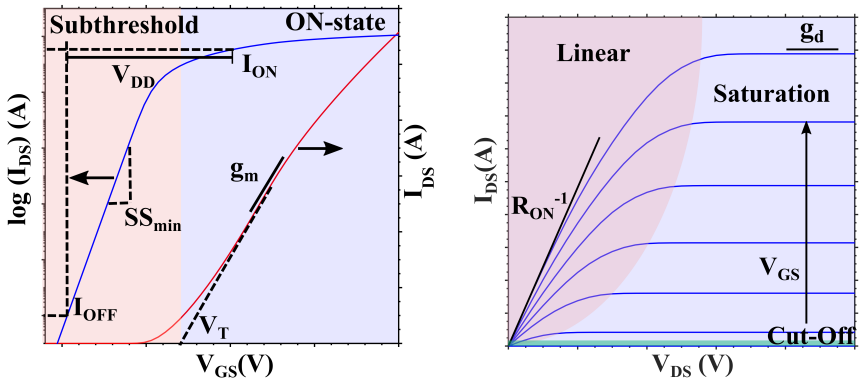

Figure 3.1: (a) Transfer and (b) output characteristics of a MOSFET.

at room temperature, SS_{min} is approximately 60 mV/dec. Larger $\frac{I_{ON}}{I_{OFF}}$ is an important metric for digital applications, and this ratio is decided for a given I_{OFF} and V_{DD} as illustrated in Figure 3.1(a). Next, transconductance is the ability of the FET to convert the small change in the input voltage to a large change in the output current and is defined as $g_m = \frac{\partial I_{DS}}{\partial V_{GS}}$ at a given V_{DS}. It is one of the important metrics for DC and analog transistors and is a measure of the gain of the transistor. MOSFET operating regions are highlighted in Figure 3.1(b). FET is in cut-off region for $V_{GS} - V_T < 0$ and triode or linear region for $(V_{GS} - V_T) > V_{DS}$ and saturation region for $V_{DS} > (V_{GS} - V_T)$. The ON resistance is calculated when FET is operating in the linear region. R_{ON} includes the channel resistance (R_{ch}), which scales linearly with L_g and total access or contact resistance $(2R_c)$. A smaller R_{ON} is required for low power consumption when the transistor is ON. In short gate length devices, large total contact resistance limits R_{ON}. The output conductance, $g_d = \frac{\partial I_{DS}}{\partial V_{DS}}$ is the measure of V_{DS} influence on I_{DS} at a given V_{GS}. Ideally, in MOSFET operating in the saturation region, the drain current is independent of V_{DS}; however, due to the dominant short channel effects (SCE) in short gate length devices, g_d increases.

According to Moore's prophecy, the device density increases $\times 2$ every 18 months [4]. It is achieved by scaling down transistor width and length while maintaining gate-to-channel electrostatics by proportionally scaling down oxide and channel thicknesses [5]. As the separation between source and drain decreases, short channel effects such as roll-off of V_T, increase in all SS_{min}, I_{OFF}, and g_d. Researchers have started to look for alternate non-traditional device geometries such as GAA and Tri-Gate FETs. Typical gate

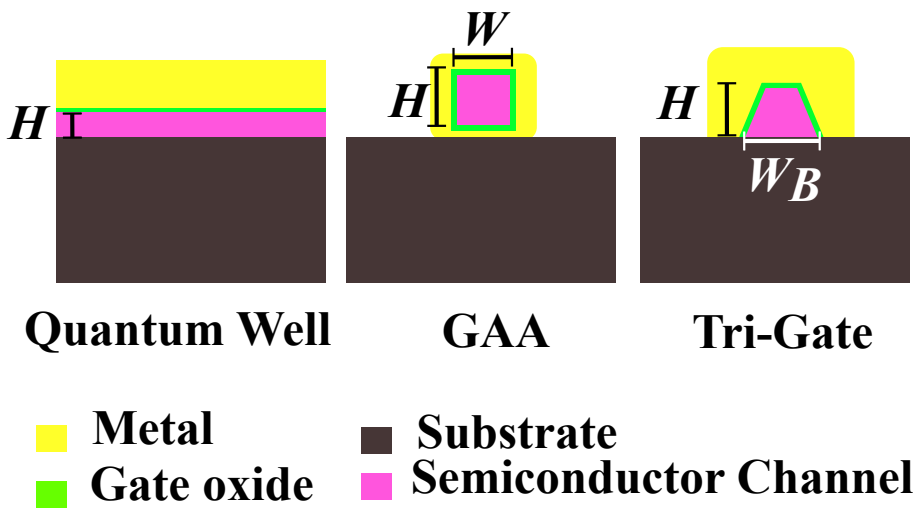

Figure 3.2: Gate metal-oxide-semiconductor channel geometry of Quantum well, GAA, and Tri-Gate FETs.

metal-oxide-channel structures of the quantum well, GAA, and Tri-Gate FET are shown in Figure 3.2.

3.2 MOS ELECTROSTATICS

In a MOSFET, the drain and source are separated by an energy barrier, with the top of the potential barrier being (E_0) [54], which ideally is controlled solely by V_{GS}. However, in real MOSFET at smaller gate lengths, drain voltage influences E_0 and thus drain to source current. Hence, 2D electrostatics are considered for current calculations to account for short-channel effects. The potential channel profile and 2D circuit level representation of the gate network are shown in Figure 3.3. In ballistic transistors, the electrons with positive k states at (E_0) move towards the drain without any scattering, and the carrier density is set by the difference between source fermi level (E_{FS}) and E_0. The electrons with negative k states move towards the source, and their density is regulated by the drain fermi energy level (E_{FD}). The E_0 is controlled by external biases and the presence of mobile carriers in the channel due to external bias. The potential in the channel due to external biases is given by,

$$U_{ext} = -q\left(\alpha_G(V_G - V_T) + \alpha_D V_D + \alpha_S V_S\right) \tag{3.1}$$

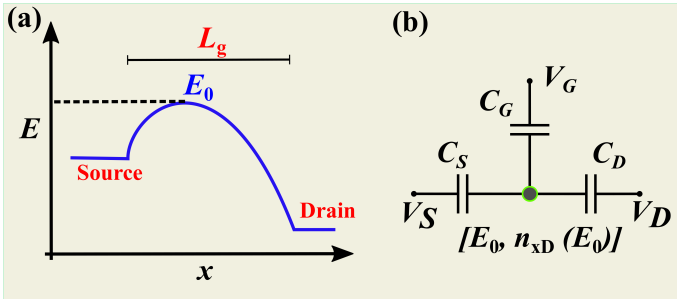

Figure 3.3: (a) Potential profile from source to drain reservoirs, (b) 2D capacitive network of the channel. The top of the barrier potential is set by source, gate, and drain capacitances through external biases and carrier concentration at the top of the barrier.

Where, $\alpha_G = \frac{C_G}{C_\Sigma}$, $\alpha_D = \frac{C_D}{C_\Sigma}$ and $\alpha_S = \frac{C_S}{C_\Sigma}$ are the gate, drain and source voltages influence on the E_0. The sum of all these coupling parameters is unity. In a MOSFET without short channel effects, $\alpha_G = 1$, and α_D and α_S are zero. The capacitance, $C_\Sigma = C_G + C_S + C_D$, and in general, C_S and C_D are negligible in the case of long channel devices. Finally, C_Σ accounts for the voltage drop across the gate oxide and band bending in the channel, and its expression is

$$C_\Sigma = \frac{C'_{ox}}{\alpha_G} \tag{3.2}$$

Where, C'_{ox} is the equivalent of series gate oxide capacitance (C_{ox}) and charge centroid capacitance (C_c). In detail, more about the various gate capacitances are discussed later in this section. The additional potential due to the presence of mobile charge carriers in the channel is,

$$U_{int} = q^2 \frac{n_0}{C_\Sigma} \tag{3.3}$$

Where n_0 is the total carrier density at E_0, which is the sum of carrier density with $+k$ momentum moving from source to drain (n_0^+) and carrier density with $-k$ momentum flowing towards the source from the drain (n_0^-), and i.e., $n_0 = n_0^+ + n_0^-$. The definition of carrier densities is as follows,

$$n_0^+ = \sum_{E_{mn/n}} \int \frac{1}{2} D_{xD}(E - E_0 - E_{mn/n}) f(E, E_{FS}) dE \tag{3.4}$$

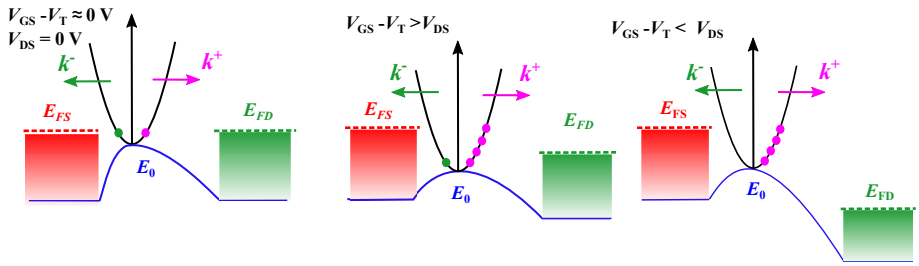

Figure 3.4: Channel potential profile from source to drain, positive and negative momentum carrier densities are highlighted, which are set by E_{FS} and E_{FD}. The channel potential profile from left to right are ballistic FETs operating in the cut-off, triode, and saturation region.

$$n_0^- = \sum_{E_{mn/n}} \int \frac{1}{2} D_{xD}(E - E_0 - E_{mn/n}) f(E, E_{FD}) dE \tag{3.5}$$

D_{xD} is the density of states of an electronic system with x dimension. For quantum well or planar MOSFETs, $x = 2$; for nanowire MOSFETs, x is 1. Factor 2 in the denominator accounts for half the k states contributing to net current. $f(E, E_F)$ is the equilibrium Fermi-Dirac statistics [55, 56], which explains the probability of occupancy of the states around fermi-level E_F at a given temperature.

Fermi-Dirac distribution is given by,

$$f(E, E_F) = \frac{1}{e^{\left(\frac{E - E_F}{k_B T}\right)} + 1} \tag{3.6}$$

The nonparabolic density of states of the 2D system is

$$D_{2D}(E) = \sum_{E_n} \frac{m_n^*}{\pi \hbar^2} (1 + 2\alpha_n (E - E_0 - E_n)) U(E - E_n) \tag{3.7}$$

Where E_n, m_n^*, and α_n are the nth subband energy magnitude given with respect to the bottom of the conduction band energy, its effective mass and nonparabolicity factor. $U(E)$ is the unit-step function. 1D nonparabolic density of states are

$$D_{1D}(E) = \sum_{E_{mn}} \frac{\sqrt{2m_{mn}^*}(1 + 2\alpha_{mn}(E - E_{mn}))}{\pi \hbar \sqrt{(E - E_{mn})(1 + \alpha_{mn}(E - E_{mn}))}} U(E - E_{mn}) \tag{3.8}$$

Where E_{mn}, m_{mn}^* and α_{mn} are the subband energy level given with respect to the conduction band energy of a 1D nanowire, its effective mass and nonparabolicity factor. The m and n are the subband indexes.

The complete expression of the top of the barrier potential E_0 is

$$E_0 = -q(\alpha_G(V_G - V_T) + \alpha_D V_D + \alpha_S V_S) + q^2 \frac{n_0}{C_\Sigma} \tag{3.9}$$

As n_0 and E_0 depend on each other, the equation (3.9) needs to be solved iteratively for E_0. The channel potential profile of ballistic FETs operating in the cut-off, triode, and saturation region are shown in Figure 3.4(a-c). The $+k$ and $-k$ states occupied at the top of the barrier are highlighted. In real MOSFETs, it is impossible to have zero scattering, and because of this, a few percentages of injected carriers from the source are backscattered into the source, and thus, the net current decreases. This scattering model assumes that the scattering is near the source region, and T is the transmission probability of the electron going from source to drain or drain to source without scattering. The carrier density at the top of the barrier with positive momentum is (n^+), and negative momentum (n^-) are given by,

$$n^+ = n_0^+ \tag{3.10}$$

n^+ is the same as the ballistic carrier density at the top of the barrier, assuming a quasi-equilibrium forward junction on the source side. The n^- consists of carriers backscattered into the source with reflection coefficient $(1 - T)$ and carriers scattered from the drain to the source with transmission probability, T.

$$n^- = (1 - T)n_0^+ + Tn_0^- \tag{3.11}$$

The E_0 in the quasi-ballistic devices becomes,

$$E_0 = -q(\alpha_G(V_G - V_T) + \alpha_D V_D + \alpha_S V_S) + q^2 \frac{(2 - T)n_0^+ + Tn_0^-}{C_\Sigma} \tag{3.12}$$

If T = 1 is substituted in the above equation (3.12), it results in E_0 of the ballistic device as given in equation (3.9).

NATURAL LENGTH (λ_N)

The minimum separation between the source and drain, when the V_{DS} starts influencing the channel potential, is called natural length (λ_n) [57]. Its magnitude is estimated by solving the 2D Poisson's equation, balancing the vertical and horizontal electric fields; hence, it is gate geometry dependent. It turns out that, for a decent transistor action without any severe SCE, $L_g > 5\lambda_n$ condition should be satisfied [58]. Natural lengths are derived for planar MOSFET with single and double gates, rectangular GAA FET, and Tri-Gate

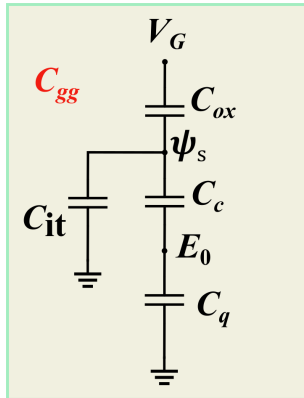

Figure 3.5: Gate capacitance network.

FET in [57,59]. The smaller the natural length for a given gate geometry, the smaller the L_g that can be realized without severe SCE for a given t_{ox} and channel thickness. GAA FET enables the ultimate gate length scaling limit followed by Tri-gate and planar FETs.

3.2.1 GATE CAPACITANCE

The gate capacitance network of a MOSFET is illustrated in Figure 3.5. The total gate capacitance is combination of C_{ox}, C_c and quantum capacitance (C_q) and an interface trap capacitance (C_{it}).

Gate oxide capacitance is dependent on gate geometry and gate oxide thickness (t_{ox}) and its dielectric constant, ε_{ox}. Higher ε_{ox} oxides such as HfO_2 and ZrO_2 allow larger C_{ox} even at larger oxide thicknesses. As the device dimensions decrease, gate oxide thickness is scaled down to maintain gate electrostatics. The expression for C_{ox} of quantum well, tri-gate, and GAA FETs are as follows: For quantum wells, it is a simple parallel plate capacitor, and it is

$$C_{ox} = \frac{\varepsilon_{ox}\varepsilon_0}{t_{ox}} \tag{3.13}$$

Where ε_0 is free space permittivity, for tri-gate FETs, i.e., heterostructure-based trapezoidal shape nanowires, the Poisson equation is solved numerically by considering Dirichlet and Neumann boundary conditions for various dimensions and the fit to the calculated capacitance is

$$C_{ox} = \frac{\varepsilon_{ox}\varepsilon_0(W_B + 2H)}{t_{ox}} + 1.35\varepsilon_{ox}\varepsilon_0 \tag{3.14}$$

Similar numerical simulaitons are perfomed for rectangular gate-all-around gate geomerty to get estimate of C_{ox} and it is also given in [60]:

$$C_{ox} = \frac{2\varepsilon_{ox}\varepsilon_0(W+H)}{t_{ox}} + 2.23\varepsilon_{ox}\varepsilon_0 \tag{3.15}$$

Next, C_c results from the presence of mobile charge carriers in the channel, which leads to an upward shift in the conduction band. The approximate shift in subband energy is calculated using first-order perturbation theory, valid in low carrier density limits and for larger subband separation. Higher-order perturbations should be accounted for at larger widths as the subband separation becomes small. Here, the Poisson equation is solved numerically for potential distribution from which the energy shift ΔE_1 is estimated using the 5-point finite element method. For quantum wells,

$$C_c = \frac{\varepsilon_s\varepsilon_0}{0.39H} \tag{3.16}$$

For tri-gate trapezoidal-shaped nanowires,

$$C_c = \frac{1.65\varepsilon_s\varepsilon_0(W_B+2H)}{W_B H} \tag{3.17}$$

For rectangular gate-all-around nanowires,

$$C_c = \frac{4.65\varepsilon_s\varepsilon_0(W^2+H^2)}{WH} \tag{3.18}$$

ε_s is the relative permittivity of the semiconductor channel. The quantum capacitance or semiconductor capacitance measures the availability of density of states in the semiconductor [61]. In metal, as the density of states is abundant, the ability to store the charge is high, and quantum capacitance is large. Whereas, in low-dimension systems such as 2D or 1D electronic system, the limited DOS lead to smaller quantum capacitance. The general definition of C_q is

$$C_q = -q^2\frac{\partial n_0}{\partial E_0} \tag{3.19}$$

The behavior of C_q replicates the DOS of the electronic system and can be used to extract subband separation. The total gate capacitance is the series equivalence of C_{ox}, C_c and C_q [62, 63]. When $C_q \gg C_{ox}$, transistor operates in MOS limit, whereas $C_q \ll C_{ox}$, the transistor operates in quantum capacitance limit (QCL) [64]. The transistor's performance would be improved by scaling gate oxide thickness in the MOS limit, whereas in the QCL limit, it would not.

The non-ideal interface between channel and gate oxides leads to interface traps, which get charged and discharged depending on the surface potential (Ψ_s). This charging and discharging of the traps add additional interface trap capacitance, $C_{it} = q^2 D_{it}$. Interface trap density, D_{it}, is set by the fermi-level and trap levels inside the band gap or near the band edges (VB or CB). Charging of the interface traps reduces the channel's carrier density, and thus I_{DS} decreases, and SS_{min} increases. The modified top of the barrier potential, including interface traps and neglecting α_S, is

$$E_0 = -q(\alpha_G(V_G - V_T) + \alpha_D V_D) + q^2 \frac{(2-T)n_0^+ + Tn_0^- + n_{trap}(\Psi_s)}{C_\Sigma} \tag{3.20}$$

n_{trap} is the interface trap carrier density. The approximate total gate capacitance is numerically calculated using, $C_{gg} = q\frac{\partial n_0}{\partial V_{GS}}$ at a given V_{DS}. For accurate gate capacitance calculations, the carrier density variation in source and drain regions below the fermi energy level should be accounted for in modeling.

BAND TAILS

Imperfections within semiconductors, including charge defects, alloy composition variations, phonon interactions, and the presence of dopants, can give rise to band tails located at the edges of the conduction band. These band tails exhibit an exponential decay characterized by the Urbach parameter (E_0) within the semiconductor's energy bandgap, as established by Urbach's seminal work in 1953 [65]. In [66], the importance of band tails at cryogenic temperatures, setting the limit on SS_{min}, is discussed. An empirical extension of the density of states utilizing generalized Fermi-Dirac integrals [67] is used to model band tails. The empirical formula of the 2D DOS accounting for exponential band tails considering a single conduction band subband is given in equation (3.21). The magnitude of E_0 is around a few meV, which decides the subthreshold slope at cryogenic temperatures.

$$D_{2D}(E) = \frac{m_1^*}{\pi\hbar^2}\left(F_{-1}\left(\frac{E-E_1}{E_0}\right) + 2\alpha_1 E_0 F_0\left(\frac{E-E_1}{E_0}\right)\right) \tag{3.21}$$

F_{-1} and F_0 are Fermi-Dirac integrals with order -1 and 0. The expression for the non-ideal SS_{min}, including interface traps and band tails, is

$$SS_{min} = ln(10)\frac{E_0 + \frac{k_B T}{q}}{E_0}\left(1 + \frac{C_q + C_{it}}{C_{ox}}\right) \tag{3.22}$$

3.3 MOS TRANSPORT

The total 1D drift-diffusion current in nondegenrate limit [68,69] in semiconductor is

$$J = qD_n \frac{\partial n_{xD}}{\partial x} + qn_{xD}\mu_n E \tag{3.23}$$

The first term accounts for diffusion transport due to spatial gradient in carrier concentration. The second term represents drift current due to applied electric field E. D_n is the diffusion constant and μ_n is the field effective mobility. In long-channel FETs, voltage-current characteristics are given by gradual channel approximation (GCA), assuming that the potential varies slowly from drain to source. The drain to source current expression in linear and saturation regions are

$$I_{DS,lin} = q\mu_n C_{ox} \frac{W_g}{L_g} \left((V_{GS} - V_T)V_{DS} - \frac{V_{DS}^2}{2} \right) \tag{3.24}$$

In saturation, it is

$$I_{DS,sat} = q\mu_n C_{ox} \frac{W_g}{L_g} (V_{GS} - V_T)^2 \tag{3.25}$$

Due to pinch-off, the above equations fail to explain the I-V characteristics in shorter gate-length devices. As the gate length becomes small, FETs operate in the ballistic limit and current is limited by velocity saturation [70] and the $I_{DS} \propto (V_{GS} - V_T)$.

Landauer approach [71,72] describes the MOSFET I-V characteristics from drift-diffusion limit to quasi ballistic to ballistic limit irrespective of dimensionality of the electronic system. According to Landauer's theory, the current between the two contact leads at a low applied field limit is:

$$I = \frac{2q}{h} \int T(E)M(E)(f_s - f_d)dE \tag{3.26}$$

The total current depends on the energy-dependent transmission probability, $T(E)$, and the number of conducting modes at a given energy, $M(E)$, and the difference between source and drain fermi levels, $(f_s - f_d)$. $T(E)$ depends on carrier mean-free path and L_g as given below.

$$T(E) = \frac{\lambda_{MF}(E)}{\lambda_{MF}(E) + L_g} \tag{3.27}$$

It is impossible to have zero scattering in the channel, and hence, the quasi-ballistic device model [73] is discussed here. The energy dependency of $T(E)$ is generally neglected and is considered constant. In ballistic device, when

$L_g \ll \lambda_{MF}, T \approx 1$, in quasi-ballistic device, $L_g \approx \lambda_{MF}, T < 1$ and in diffusive device, $L_g \gg \lambda_{MF}, T \ll 1$. The number of modes is proportional to the product of average velocity, $< v^+(E) >$ and D_{xD}. In a 2D quantum well, the number of modes per unit width is

$$M_{2D}(E) = g_v \frac{\sqrt{2m^*(E - E_c)}}{\pi\hbar} U(E - E_c) \tag{3.28}$$

Where g_v is the valley degeneracy factor. In the 1D case,

$$M_{1D}(E) = U(E - E_c) \tag{3.29}$$

In the linear region for small V_{DS}, $(f_s - f_d) \approx -\frac{\partial f_0}{\partial E}$. The conductance, G, is calculated from equation (3.26),

$$G = \frac{2q^2}{h} \int T(E) M_{xD} \left(-\frac{\partial f_0}{\partial E} \right) dE \tag{3.30}$$

G is the product of quantum conductance, $G_0 = \frac{2q^2}{h}$, transmission probability, and the number of conducting modes.

Quasi-ballistic current is calculated by combining equations (3.20) and (3.26). In Figure 3.6, InAs/InP quantum well, GAA, and tri-gate FET transfer characteristics are plotted, and interface trap carrier density is neglected. The channel dimensions of the FETs are as follows: the quantum well thickness is 10 nm, the dimension of the free-standing rectangular GAA nanowire is 10 nm $\times$ 10 nm, and the InAs/InP tri-gate nanowire with a height of 10 nm and top and bottom width of 5 and 25 nm, respectively. All the conducting subbands' electronic properties are derived from the 8-band k·p theory at room temperature. In Figure 3.6(a), *ballistic* transfer characteristics assuming ideal 1D electrostatics are plotted. QWFETs outperform the remaining transistor technologies if SCE is ignored. Tri-gate FETs drive current and transconductance are slightly larger than GAA FET due to a larger cross-section and more conducting subbands. In Figure 3.6(b), intrinsic transconductance, g_{mi} and SS_{min} vs. gate length with SCE is plotted. QWFET performance degrades for $L_g < 100$ nm and is worse below 40 nm. GAA FETs are more resistant to SCE due to excellent gate control. A similar behavior is observed in SS_{min} as a function of L_g. QWFETs have SS_{min}, larger than 100 mV/dec for $L_g < 40$ nm, whereas GAA FETs have 90 mV/dec at L_g = 15 nm. In Figure 3.6(c), g_{mi} vs. L_g are plotted for quasi-ballistic devices with λ_{MF} = 100 nm. At larger gate lengths, $g_m \propto \frac{\lambda_{MF}}{\lambda_{MF}+L_g}$ and at very short gate lengths, strong SCE appears and g_m decreases. 3D vertically stacked horizontal InAs/InP nanosheet/wire and quantum well's electrical performances are calculated using the above-discussed Landauer model, and the results are presented in Paper V.

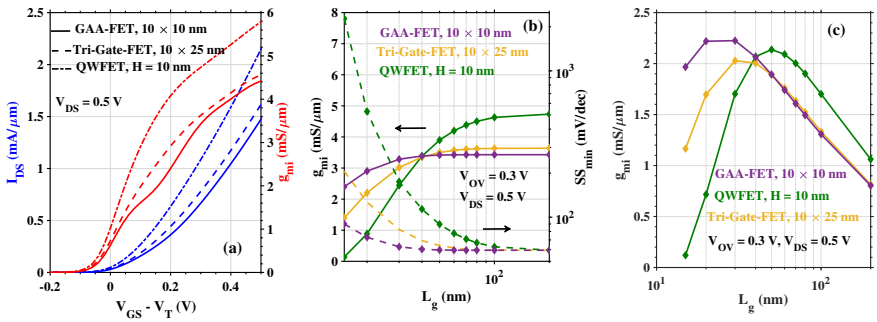

Figure 3.6: (a) Ballistic I_{DS} and g_{mi} vs. V_{GS} assuming ideal electrostatics, (b) ballistic g_{mi} and SS_{min} vs. L_g with SCE and (c) quasi-ballistic g_{mi} vs. L_g considering λ_{MF} = 100 nm and SCE. g_{mi} is taken at an overdrive voltage ($V_{OV} = V_{GS} - V_T$) of 0.3 V and V_{DS} = 0.5 V.

ON RESISTANCE (R_{ON})

In the triode region, the transistor acts as a resistor, and it is the sum of channel resistance and the total contact resistance.

$$R_{ON} = R_{ch} + R_s + R_d \tag{3.31}$$

R_s and R_d are the source and drain contact resistance, respectively. In a quasi-ballistic transport regime, the channel resistance becomes

$$R_{ON} = \frac{h}{2q^2M}\left(1 + \frac{L_g}{\lambda_{MF}}\right) + R_s + R_d \tag{3.32}$$

M is the number of conducting modes. The number of modes is estimated by subtracting the access resistance ($R_s + R_d$) from the y-intercept. The contact resistances are estimated from the contact layer's transfer length methods (TLMs). From the slope of R_{ON} vs. L_g and with known M, mean free path, λ_{MF} is extracted.

Transistor access regions and highly doped contact/metal junctions add additional resistances and are significant at shorter gate lengths. The extrinsic transconductance (g_{me}) and output conductance (g_{de}) are reduced compared to intrinsic values and are given in equations (3.33) and (3.34), respectively.

$$g_{me} = \frac{g_{mi}}{1 + R_s g_{mi} + (R_s + R_d) g_{di}} \tag{3.33}$$

$$g_{de} = \frac{g_{di}}{1 + R_s g_{mi} + (R_s + R_d) g_{di}} \tag{3.34}$$

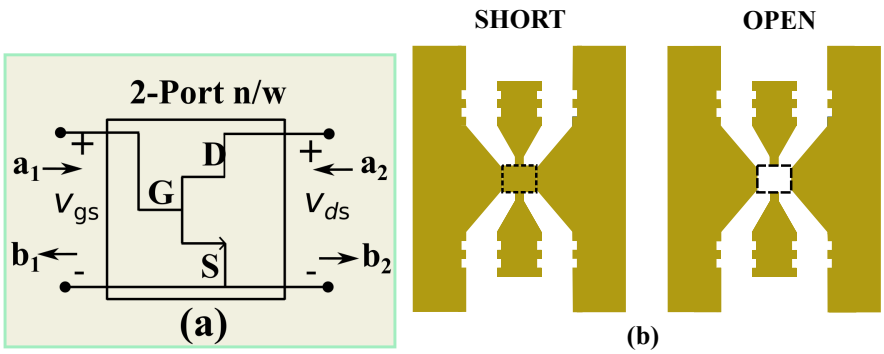

Figure 3.7: (a) MOSFET as a 2-port network in common source configuration, (b) short and open de-embedding structures with the highlighted DUT's active device area.

3.4 HIGH-FREQUENCY SMALL-SIGNAL CHARACTERIZATION

Two-port parameters such as impedance (Z), admittance (Y), and scattering (S)-parameters are used to describe the high-frequency characteristics of MOSFET in its common source configuration. S-parameters are easy to measure at high frequencies, requiring only 50 Ω terminations. Figure 3.7(a) shows the common source configuration of MOSFET. The scattering matrix elements S_{ij} [74] in terms of incident and reflected electromagnetic waves, a_i and b_i, respectively are

$$\begin{bmatrix} b_1 \\ b_2 \end{bmatrix} = \begin{bmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \end{bmatrix} \tag{3.35}$$

Where, S_{11} is the reflection coefficient at the port-1 when port-2 is terminated with 50 Ω, and S_{22} is the reflection coefficient at the port-2 when port-1 is matched with 50 Ω. S_{21} and S_{12} are the forward and reverse transmissions when the opposite port is terminated with 50 Ω. The S-parameters can be easily transformed into Y and Z, from which the device's intrinsic parameters, such as capacitance, resistances, and conductance, are extracted. The S-parameter measurements are linear because we characterize the frequency behavior of FET near the bias voltage.

A vector network analyzer (VNA) measures the magnitude and phase of the transmitted and reflected waves in the specified frequency range. An off-chip through-match-reflect-reflect (TMRR) [75] calibration is performed on impedance standard substrate (ISS) to move the reference plane from the VNA

ports to the probe tips or, in other words, the effect of cable loss is estimated, which would be removed from the measured S-parameters.

The measurement pads introduce significant capacitance and inductances compared to the device under test, and these capacitances are removed through de-embedding. Figure 3.7(b) shows an on-chip open and short de-embedding structures used in this work. A standard 2-step de-embedding procedure [76] is used in this work, and the procedure is as follows:

$$[S]_{device,em} \rightarrow [Y]_{device} \tag{3.36}$$

$$[S]_o \rightarrow [Y]_o \tag{3.37}$$

$$[S]_s \rightarrow [Y]_s \tag{3.38}$$

$$[Y]_{s-o} = [Y]_s - [Y]_o \rightarrow [Z]_{s-o} \tag{3.39}$$

$$[Y]_{device,o} = [Y]_{device} - [Y]_o \tag{3.40}$$

$$[Y]_{device,o} \rightarrow [Z]_{device,o} \tag{3.41}$$

$$[Z]_{device,de-em} = [Z]_{device,o} - [Z]_{s-o} \tag{3.42}$$

$$[Z]_{device,de-em} \rightarrow [S]_{device,de-em} \tag{3.43}$$

$[S]_o$ and $[S]_s$ are the S-parameters of the open and short de-embedding structures, respectively. $[S]_{device,em}$ and $[S]_{device,de-em}$ are the device S-parameters before and after de-embedding procedure.

3.4.1 HIGH-FREQUENCY SMALL-SIGNAL MODEL

A simple hybrid-π model to pervasive circuit models are used to capture various physical phenomena that happen at low frequencies to very high frequencies [77–79]. The extended hybrid-π model of the MOSFET used in this work is shown in Figure 3.8. The analytical expressions for Y-parameters and extrinsic and intrinsic parameter extraction are detailed in this section.

MOSFET, in its OFF state, is called cold-FET. In this state, all intrinsic parameters are negligible except parasitic components [80]. The cold-FET circuit representation is shown in Figure 3.9. The external contact resistances and mutual parasitic capacitance are extracted in this state. Parasitic capacitances are divided into a constant extrinsic parasitic capacitance

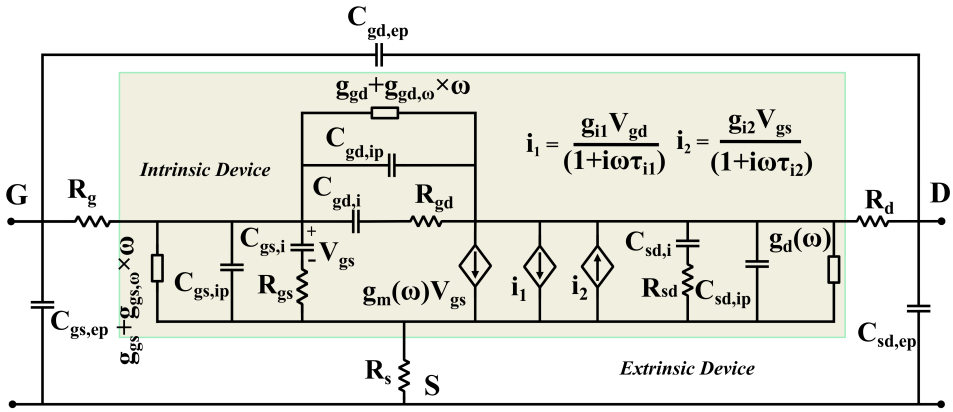

Figure 3.8: MOSFETs non-quasi-static small-signal model and intrinsic and extrinsic parts are highlighted.

$(C_{gs,ep}, C_{gd,ep}, C_{sd,ep})$, which is due to mutual coupling among metal contacts and are approximately constant for all device widths, and an intrinsic parasitic capacitance, $(C_{gs,ip}, C_{gd,ip}, C_{sd,ip})$, which is gate width dependent and is due to the overlap of gate metal contact on the access region or highly doped regrown source and drain contacts. Generally, off-state capacitances are plotted for various device widths. The zero device width gives the constant extrinsic parasitic capacitance, and the linear slope gives the width-dependent parasitic capacitance. The measured parasitic capacitances are validated by simulating the structures in the COMSOL Multiphysics FEM solver using distributed capacitance, which apprehends the metal-oxide-semiconductor stack. The equivalence of the quantum capacitance of the semiconductor is the input to the electrostatic model, where the effective relative permittivity of dielectric is calculated assuming a 1 nm thick dielectric. Contact resistances can also be extracted from the $real(Z)$ parameters at high frequencies. Typical, Z-parameters in the OFF-state are plotted in Figure 3.10. However, the cold-FET measurements do not result in reliable resistances; thus, TLMs are generally used for contact resistance estimations.

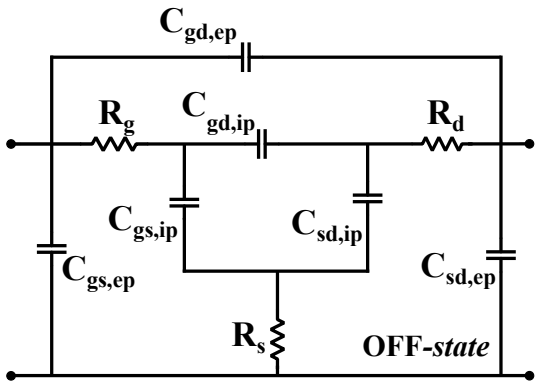

Figure 3.9: The cold-FET small-signal circuit representation.

The intrinsic Y-parameters, Y_{int} are obtained by subtracting the extrinsic resistances [81] and capacitances from the measured and de-embedded total Y_{de} as follows:

$$Y_{int} = \left[\left[Y_{de} - \begin{bmatrix} C_{gs,ep} + C_{gd,ep} & C_{gd,ep} \\ C_{gd,ep} & C_{sd,ep} + C_{gd,ep} \end{bmatrix} \right]^{-1} - \begin{bmatrix} R_s + R_g & R_s \\ R_s & R_s + R_d \end{bmatrix} \right]^{-1} \tag{3.44}$$

g_{gs} and g_{gd} account for conductance loss due to oxide traps at very low frequencies. $g_{gs,\omega}$ and $g_{gd,\omega}$ reflect the frequency-dependent conductance loss due to border traps in the gate oxide. In this non-quasi-static model, every intrinsic capacitance is replaced with a capacitance in series with a resistance. The carrier response delay to gate voltage is modeled by series $C_{gs,i}$ and R_{gs} branch. The limit on channel resistance is given in equation (3.45), where the lower and upper limits are set by diffusion and ballistic limited transport, respectively.

$$\frac{1}{5g_{mi}} \leq R_{gs} \leq \frac{1}{1.4g_{mi}} \tag{3.45}$$

Logarithmic frequency dependency of transconductance, $g_m(\omega)$ and output conductance $g_d(\omega)$ account for the border traps in the oxide [82], and the expressions are given in equation (3.46) and (3.47), respectively.

$$g_m(\omega) = g_{m,i}\left(1 + \alpha log\left(\frac{\omega}{\omega_0}\right)\right) \tag{3.46}$$

$$g_d(\omega) = g_{d,i}\left(1 + \beta log\left(\frac{\omega}{\omega_0}\right)\right) \tag{3.47}$$

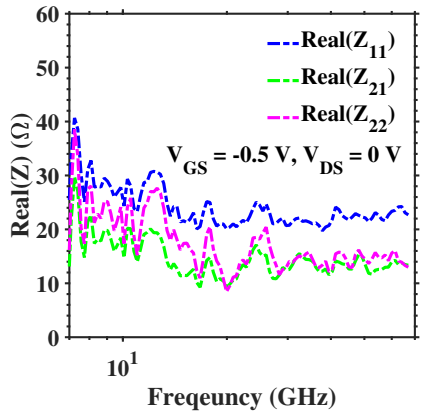

Figure 3.10: $Real(Z_{11})$, $Real(Z_{12})$ and $Real(Z_{22})$ vs. frequency in the device OFF-state. This data is of $In_{0.71}Ga_{0.29}As/InP$ QWFET fabricated using sacrificial a:Si spacer and device has gate width, W_g = 40 μm and L_g = 80 nm.

Where α and β account for the slope of the conductances and ω_0 is the dispersion angular cut-off frequency. In addition, i_1 and i_2 are two additional current sources accounting for impact ionization and band-band tunneling, which are prevalent in narrow bandgap semiconductors. g_{i1} and g_{i2} are the magnitude of the current sources and τ_{i1} and τ_{i2} are the time-constants [83]. The order of these time constants is in the range of 10's of nS; thus, these current sources' contribution is negligible at high frequencies. The Y_{int} parameters expressions of the complete small-signal model are given below:

$$\begin{aligned} Y_{11} =& g_{gs} + g_{gs,\omega}\omega + g_{gd} + g_{gd,\omega}\omega + \frac{i\omega C_{gs,i}}{1 + i\omega R_{gs}C_{gs,i}} + \frac{i\omega C_{gd,i}}{1 + i\omega R_{gd}C_{gd,i}} \\ &+ i\omega\left(C_{gs,ip} + C_{gd,ip}\right) \end{aligned} \tag{3.48}$$

$$Y_{12} = -g_{gd} - g_{gd,\omega}\omega - \frac{i\omega C_{gd,i}}{1 + i\omega R_{gd}C_{gd,i}} - i\omega C_{gd,ip} \tag{3.49}$$

$$\begin{aligned} Y_{21} =& \frac{g_{mi}\left(1 + \alpha log\left(\frac{\omega}{\omega_0}\right)\right)}{1 + i\omega R_{gs}C_{gs,i}} - g_{gd} - g_{gd,\omega}\omega - \frac{i\omega C_{gd,i}}{1 + i\omega R_{gd}C_{gd,i}} - \frac{g_{i1}}{1 + i\omega\tau_{i1}} \\ &+ \frac{g_{i2}}{1 + i\omega\tau_{i2}} - i\omega C_{gd,ip} \end{aligned} \tag{3.50}$$

$$Y_{22} = g_{di}\left(1+\beta log\left(\frac{\omega}{\omega_0}\right)\right) + g_{gd} + g_{gd,\omega}\omega + \frac{i\omega C_{gd,i}}{1+i\omega R_{gd}C_{gd,i}} + \frac{g_{i1}}{1+i\omega\tau_{i1}} + \frac{i\omega C_{sd,i}}{1+i\omega R_{sd}C_{sd,i}} + i\omega\left(C_{gd,ip}+C_{sd,ip}\right) \tag{3.51}$$

In a limit, $(\omega R_{ij}C_{ij})^2 \ll 1$, where $\{ij\} \in \{gs, gd, sd\}$, the above mentioned Y-parameters are simplified as follows.

$$Y_{11} = g_{gs} + g_{gs,\omega}\omega + g_{gd} + g_{gd,\omega}\omega + \omega^2\left(R_{gs}C_{gs,i}^2 + R_{gd}C_{gd,i}^2\right) + i\omega\left(C_{gs,i}+C_{gd,i}+C_{gs,ip}+C_{gd,ip}\right) \tag{3.52}$$

$$Y_{12} = -g_{gd} - g_{gd,\omega}\omega - \omega^2 R_{gd}C_{gd,i}^2 - i\omega(C_{gd,i}+C_{gd,ip}) \tag{3.53}$$

$$Y_{21} = g_{mi}\left(1+\alpha log\left(\frac{\omega}{\omega_0}\right)\right)(1-i\omega R_{gs}C_{gs,i}) - g_{gd} - g_{gd,\omega}\omega - \omega^2 R_{gd}C_{gd,i}^2 - \frac{g_{i1}}{1+i\omega\tau_{i1}} + \frac{g_{i2}}{1+i\omega\tau_{i2}} - i\omega(C_{gd,i}+C_{gd,ip}) \tag{3.54}$$

$$Y_{22} = g_{di}\left(1+\beta log\left(\frac{\omega}{\omega_0}\right)\right) + g_{gd} + g_{gd,\omega}\omega + \omega^2\left(R_{gd}C_{gd,i}^2 + R_{sd}C_{sd,i}^2\right) + \frac{g_{i1}}{1+i\omega\tau_{i1}} + i\omega(C_{gd,ip}+C_{gd,i}+C_{sd,i}+C_{sd,ip}) \tag{3.55}$$

3.4.2 INTRINSIC PARAMETER EXTRACTION

In this section, various intrinsic parameters extraction from the Y_{int} is discussed [79].

$$C_{gs,i} = \frac{Im(Y_{11}+Y_{12})}{\omega} - C_{gs,ip} \tag{3.56}$$

$$C_{gd,i} = -\frac{Im(Y_{12})}{\omega} - C_{gd,ip} \tag{3.57}$$

$$C_{sd,i} = \frac{Im(Y_{22}+Y_{12})}{\omega} - C_{sd,ip} \tag{3.58}$$

$$g_{mi} = Re(Y_{21} - Y_{12}) \tag{3.59}$$

$$g_{di} = Re(Y_{22} + Y_{12}) \tag{3.60}$$

g_{mi} and g_{di} are taken at low frequencies. In the frequency limit, $(\omega^2 R_{ij} C_{ij}^2) \ll 1$, where $\{ij\} \in \{gs, gd, sd\}$ or at high-frequency limit,

$$R_{gs} = \frac{Re(Y_{11} + Y_{12})}{\omega^2 C_{gs,i}^2} \tag{3.61}$$

$$R_{gd} = -\frac{Re(Y_{12})}{\omega^2 C_{gd,i}^2} \tag{3.62}$$

$$R_{sd} = \frac{Re(Y_{22} + Y_{12})}{\omega^2 C_{sd,i}^2} \tag{3.63}$$

3.4.3 HIGH-FREQUENCY GAINS

Any active two-port device has current gain and power gain, which can be defined in multiple ways. The short circuit current gain is defined using Y-parameters as follows:

$$|h_{21}|^2 = \left|\frac{Y_{21}}{Y_{11}}\right|^2 \tag{3.64}$$

Substituting the Y-parameters in the above equation results in $|h_{21}|^2 \propto \frac{1}{\omega^2}$ or -20 dB/dec nature of current gain with frequency. The transition frequency, f_T, is the frequency, where $|h_{21}|^2 = 0$ dB. The analytical expression of FET's f_T with nonzero R_s and R_d is

$$\frac{1}{2\pi f_T} = \frac{C_{gs} + C_{gd}}{g_{mi}} + \frac{C_{gs} + C_{gd}}{g_{mi}}(R_s + R_d)g_{di} + (R_s + R_d)C_{gd} \tag{3.65}$$

FETs with larger transconductance, smaller capacitances, smaller contact resistances, and negligible short channel effects would have larger f_T. At longer gate lengths, $f_T \propto \frac{1}{L_g^2}$ and at shorter gate lengths, g_{mi} saturates due to SCE and $f_T \propto \frac{1}{L_g}$ and later f_T decreases drastically with L_g due to severe SCE.

Actual power gain (G_A) is the ratio of power delivered to the load and the transistor. The power gain is infinite in an ideal MOSFET because of pure input reactive impedance. Maximum available gain (MAG) is defined

as the ratio of maximum available power at the load to the maximum power available at the source. This gain is valid only when the stable transistor or the stability factor (K) is larger than 1. Maximum stable gain (MSG) is used as a power gain metric when the transistor is unstable or $K < 1$. Oscillation frequency, f_{max} is where MAG is 1 or 0 dB. Unlike $|h_{21}|^2$, MAG behavior is complex at high frequencies because of its dependency on the stability factor. Mason's unilateral power gain, $(|U|)$, is defined for a unilateral 2-port network and is independent of K, and it becomes 0 dB at f_{max} of the transistor. Another advantage of using $|U|$ for device analysis is that it is independent of the transistor's configuration and can be extrapolated with -20 dB/dec to estimate f_{max}. However, in III-V-based transistors, the $|U|$ roll-off becomes larger than -20 dB/dec at high frequencies because of its non-quasi-static nature.

The definition of power gains in terms of Y-parameters are:

$$MSG = \frac{|Y_{21}|}{|Y_{12}|} \tag{3.66}$$

$$MAG = \frac{|Y_{21}|}{|Y_{12}|}\left(K - \sqrt{(K^2 - 1)}\right) \tag{3.67}$$

and

$$U = \frac{|Y_{21} - Y_{12}|^2}{4(Real(Y_{11})Real(Y_{22}) - Real(Y_{12})Real(Y_{21}))} \tag{3.68}$$

The stability factor,

$$K = \frac{2Real(Y_{11})Real(Y_{22}) - Real(Y_{12}Y_{21})}{|Y_{12}Y_{21}|} \tag{3.69}$$

The approximate expression for f_{max}

$$f_{max} = \sqrt{\frac{f_T}{8\pi R_g C_{gd}\left(1 + \frac{2\pi f_T}{C_{gd}}\psi\right)}} \tag{3.70}$$

Where,

$$\psi = \frac{g_{di}}{g_{mi}^2}\left[C_T^2 + \frac{R_d}{R_g}(C_{gd,i} + C_{gd,ip})^2 + \frac{R_s}{R_g}(C_{gs,i} + C_{gs,ip})^2 + \frac{R_i}{R_g}(C_{gs,i})^2\right] \tag{3.71}$$

Where, $C_T = C_{gs,i} + C_{gd,i} + C_{gs,ip} + C_{gd,ip}$. Transistors with larger f_T, smaller gate resistance, output conductance, and C_{gd} will result in larger f_{max}.

4

MOSFET Fabrication

This chapter details essential steps in fabricating RF MOSFETs designed and implemented in this thesis.

4.1 INGAAS QUANTUM WELL MOSFETS WITH NITRIDE SPACERS

InGaAs composite channel MOSFETs are fabricated in a top-down approach. 2-finger devices are, in general, designed and fabricated, which allows for G-S-G probe configuration for RF measurements. In this work, devices with each finger width varying from 5 to 20 μm and gate lengths from 40 nm to 140 nm are fabricated.

4.1.1 MBE EPITAXIAL STACK

The schematics of the self-aligned InGaAs MOSFET with nitride spacers and the heterostructure structure used for fabrication are shown in Figure 4.1. The epitaxial stack is designed to minimize contact resistance by tuning barrier/etch stop layer thickness, indium composition, and doping. Composite channel consisting of $In_{0.53}Ga_{0.47}As/In_{0.8}Ga_{0.2}As/In_{0.53}Ga_{0.47}As$ (3/5/4 nm) is used for carrier conduction medium, where the peak carrier concentration is confined in the high indium content epi layer and away from high-κ/semiconductor interface resulting in larger carrier mobility. This type of composite channel design has been used in high-performing HEMTs, MOS-HEMTs, and MOSFETs [33, 34, 84]. The n-InP etch stop layer and back side δ-doped InAlAs are expected to provide carriers to the access region, thereby minimizing the access resistance [85]. The n-InP top barrier is anticipated to yield a lower contact resistance because of the smaller conduction band offset

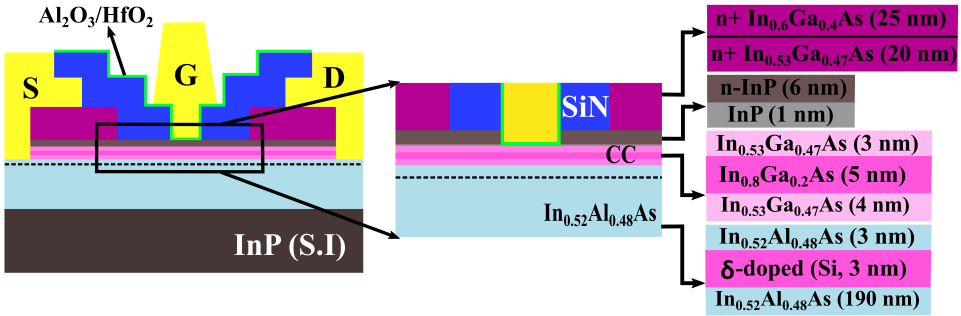

Figure 4.1: From left to right, schematics of the self-aligned InGaAs composite channel MOSFET with nitride spacers and various epitaxial layers in the heterostructure.

between InP and InGaAs materials compared to the conventional InAlAs top barrier in HEMTs. However, the InP/InGaAs interface suffers from a high resistive interface layer compared to the InAlAs/InGaAs interface, resulting in lower electron mobility.

4.1.2 N^+ ETCH OR L_{SD} DEFINITION

A highly doped higher indium content InGaAs cap layer is etched using Al_2O_3/SiN (approx. 5/40 nm) as a mask. To minimize the horizontal etching of InGaAs, atomic layer deposition (ALD) deposited Al_2O_3 (50 cycles at 100°C) is used as it has better adhesion to InGaAs. SiN is deposited using plasma enhanced chemical vapor deposition (PECVD) at 250°C, and $SiH_4/Ar/N_2$ are the precursor gases. A 200 nm thick ARP 6200.09 e-beam resist is patterned and used as a mask to define the L_{sd} in the hard mask. The SiN is etched for 80 s using reactive ion etching (RIE), CHF_3/O_2 (50/5 sccm, 75 W) gases at 20 mbar pressure. The etch rate is calibrated on the planar sample, approximately 32 nm/min. A small percentage of oxygen is added to the plasma to eliminate the formation of fluorocarbons. Later, Alumina is etched in diluted HF(1:400) for 15 s, and the expected etch rate is close to 0.3 nm/s. The n^+ cap layer is etched for 45 s in freshly prepared H_3PO_4:H_2O_2:H_2O (1:1:25). The isotropic nature of the wet chemical etching has resulted in larger openings in n^+ cap layer compared to the openings in etch mask. The schematic of this process step or gate recess step I is shown in Figure 4.2. This hard mask is removed before the next process step. The width of the openings in the n^+ cap layer ranges from 500 to 900 nm.

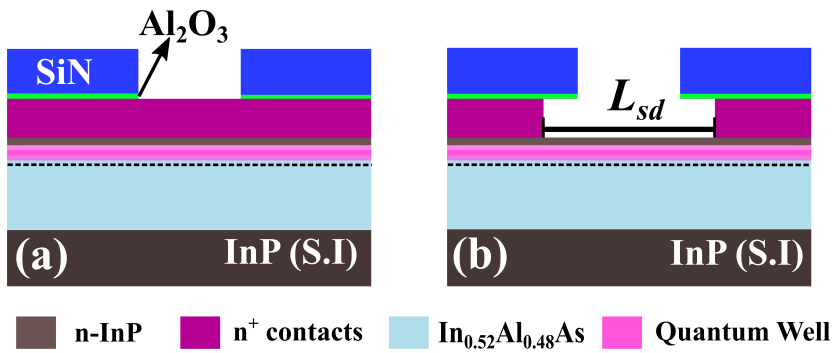

Figure 4.2: Gate recess etch step I using (a) (5/40 nm) Al_2O_3/SiN hard mask, (b) Post n^+ cap layer wet etching.

4.1.3 DEVICE ISOLATION

Device isolation eliminates short circuits between the same device's source and drain contacts and isolates all the devices on the chip. Photoresist S1813/05 or e-beam resist hydrogen silsesquioxane (HSQ) is used as an etch mask, and the mask width defines the finger width. Laser or electron-beam lithography (EBL) is used to define the mask width, and the alignment accuracy of EBL is higher than laser writer (offset of around 500 nm), however, at the expense of cost and writing time. Post development, the S1813/05 etch mask is baked at 120°C for 10 min to improve adhesion. MESA etch starts from the top n^+ cap etching in H_3PO_4:H_2O_2:H_2O (1:1:25) for 45 s, InP etch stop layer etching in HCl:H_2O (1:1) for 5 to 7 s, and InGaAs CC in citric acid:H_2O_2 (1:1) for 14 s. All the acid-based etchants are freshly prepared to get reproducible etch rates. The citric acid:H_2O_2 (1:1) has etch selectivity between InGaAs and $In_{0.52}Al_{0.48}As$ and is used to have more control on MESA height while etching the composite channel.

4.1.4 CONTACTS METALLIZATION

Ti, Ni, and Mo are a few of many metals that displayed smaller contact resistivity to highly doped n-InGaAs. The choice of metal contact is limited by the thermal budget rather than the work function. This work uses non-alloy Ti/Pd/Au metal for contacts. A thin Titanium metal improves adhesion to the semiconductor, and a thin palladium metal acts as a diffusion barrier to gold during annealing. Photoresist ma-N 1420 or e-beam PMMA A6 resist is used for the lift-off process for the Ti/Pd/Au (10/10/100 nm) source and drain contact, which are deposited through e-beam evaporation. Figure 4.3

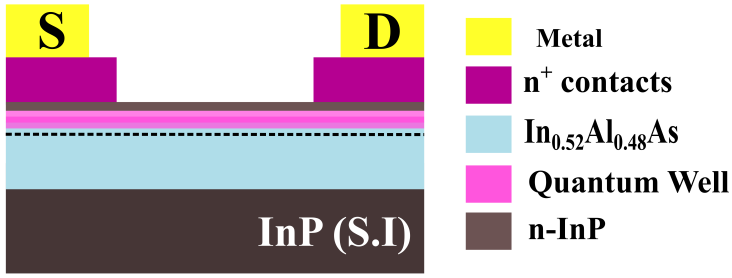

Figure 4.3: Schematic of the MOSFET cross-section post source and drain contact metallization.

shows the schematic view of the device's cross-section post source and drain contact metallization.

4.1.5 GATE RECESS ETCH ETEP-II

Gate length is defined in this step using nitride as a self-aligned spacer. A 40 nm thick SiN is deposited using PECVD, and an e-beam resist ARP 6200.09 is used as a mask for openings in the nitride spacer. The nitride layer is etched using the same RIE etch recipe used in mask preparation for n^+ cap layer etching. The openings in the SiN are defined as the gate length of the MOSFET. The n-InP etch stop layer is removed in the gated region for effective gate modulation of the channel. This epitaxial layer can be removed by anisotropic dry etching or digital etches while minimizing the lateral etch. The latter process is implemented in this work as this process is expected to create less damage to the channel surface. Gate lengths ranging from 40 to 140 nm are realized. Scanning electron microscope (SEM) top view of the device, post gate length definition in SiN spacer is shown in Figure 4.4.

DIGITAL ETCHING (DE)

Digital etching is an etching technique used in general to thin down a semiconductor layer by a couple of nanometers, and this is widely used in either thinning down nanowires or selectively removing a thin semiconductor layer. One digital etch cycle consists of an oxidation step and selective wet etching of the thin oxidized layer. The type of oxidation process decides the oxidized layer thickness and the quality of the etch process. H_2O_2 was used as an oxidizer to create an oxide layer, which will be removed using diluted acid [86]. However, chemical oxidation has cross-contamination issues; hence,

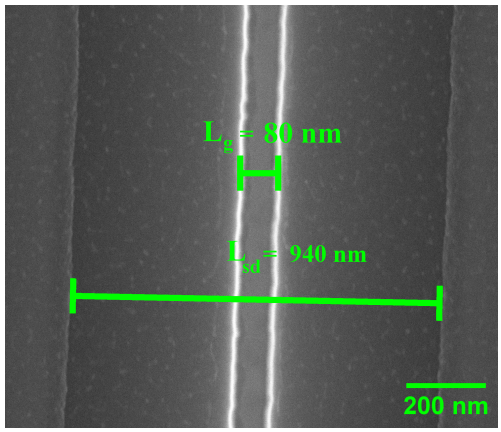

Figure 4.4: Post gate recess step II and resist removal. The SEM top view of the device with L_g = 80 nm and L_{sd} = 940 nm.

it is replaced with UV ozone cleaning or O_2 plasma. Various digital etching recipes for III-V semiconductors were introduced using HCl: IPA, H_2SO_4 and Citric acid (1:1) in [87–89] in combination with UV ozone cleaning. The oxide layer thickness might vary depending on the semiconductor material and oxidation process. In this device fabrication, one digital etch cycle involves 8 min ozone cleaning and 15 s oxide etch in HCl (1:10). Etch tests are performed on the test structures with SiN mask, and after each cycle, InGaAs etch is performed. It is noted that the InGaAs channel is removed after three DE cycles, which means InP is removed entirely after 3 DE.

4.1.6 GATE CONTACT FORMATION

PASSIVATION AND HIGH-κ OXIDE DEPOSITION

Sulfur passivation of InGaAs surface in ammonium sulfide (NH_4S) (1:1) before gate oxide deposition has shown to produce reasonable SS_{min} in InGaAs MOSFETs [90–92]. In-situ cleaning of the InGaAs surface without degrading the surface quality is also promising [93]. ALD deposited high-κ dielectrics such as Al_2O_3, HfO_2 and bilayer of these two oxides [94–96] and ZrO_2 [97,98] and La_2O_3 [99] are promising choices for the gate oxide. Bilayer Al_2O_3/HfO_2 (7/37 cyc) is deposited using ALD, where Al_2O_3 is deposited at high-temperature 250°C to minimize interface defects and high permittivity HfO_2 is deposited to increase the effective dielectric constant of the gate oxide. This bilayer has an effective oxide thickness $\left(EOT = t_{ox}\frac{\varepsilon_{SiO_2}}{\varepsilon_{ox}}\right)$ close to 1 nm [100].

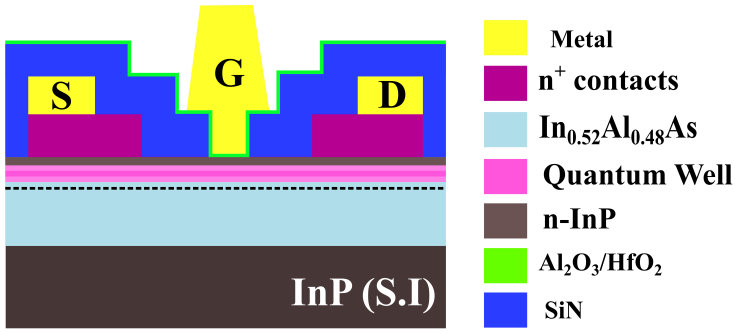

Figure 4.5: Schematic of the MOSFET cross-section after gate contact formation.

GATE CONTACT METALLIZATION

W [101], Mo [102], TiN/W [103] and metal stacks like Ti/Pd/Au [104], Ni/Pd/Au [105] have been used as gate contacts. These can be deposited through different deposition techniques like evaporation, sputtering, and ALD, where the latter two give superior sidewall coverage. As the FET is fabricated using a self-aligned process, gate contact placement seems quite direct; however, it is to be noted that large gate contact overlap with highly doped contacts or spacer would result in large parasitic capacitance. Bilayer and tri-layer resist stacks are used to implement T-gate, particularly in HEMTs [106, 107]. A thick and sacrificial inorganic spacer is also used to implement the T-gate in MOSFETs. We used a:Si and SiN for the T-gate definition in two different MOSFET fabrications. MOSFETs with InP $\wedge$-ridge spacer, bilayer e-beam resist stack PMMA A2/MMA EL6 is used to define T-gate. Ti/Pd/Au (10/10/45 nm) non-alloy metal is evaporated and developed using a lift-off process. InGaAs MOSFET with SiN spacer, post gate contact formation is shown in Figure 4.5 and the SEM top view of the device post-completion is shown in Figure 4.6.

4.2 INP $\wedge$-RIDGE SPACER MOSFET PROCESS

A novel modulation doped InP $\wedge$-ridge spacer process was introduced, with the flexibility of tuning the InP spacer width, and the fabrication is briefly explained in paper II. This FET design involves three epitaxial layer growths and lithography-sensitive steps, which has reduced the device yield. Semi-insulating InP (100) substrate is used to fabricate the devices. 2% HSQ

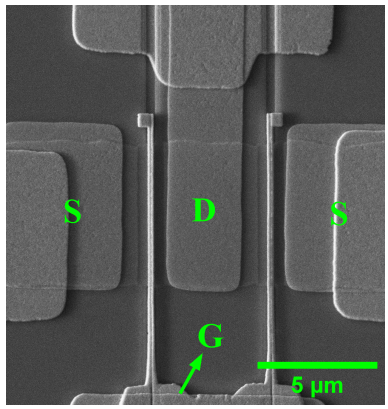

Figure 4.6: SEM top view of the 2-finger device with total gate width, W_g = 2 $\times$ 5 μm.

single pixel lines (SPLs) are defined using EBL, whose width and separation decide the nanowire separation and width, respectively. The precursor gases used for the InGaAs epitaxial layer are trimethyl indium (TMIn), trimethyl gallium (TMGa), and arsine (AsH_3). The nanowires are oriented in the [100] direction, and side wall facets are set by {110}. A 7 nm thick and 30 nm wide $In_{0.85}Ga_{0.15}As$ nanowires are grown at 600° using metal-organic chemical vapor deposition technique (MOCVD). Again, 2% HSQ is used to create a mask aligned in the [011] direction for the InP $\wedge$-ridge spacers growth. The side walls of the InP spacers {111}B have a negligible growth rate compared to the top facet (100), thus resulting in a $\wedge$-ridge-shaped spacer. The HSQ dummy gate (DG) is defined on top of the array of nanowires and in between the InP spacers and 20 nm thick Sn-doped $In_{0.63}Ga_{0.37}As$ ($N_d = 5\times 10^{19}/cm^3$) contacts are regrown. Smaller spacer widths below 40 nm are challenging to realize due to proximity effects during mask patterning. InGaAsP layer formation between InP and InGaAs nanowire layers has led to problems during MESA etching. The schematic of the device cross-section and SEM top view of the device post-InP spacer formation are shown in Figure 4.7(a) and (b), respectively.

4.3 SACRIFICIAL A:SI SPACER MOSFET PROCESS

A quantum well $In_{0.71}Ga_{0.29}As$/InP MOSFETs with a self-aligned sacrificial amorphous silicon (a:Si) spacer are fabricated. Its fabrication process is explained briefly in Paper I. An a:Si spacer reduces the gate contact overlap on highly doped contacts and reduces parasitic capacitances. The schematic

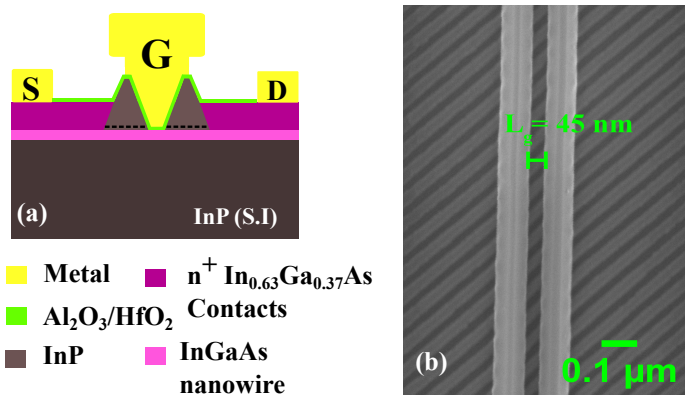

Figure 4.7: (a) Schematic of the cross-section of RF MOSFET with InP ∧-ridge spacer, (b) SEM top view of the device with gate length, L_g = 45 nm, post InP ∧-ridge spacers formation.

of the device cross-section is shown in Figure 4.8. $In_{0.71}Ga_{0.29}As$ /InP (12/2 nm) epitaxial layers are grown on InP (100) substrate using MOCVD. An HSQ dummy gate is defined using e-beam lithography, and the InP barrier in the contact region is removed by doing 3 digital etches. Next, 25 nm thick Sn-doped $In_{0.63}Ga_{0.37}As$ (N_d = 5× 10 ^{19}/cm^3) contacts are grown at 600 °C. Approximately 80 to 90 nm thick a:Si is deposited using PECVD to get a conformal spacer around DG covering n^+ cap layer. S1805:PGMA (1:1.5) planarization layer is used as a mask to remove a:Si on top of the DG. Next, HSQ DG is selectively removed, leaving a self-aligned a:Si spacer around the channel covering the face away {111}B facets of the highly doped contacts. After gate contact formation and gate oxide etch, a:Si is removed using dry isotropic etching, creating an air spacer and T-gate structure. The advantage of using a:Si spacer is to attain wet etch selectivity to HSQ DG and gate oxide. Figure 4.8 shows the schematic of the spacer formation, complete device, and SEM top view of the spacer formation around the channel.

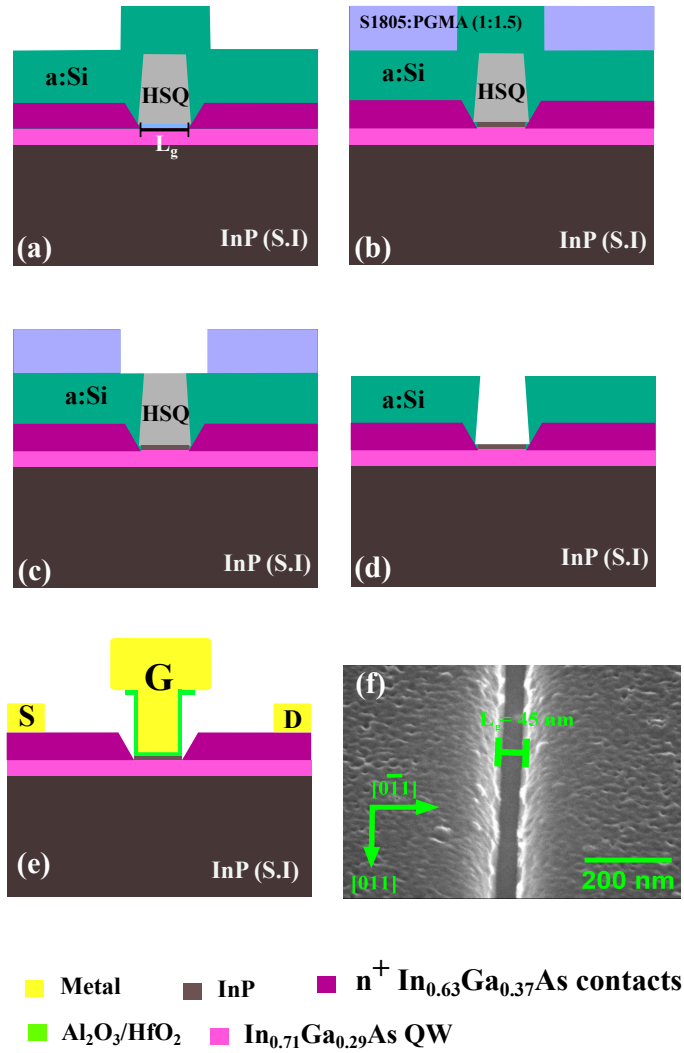

Figure 4.8: (a-d) Schematic of the cross-section of self-aligned a:Si spacer formation and (e) complete device after an air spacer formation and contacts metallization and (f) SEM top view of the device with L_g = 45 nm, post a:Si spacer formation around the channel.

5

Device Characterization

This chapter presents simulation and experimental results in detail, starting with 8-band $\mathbf{k} \cdot \mathbf{p}$ band structure simulations to various radio frequency MOSFET electrical and high-frequency characterization and benchmarking of the current work with the existing experimental work.

5.1 8-BAND $\mathbf{K} \cdot \mathbf{P}$ BAND STRUCTURE SIMULATION RESULTS

In Chapter 2, 8-band $\mathbf{k} \cdot \mathbf{p}$ model and its implementation in the COMSOL FEM solver are described in detail for 2D and 1D confined heterostructures. This section discusses the band structure simulation results and calculated electronic parameters behavior with strain and confinement. Bandgap is calculated at $\mathbf{\Gamma}$-point or at the wave vector, $\mathbf{k} = 0$. It is defined as the energy difference between the lowest conduction band subband and the highest valence band subband. The effective mass of the first conduction band subband $\left(\frac{1}{m_1^*} = \frac{1}{\hbar^2}\frac{\partial E_1^2(k)}{\partial k^2}\right)$ is calculated from the curvature of the $E - k$ relation at the same $\mathbf{\Gamma}$-point as well. Unless otherwise stated, All calculations are performed at temperature T = 300 K. All the subband energy levels are calculated and given with respect to InP valence band energy.

5.1.1 $\mathbf{IN_X GA_{1-X} AS/INP}$ QUANTUM WELL

Paper III discusses the electronic parameters of strained $In_xGa_{1-x}As/InP$ quantum well in detail. In this work, $In_xGa_{1-x}As/InP$ heterostructure quantum well band structure is calculated for technologically relevant quantum well thicknesses, H = 5, 7, 10 and 13 nm and compositions x = 0.2 to 1.

Electronic parameters in confined heterostructures are different from the bulk due to the inclusion of quantum confinement and strain. In Figure 5.1 (a), band-edge diagram illustrating CB and various VBs of $In_xGa_{1-x}As$ and InP, the lowest CB subband energy level, E_{c1} and highest VB subband energy, E_{v1} of strained (quantum confinement + strain) and unstrained (only quantum confinement) quantum well are plotted for various quantum well thickness and composition. The *lh* and *hh* degeneracy are removed except for lattice-matched condition; *lh* has higher energy than *hh* for bi-axial tensile strain or for $x < 0.53$ and the opposite for bi-axial compressive strain or for $x > 0.53$. Conduction band offset is defined as $CBO = E_{c,InP} - E_{c1}$ and it is the measure of the carriers' confinement in the quantum well. 13 nm InAs/InP has the largest CBO of approximately 0.4 eV. In Figure 5.1 (b) and (c) strained and unstrained $In_xGa_{1-x}As$/InP quantum well's E_g and m_1^* are plotted for various quantum well thickness and indium composition. Thin quantum wells' bandgap at a given composition is larger than thick quantum wells due to the strong quantum confinement effect, as the bi-axial strain is independent of H. Strained E_g and m_1^* are larger than their respective unstrained values in the compressive strain region and the opposite in the tensile strain region. The increase in effective mass with the decrease in both indium composition and well thickness is due to strong confinement, which increases the wavefunction leakage into the InP and nonparabolicity. The nonparabolicity factor of E_{c1} in the compressively strained quantum well is smaller than the unstrained value and the opposite for the tensile strained quantum well. It is concluded that neglecting strain in band structure simulations would result in inaccuracy in the magnitude of electronic parameters.

5.1.2 $IN_XGA_{1-X}AS$/INP NANOWIRE

Paper IV is focused on the electronic parameters behavior of strained $In_xGa_{1-x}As$/InP nanowires. Strain is inhomogenous in the case of heterostructure nanowires, and calculating strain tensor is not exact. In COMSOL Multiphysics, solid mechanics coupled with the coefficient form PDE is solved to calculate the eigenstates. Nanowires with H = 5, 13 nm and smallest W_B = $2\times H$ and wider 80 nm are considered for strain analysis. In Figure 5.2(a) and (b), the change in bandgap due to strain defined as $\Delta E_g(\%) = \frac{E_{g,strain} - E_{g,unstrain}}{E_{g,unstrain}}$ vs. uniaxial strain and ΔE_g vs. W_B are plotted, respectively. The uniaxial strain is $\epsilon_{\parallel} - \epsilon_{\perp}$, where $\epsilon_{\parallel}$ is the lattice-mismatch and $\epsilon_{\perp} = \frac{\epsilon_{\parallel}}{\sigma}$ is the perpendicular-to-plane strain, and σ is Poisson's ratio. It is observed that the ΔE_g behavior is nanowire dimension dependent. Compressively strained nanowires, especially smallest width InAs/InP strained nanowires E_g are smaller than that of unstrained nanowires, whereas, for larger nanowire

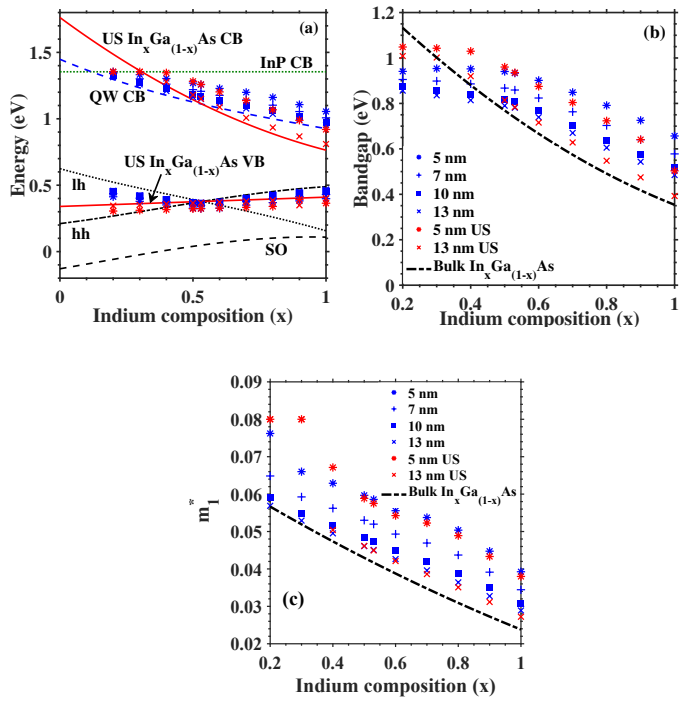

Figure 5.1: (a) Band edge diagram showing CB of bulk InP, bulk CB, and various VBs of strained and unstrained $In_xGa_{1-x}As$, E_{c1} and E_{v1} of the strained (blue symbols) and unstrained (red symbols) $In_xGa_{1-x}As$/InP well for various QW thickness. All energy levels are given with respect to InP VB. (b) Bandgap of strained, unstrained $In_xGa_{1-x}As$/InP QW with various thicknesses and bulk $In_xGa_{1-x}As$ bandgap vs. indium composition. (c) The m_1^* of strained and unstrained QW of various thickness, bulk $In_xGa_{1-x}As$ vs. indium composition. The $m_1^* = 0.08m_0$ corresponds to wave function localization in InP. For both plots, in the strained case, H = 5 nm (asterisk), 7 nm (plus), 10 nm (white square, filled), and 13 nm (multiplication), and for the unstrained case, only H = 5 and 13 nm, are shown.

widths, the strained bandgap is slightly larger than unstrained nanowire. The behavior of compressively strained larger-width nanowires is similar to that of compressively strained quantum wells. The change in the effective mass of the first conduction band subband due to strain is defined as $\Delta m_1^*(\%) = \frac{m_{1,strain} - m_{1,unstrain}}{m_{1,unstrain}}$ and is plotted against the uniaxial strain in Figure 5.3(a). The effective mass reduces as the compressive strain increases, irrespective of the nanowire size, and Δm_1^* has become width-dependent in the tensile region.

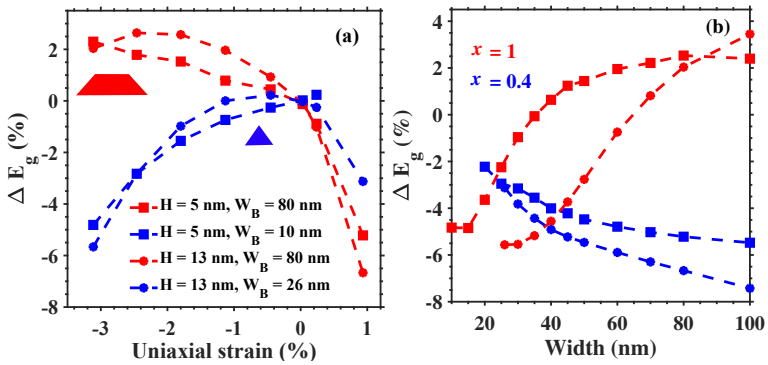

Figure 5.2: (a) ΔE_g vs. uniaxial strain, (b) ΔE_g vs. nanowire width, W_B for various nanowire heights, H = 5 nm (square) and 13 nm (star).

InAs/InP with smaller dimensions have a smaller decrease in effective mass compared to wider nanowires, and an increase in effective mass in compressively strained wider nanowires is expected, exhibiting similar behavior as quantum wells. Δm_1^* of the nanowire with H = 13 nm is approximately linear with uniaxial strain, whereas it is nonlinear for H = 5 nm. Hence, Δm_1^* behavior with nanowire width is very complex due to the interplay between quantum confinement and inhomogeneous strain, as shown in Figure 5.3 (b). When x = 1, Δm_1^* decreases as the width of the nanowire decreases till a certain width, and then, it increases irrespective of the nanowire height. The increase in Δm_1^* at narrow widths could be due to the quantum confinement effect dominating the strain in the nanowire. In the tensile strain region or x = 0.4, a slight increase in Δm_1^* is observed as the width increases, and after a certain width, Δm_1^* reduces.

5.1.3 LONG L_G MOSFET MODELING

A 13 nm thick $In_{0.71}Ga_{0.29}As$ MOSFETs with L_g = 6 μm and gate width, W_g = 60 μm are fabricated, and I-V, low-frequency C-V measurements are conducted. In Figure 5.4 (a) and (b), measured and modeled transfer and low-frequency C-V at V_{DS} = 50 mV are shown, respectively. The model uses the Urbach parameter (E_0), threshold voltage, electron mobility, and interface trap density as fitting parameters. All electronic parameters are obtained from the 8-band $\mathbf{k} \cdot \mathbf{p}$ simulations at both T = 300 K, and 13 K. A good agreement is found between measured and modeled data. An interesting observation is that the effect of interface traps is negligible at 13 K compared to RT, which could be due to multi-phonon-activated interface traps at room

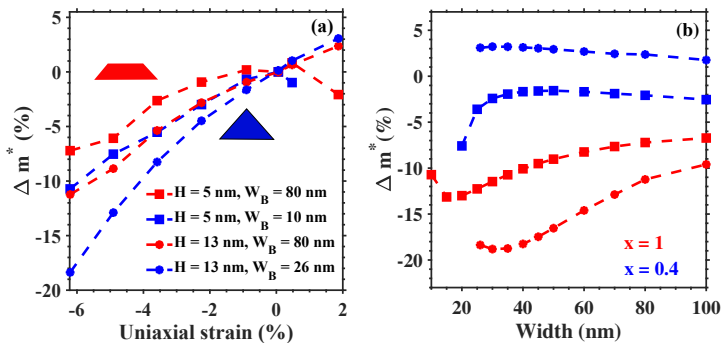

Figure 5.3: (a) Δm_1^* vs. uniaxial strain, (b) Δm_1^* vs. nanowire width, W_B for various nanowire heights, H = 5 nm (square) and 13 nm (star).

temperature. The measured and modeled minimum subthreshold swing at 300 K is 78 mV/dec for both; at 13 K, they are 19 mV/dec and 18.4 mV/dec, respectively. A distributed RC model, including interface trap capacitance (C_{it}), is implemented to fit the measured low-frequency C-V. The band tail contribution is observed in the steepness of the transition between above and below the threshold, as well as the sharpness of the subthreshold slope. The second order shift in the subband, E_1 in strong inversion, is dominant and the cause for the increase in capacitance, and it can be observed at 13 K. The sharp increase in measured capacitance at both temperatures in the ON state is due to electron accumulation in the gate probing pad, which is placed on the semi-insulating InP substrate. In Figure 5.4(c), measured and modeled gated Hall carrier concentrations at T = 13 K are plotted as a function of gate voltage, and the sheet carrier concentration increases with the applied gate voltage.

5.2 INGAAS RADIO FREQUENCY MOSFETS

In the second half of this chapter, electrical and high-frequency characterization of InGaAs radio frequency transistors with different spacer technologies are presented. Mainly, f_T/f_{max}, parasitic capacitances, and transconductance are compared for all the transistor devices fabricated in this work.

5.2.1 $IN_{0.85}GA_{0.15}AS$ NANOWIRE MOSFETS WITH INP ∧-RIDGE SPACERS

InGaAs nanowire RF MOSFET with InP, ∧-ridge spacers are fabricated on an InP semi-insulating (100) substrate. The novel δ-doped InP ∧-ridge spacers

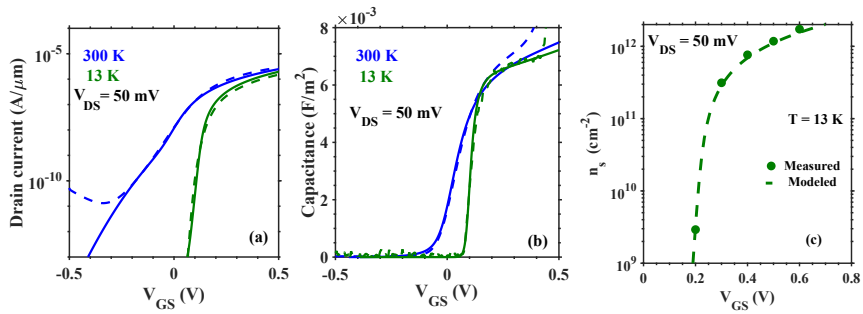

Figure 5.4: Measured (dash line) and modeled (solid line) (a) drain current vs. V_{GS}, (b) low-frequency gate capacitance vs. gate voltage at T = 300 K (blue) and T = 13 K (green), (c) measured and modeled channel carrier density against V_{GS} at T = 13 K.

have two-fold advantages: First, a wide bandgap InP spacer would increase the breakdown voltage, and second, selective growth of the spacer does not increase the contact resistance. Paper II elaborates on the fabrication and characterization of these MOSFETs in detail. RF transistors of gate length, L_g = 40 nm with InP spacer width of 80 nm and spacer-less are designed and fabricated for the performance comparison analysis. In Figure 5.5 (a) and (b), the measured gate-to-source and gate-to-drain capacitance are plotted, respectively, as a function of bias voltages for devices with and without InP spacer. The parasitic gate capacitance of C_{ggp} = 1.0 fF/μm is achieved. The measured parasitic capacitances are validated with the help of COMSOL electrostatic simulations. Figure 5.5 (c) plots the simulated gate-to-source or gate-to-drain parasitic capacitance for two scenarios with and without InP spacers. It was concluded that the RF MOSFETs with spacer allow relaxation of gate-contact alignment as this would give the identical capacitances as that of spacer-less devices with a slight overlap. The device with gate length 32 nm and width 2×7 μm has exhibited f_T/f_{max} = 75/100 GHz.

5.2.2 $IN_{0.71}GA_{0.29}AS$/INP QUANTUM WELL MOSFETS WITH SACRIFICIAL A:SI SPACERS

Paper I presents the fabrication, electrical characterization, and high-frequency analysis of $In_{0.71}Ga_{0.21}As$/InP (12 nm / 2 nm) quantum well MOSFET. Self-aligned and sacrificial a:Si spacer reduces the metal gate overlap on heavily doped contacts to achieve the least parasitic capacitance. A device with the gate length L_g = 80 nm and W_g = 2 × 20 μm has peak $g_{me,max}$ = 1 mS/μm at V_{DS}=0.5 V. Figure 5.6 (a) shows a cross-section schematic of the

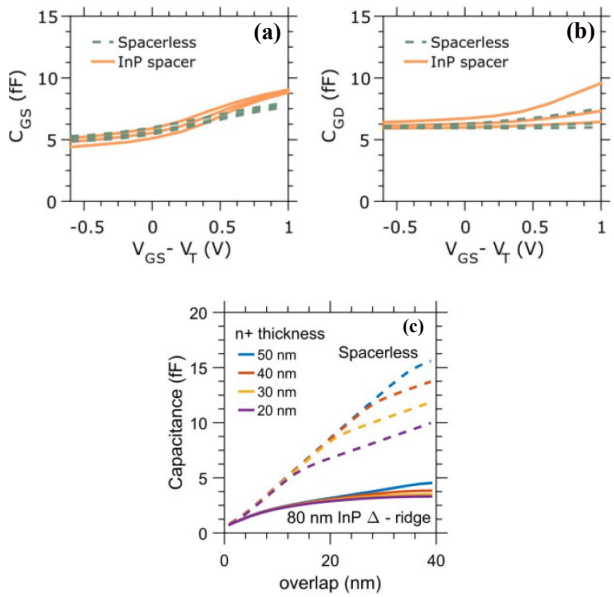

Figure 5.5: (a) Gate to source and (b) gate to drain capacitance of MOSFETs with InP spacer width of 80 nm and spacer-less at various bias voltages. Drain to source voltage, V_{DS} = 0.2, 0.6 and 1 V. (c) Simulated overlap capacitance against the direct overlap on n^+ contact and 80 nm InP spacer.

device with gate intrinsic and intrinsic parasitic capacitances. To understand various capacitance components in the device, gate-to-source, gate-to-drain, and source-to-drain total parasitic capacitance are plotted for various gate widths in Figure 5.6 (b). The linear fit of the capacitances are as follows, $C_{gs,p}$ = 0.25 fF/μm + 4.4 fF, $C_{gd,p}$ = 0.3 fF/μm + 3.7 fF and $C_{sd,p}$ = 0.2 fF/μm + 1.6 fF. The parasitic capacitances achieved here are smaller than the InP $\wedge$-ridge spacer MOSFET process. $C_{gs,ip}$ (or $C_{gd,ip}$) vs. the gate contact overlap that is calculated from COMSOL Multiphysics electrostatic simulations for various spacer thicknesses (t_{sp}) are shown in Figure 5.6 (c). The intrinsic parasitic capacitance increases with the gate contact overlap, and it is expected in the fabricated MOSFET to be around 3 to 5 nm due to possible a:Si erosion during the DG removal. The parasitic capacitance also saturates with the increase in spacer thickness, and there is not much to gain in capacitances for spacer thickness above 50 nm.

Intrinsic gate-to-source and gate-to-drain capacitances ($C_{gs,i}$ and $C_{gd,i}$) are obtained by subtracting the OFF-state capacitance from ON-state capacitance

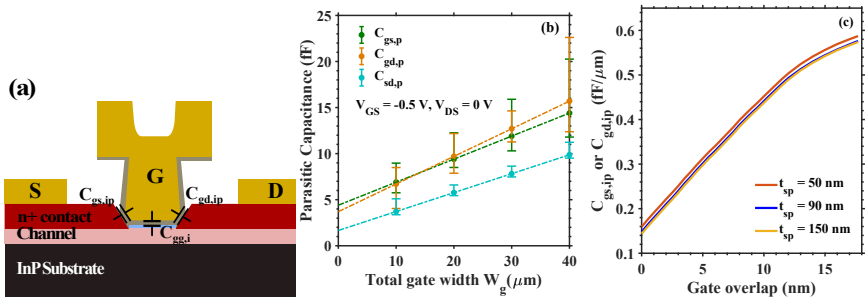

Figure 5.6: (a) Schematic illustration of $C_{gg,i}$, $C_{gs,ip}$ and $C_{gd,ip}$ on device cross-section, (b) various total measured parasitic capacitances versus total gate width and their linear fit are shown in dashed line, (c) calculated $C_{gs,ip}$ or $C_{gd,ip}$ for various gate contact overlap on regrown contacts for various spacer thickness.

at different bias conditions. They are shown against device L_g in Figure 5.7 (a). Both $C_{gs,i}$ and $C_{gd,i}$ scales with the gate length in the triode region, whereas in saturation or at higher V_{DS}, $C_{gd,i}$ is almost constant and negligible. Small-signal model, including high-κ gate oxide border traps, interface trap loss, impact ionization, and band-to-band tunneling, are used to model high-frequency characteristics. The calculated f_T/f_{max} using small-signal model vs. device gate width are plotted in Figure 5.7 (b). Transition frequency increases with device width because the extrinsic parasitic capacitance becomes smaller than intrinsic parasitic as W_g increases. FET with L_g = 80 nm and $W_g = 2 \times 20$ μm has exhibited f_T/f_{max} = 243/147 GHz. On the other hand, f_{max} decreases with an increase in W_g due to an increase in gate resistance. The performance of these devices was limited by high contact resistance and extrinsic constant parasitic capacitance due to metal routing. The metal contacts layout optimization, i.e., reducing the separation between two gate fingers and moving the reference plane close to the device, would be expected to decrease the extrinsic parasitic capacitances. This contact layout optimization is performed in the subsequent versions of RF MOSFETs.

Pulsed IV characteristics are performed on these transistors, and the chosen pulse widths are 50 ns, 500 ns, and 5 μs. These measurements provide information about the oxide traps and self-heating effects and show frequency dispersion at low-frequency analysis. In Figure 5.8 (a) and (b), the transfer and output characteristics are plotted for different pulse widths. With reducing pulse width, an increase in $g_{me,max}$, shift in V_T, and improvement in SS_{min} are observed. The measured SS_{min} of DC and short pulse width are 150 mV/dec

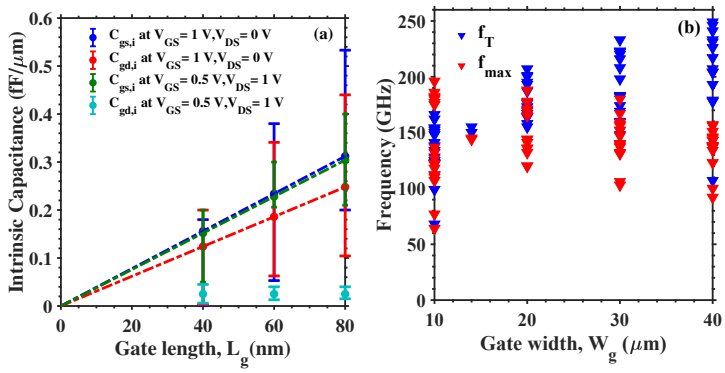

Figure 5.7: (a) Intrinsic capacitances normalised to W_g versus L_g and their linear fit is shown in dashed line. (b) Extracted peak f_T and f_{max} from small-signal model vs. gate width, W_g.

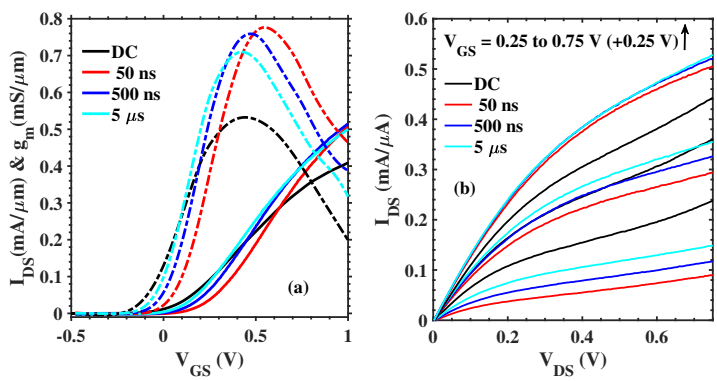

Figure 5.8: DC and pulsed IV of a quantum well InGaAs MOSFET at room temperature. (a) Transfer characteristics at V_{DS} = 0.5 V and (b) output characteristics of the device with L_g = 60 nm, W_g = 20 μm.

and 90 mV/dec, respectively. There is a nearly 50% increase in $g_{me,max}$ for a pulse width of 50 ns compared to DC $g_{me,max}$. No significant change in R_{ON} and g_{de} is observed. These measurements provide information about the oxide traps and self-heating effects and exhibit frequency dispersion at low-frequency analysis.

5.2.3 INGAAS COMPOSITE CHANNEL QUANTUM WELL MOSFET WITH SIN SPACERS

InGaAs composite channel MOSFETs are fabricated with 40 nm thick SiN spacers in a top-down approach. In the first attempt to fabricate these devices, the misalignment in drain contact resulted in constant leakage current in the device's right finger and asymmetric finger current levels. However, from the electrical characterization of transistors and transfer length methods (TLMs), DC performance metrics and resistances are extracted. Figure 5.9 plots the transfer and output characteristics of two transistors with gate length 80 nm and different source to drain contact separation. A transistor with L_{sd} = 340 nm has $g_{me,max}$ = 1.16 mS/μm and L_{sd} = 550 nm has $g_{me,max}$ = 1.0 mS/μm. The SS_{min} is quite large due to bad OFF-state. R_{ON} is calculated from the output characteristics at smaller V_{DS} when the transistor operates in the linear region. The R_{ON} of transistor with L_{sd} = 340 nm has 200 $\Omega - \mu m$ and L_{sd} = 550 nm has 262 $\Omega - \mu m$. Higher R_{ON} is expected for larger L_{sd}, as this device has a larger access region. The extrinsic output conductance, g_{de}, is calculated in the saturation. The magnitude of g_{de} is in the order of 1.2 mS/μm for both the transistors, and as it can be seen that the magnitude of g_{de} is the same order of g_{me}, this has decreased the DC intrinsic device gain significantly.

The measured resistances are plotted along with the linear fit in Figure 5.10 (a). The extracted sheet resistances of n^+ cap layer (R_{sh,n^+cap}) and n-InP etch stop layer ($R_{sh,InP}$) are 40 $\Omega/\square$ and 185 $\Omega/\square$. The InP barrier resistivity, $\rho_{barrier}$ should be decreased further to reduce the contact resistance. The calculated barrier resistance here is similar to high-performance InP HEMTs [22]. Transfer lengths in both barrier layer ($L_{T,access}$) and contact layer ($L_{T,c}$) and contact resistivity (ρ_c) are also calculated and all are summarised in Table 5.1.

Parameter	Value
$R_{sh,n^+cap}(\Omega/\square)$	40
ρ_c $(\Omega - cm^2)$	1×10^{-8}
$L_{T,c}$ (μm)	0.25
$R_{sh,InP}(\Omega/\square)$	185
$\rho_{barrier}(\Omega - cm^2)$	1.8×10^{-6}
$L_{T,access}(\mu$m)	0.61

Table 5.1: Summary of the results obtained from TLMs.

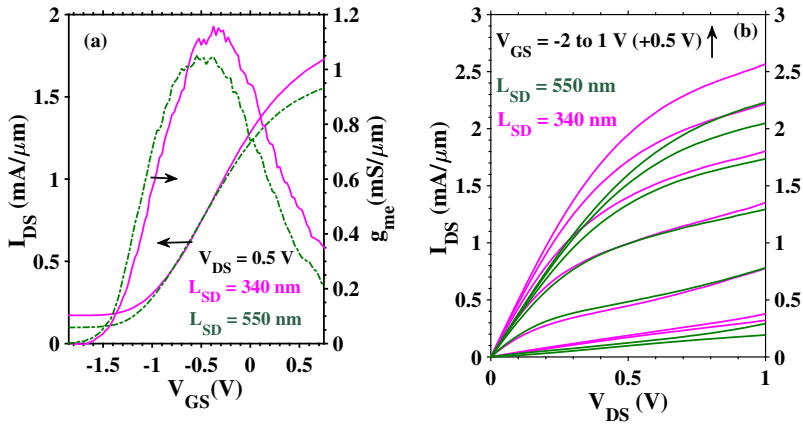

Figure 5.9: (a) Transfer and (b) output characteristics of InGaAs CC MOSFETs L_g = 80 nm and W_g = 2 × 20 μm. Two devices have a different source to drain contact separation, L_{sd} = 550 and 340 nm. Transistor with smaller L_{sd} has slightly higher maximum transconductance.

High-frequency characterization is performed, and the parasitic capacitance is more prominent due to significant leakage current. In Figure 5.10(b), the OFF-state parasitic capacitances are plotted for various device widths, and from the linear fit, the total gate width dependent capacitance is 0.7 fF/μm, and constant parasitic capacitance is 22 fF, which is expected to decrease with correct drain contact alignment. The expected total gate width capacitance with the given spacer and designed gate contact overlap is 0.65 fF/μm. High-frequency metrics are not significant due to large capacitances and larger g_{de}, the best device with L_g = 80 nm and W_g = 2× 25 μm, has f_T/f_{max} = 207/90 GHz.

5.2.4 HIGH-FREQUENCY SMALL-SIGNAL MODELING

The high-frequency small-signal modeling is crucial for accurately estimating the intrinsic device parameters and f_T/f_{max}. The transistor with the gate length, L_g = 80 nm and W_g = 2 × 12 μm is fabricated with minimised contact misalignment. The total gate constant parasitic capacitance post-correction of contact misalignment is 2.5 fF. The measured total gate-to-source and gate-to-drain capacitance (C_{gs} and C_{gd}) behavior with external bias voltages is plotted in Figure 5.11(a). C_{gs} increases with both V_{GS} and V_{DS}, whereas the C_{gd} decreases as the V_{DS} increases. The ON-state C_{gd} at larger V_{DS} is smaller than parasitic capacitance, indicating strong channel depletion on the gate-to-drain

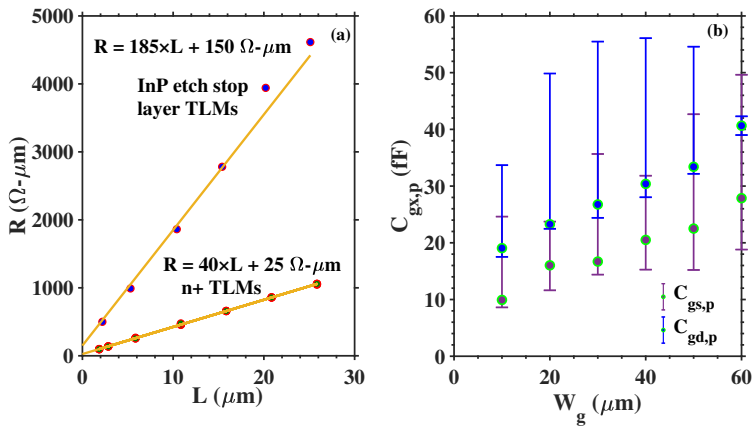

Figure 5.10: (a) n^+ cap layer and n-InP barrier layer TLMs. The measured resistances are normalized to width and are plotted for different contact separations. The linear fit is plotted in a solid line, displaying the slope and Y-intercept. (b) Gate-to-source and gate-to-drain parasitic capacitances are plotted for total device width. The linear increase in capacitance with device width is highlighted, and the spread in capacitance is due to significant device-to-device variation because of contact misalignment.

access region. The measured and modeled unity power gain and MAG or MSG are plotted for three external bias voltages in Figure 5.11(b). The non-zero and significant series resistances of C_{gd} and $C_{sd,i}$ are considered to get a good fit to intrinsic Y-parameters. Interestingly, unity power gain has three different slopes depending on the frequency range; at low frequencies, $|U|$ exhibits -10 dB/dec, indicating conductance loss due to traps in gate oxide. It has a -20 dB/dec in a few tens of GHz range and roll-off larger than -20 dB/dec close to f_{max}. Simple -20 dB/dec extrapolation of $|U|$ would only overestimate the magnitude of f_{max}. It is interesting to see this particular FET has a high gain of MSG = 19 dB at V_{DS} = 1 V and V_{GS} = 0.5 V due to smaller C_{gd} = 0.16 fF/μm. Paper VI summarises the finding of nontraditional $|U|$ roll-off behavior in depth.

5.2.5 BENCHMARKING

A summary of the RF MOSFETs fabricated in this work is given in Table 5.2. Further, the modeled S-parameters are imported into the Keysight ADS, and maximum available gain, minimum noise figure, and P_{DC} are summarised for three types of FETs evaluated at 60 GHz in Table 5.3. Figure 5.12 compares

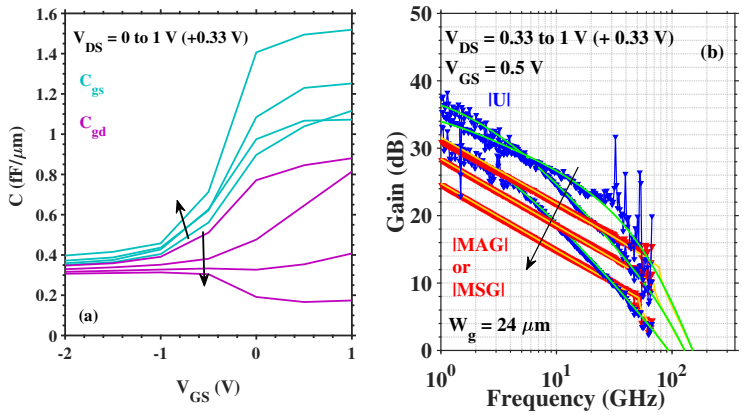

Figure 5.11: (a) Total gate-to-source and gate-to-drain capacitance vs. V_{GS} for different V_{DS}. (b) Using a full non-quasi-static model, measured and modeled $|U|$ and |MAG| or |MSG| vs. frequency.

this work with the existing high-performance FETs. The total intrinsic gate capacitance and RF g_{me} are the metrics for comparison.

Device	L_g/W_g (nm/μm)	$C_{gg,p}$ (fF/μm)	RF g_{me} (mS/μm)	f_T/f_{max} (GHz)	MSG (dB) at 20 GHz
[104] InP $\wedge - ridge$	32/14	1	0.5	75/100	9.3
[108] a:Si	80/40	0.8	1.7	243/147	15.4
Nitride	80/50	0.72	1.5	207/90	9.9
Nitride	80/24	0.7	1.53	160/150	19

Table 5.2: Summary of various RF MOSFETs fabricated in this work.

Device	L_g/W_g (nm/μm)	f_T/f_{max} (GHz)	MAG/NF (dB)	Gain/NF_{min} (dB)	P_{DC} (mW)
[104] InP $\wedge - ridge$	32/14	75/100	4.4/1.3	-	5.32
[108] a:Si	80/40	243/147	11/7	5.4/1.35	26
Nitride	80/24	160/150	14.6/5.8	6.6/0.774	10.5

Table 5.3: Small-signal model parameters used for circuit design.

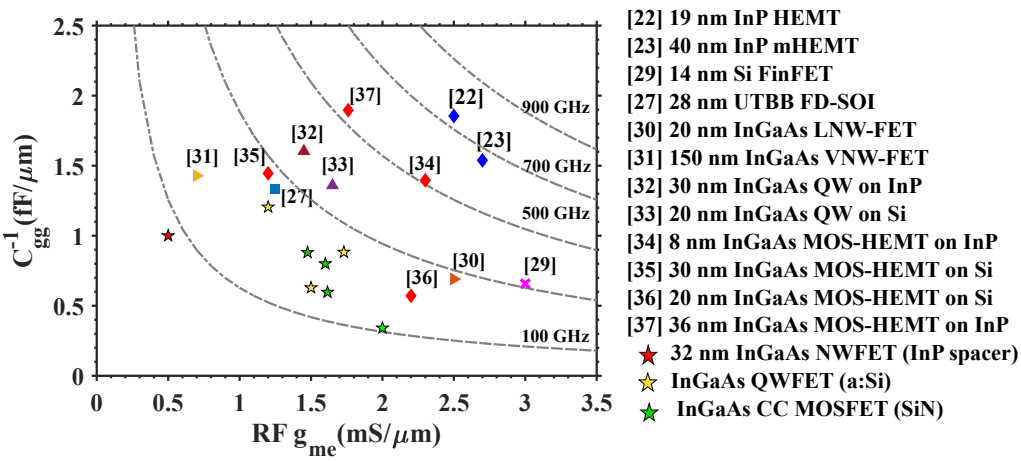

Figure 5.12: Benchmarking of RF g_{me} and $\frac{1}{C_{gg}}$ of this work and various high-performance FET technologies. Gate capacitance is estimated from $f_T \approx \frac{g_{me}}{2\pi C_{gg}}$.

6

Summary and Future Work

6.1 SUMMARY

Here is a summary of the work I have been involved in during my Ph.D. studies and the results published in peer-reviewed journals.

In Paper III and IV, the semi-empirical 8-band $\mathbf{k} \cdot \mathbf{p}$ method is used to calculate the electronic parameters of strained $In_xGa_{1-x}As/InP$ heterostructure-based quantum wells and nanowires. The work highlights the importance of considering strain in calculating the electronic parameters, which are used to model long channel quantum well FETs, mainly current and capacitances.

In Paper I, $In_{0.71}Ga_{0.29}As/InP$ quantum well MOSFETs with sacrificial a:Si are fabricated, and they have exhibited small parasitic capacitances, almost the same magnitude as that of high-performing InGaAs/InP MOSFETs and HEMTs. However, f_T/f_{max} of these MOSFETs is limited by large contact resistance due to the extra interfacial resistance layer between channel and regrown contacts.

In Paper II, we tried to introduce a novel InP $\wedge$-spacer into the InGaAs nanowire MOSFETs to reduce capacitances; however, due to a larger overlap of gate contact with undepleted and high permittivity, the InP spacer has resulted in small f_T/f_{max}. However, this process relaxes the limit on gate contact alignment compared to spaceless devices to get the same capacitances.

In Paper VI, InGaAs composite channel quantum well MOSFET fabricated using a top-down approach, and a nitride spacer is used in this self-aligned process. For example, further process development can improve f_T/f_{max} by increasing the spacer thickness. An extended small-signal model is used to fit the measured Y-parameters, and more than -20 dB/dec roll-off in $|U|$ at high frequencies is observed. The traditional -20 dB/dec extrapolation of $|U|$

generally results in overestimating f_{max}, and a precise small-signal model is in demand.

In Paper V, the theoretical high-frequency performance of ballistic 3D-nanosheet/nanowire FETs is studied for constant footprint and constant gate-effective width. 3D-vertical stacking of the channels has resulted in lower parasitic capacitances than 2D quantum well MOSFETs because of perturbation in the electric field. Another interesting observation is that, for constant transconductance or gain design, 3D-nanosheet/nanowire FETs have $\times 4$ smaller parasitic capacitance than the quantum well FETs with the additional advantage of a smaller device area. In addition, 3D-nanosheet/nanowire FETs also have significant gains at smaller current levels, which is interesting for low-power cryogenic applications.

6.2 CONCLUSION

InGaAs MOSFETs with various spacer technologies are fabricated to decrease parasitic capacitances and attempted to reach the limit of high-performance HEMTs. In the first generation of RF FETs with high permittivity, InP spacers integrated on InGaAs nanowires have large parasitic capacitances, 1 fF/μm. In the next generation, quantum well MOSFETs with air spacers have resulted in the lowest parasitic capacitance of 0.55 fF/μm, but high contact resistance has limited f_T/f_{max}. InGaAs composite channel FETs with nitride spacers have exhibited parasitic capacitance of 0.7 fF/μm, and further optimization of this process could improve high-frequency metrics. The nitride spacer process is promising among realized RF FETs in achieving high f_T/f_{max} with low capacitances, especially C_{gd} and smaller contact resistances. The highest and best combination of f_T/f_{max} = 243/147 GHz is achieved for InGaAs quantum well MOSFETs with sacrificial a:Si spacers, followed by nitride spacer FETs with f_T/f_{max} = 160/150 GHz.

The long channel and ballistic devices are also modeled during this thesis work. Cryogenic I-V and C-V models, including band tails and interface trap density, are modeled to validate the measured data and understand device performance limitations at low temperatures. At the end of the thesis, 3D InGaAs/InP 3D nanowire and nanosheet FETs are proposed, and their RF performance is studied. Due to the unavailability of TCAD tools for device simulations, the complex design is separated into the intrinsic part, modeling ballistic current and intrinsic gate capacitances, and the extrinsic part to model the fringe or parasitic capacitances. Vertical stacking of nanosheets/wires results in larger drive currents and lower parasitics due to electric field screening in a given physical footprint. These advantages come with a challenging fabrication price, including fins etching, controlled channel release,

and contacts regrowth, to name a few. However, this study is not limited to InGaAs/InP and can easily be transferred to Si/SiGe nanosheets/nanowires.

The electronic parameters behavior analysis of InGaAs/InP quantum well and nanowires has highlighted the importance of strain consideration in band structure simulations. Particularly in the case of heterostructure nanowires where the strain distribution becomes inhomogeneous and electronic parameters become dependent on nanowire composition and dimension. Implementing this semi-empirical model in COMSOL FEM solvers makes the model diverse in terms of geometry and heterostructure design.

6.3 FUTURE WORK

Here are several potential strategies that can be employed to enhance the work presented:

1. InGaAs MOSFET with thick and low permittivity spacer and optimized gate-to-source and gate-to-drain access region and T-gate design would improve the high-frequency metrics. Making an initial circuit effort by modeling a simple LNA or RF switch would also be interesting.
2. Estimating f_T/f_{max} requires accurate measurement and modeling, especially f_{max}, due to unpredictable behavior of unity power gain at high frequencies. Accurate de-embedding like TRL or iterative de-embedding open-short de-embedding procedures can be implemented for better extraction of intrinsic device S-parameters.
3. It would be interesting to combine the developed semi-empirical 8-band $\mathbf{k} \cdot \mathbf{p}$ with the Poisson solver in COMSOL, as this would allow to model the electrical characteristics of the FETs more accurately.
4. Finally, the preliminary endeavor in modeling ballistic 3D-nanosheet transistors could be expanded to encompass more realistic device physics. This extension might involve incorporating carrier scattering mechanisms and refining the modeling of intrinsic capacitance.

Bibliography

[1] M. Guarnieri, "The Age of Vacuum Tubes: Merging with Digital Computing [Historical]," *IEEE Industrial Electronics Magazine*, vol. 6, no. 3, pp. 52–55, 2012. doi: 10.1109/MIE.2012.2207830

[2] J. Bardeen and W. H. Brattain, "The transistor, a semi-conductor triode," *Physical Review*, vol. 74, no. 2, p. 230, 1948.

[3] W. Arden, M. Brillouët, P. Cogez, M. Graef, B. Huizing, and R. Mahnkopf, "More-than-Moore white paper," *Version*, vol. 2, p. 14, 2010.

[4] G. E. Moore *et al.*, "Cramming more components onto integrated circuits," 1965.

[5] R. H. Dennard, F. H. Gaensslen, H.-N. Yu, V. L. Rideout, E. Bassous, and A. R. Leblanc, "Design of ion-implanted MOSFET's with very small physical dimensions," *Proceedings of the IEEE*, vol. 87, no. 4, pp. 668–678, 1999.

[6] D. J. Frank, "Power-constrained CMOS scaling limits," *IBM Journal of Research and Development*, vol. 46, no. 2.3, pp. 235–244, 2002.

[7] C.-H. Jan, U. Bhattacharya, R. Brain, S.-J. Choi, G. Curello, G. Gupta, W. Hafez, M. Jang, M. Kang, K. Komeyli *et al.*, "A 22nm SoC platform technology featuring 3-D tri-gate and high-k/metal gate, optimized for ultra low power, high performance and high density SoC applications," in *2012 International Electron Devices Meeting*. IEEE, 2012, pp. 3–1.

[8] H. Wong, "On the CMOS device downsizing, more Moore, more than Moore, and more-than-Moore for more Moore," in *2021 IEEE 32nd International Conference on Microelectronics (MIEL)*. IEEE, 2021, pp. 9–15.

[9] D. G. Julien Frougier, "Introducing the world's first 2 nm node chip," https://research.ibm.com/blog/2-nm-chip, 2021, [Online; accessed 15-May-2023].

[10] C. M. Compagnoni, A. Goda, A. S. Spinelli, P. Feeley, A. L. Lacaita, and A. Visconti, "Reviewing the evolution of the NAND flash technology," *Proceedings of the IEEE*, vol. 105, no. 9, pp. 1609–1633, 2017.

[11] J. A. Del Alamo, "Nanometre-scale electronics with III-V compound semiconductors," *Nature*, vol. 479, no. 7373, pp. 317–323, 2011.

[12] D.-H. Kim, J. Del Alamo, D. Antoniadis, and B. Brar, "Extraction of virtual-source injection velocity in sub-100 nm III-V HFETs," in *2009 IEEE International Electron Devices Meeting (IEDM)*. IEEE, 2009, pp. 1–4.

[13] J. Lin, X. Cai, Y. Wu, D. A. Antoniadis, and J. A. del Alamo, "Record Maximum Transconductance of 3.45 ms/μm for III-V FETs," *IEEE Electron Device Letters*, vol. 37, no. 4, pp. 381–384, 2016.

[14] C. B. Zota, L.-E. Wernersson, and E. Lind, "Single suspended InGaAs nanowire MOSFETs," in *2015 IEEE International Electron Devices Meeting (IEDM)*. IEEE, 2015, pp. 31–4.

[15] O.-P. Kilpi, M. Hellenbrand, J. Svensson, A. R. Persson, R. Wallenberg, E. Lind, and L.-E. Wernersson, "High-Performance Vertical III-V Nanowire MOSFETs on Si With $g_m > 3$ mS/μm," *IEEE Electron Device Letters*, vol. 41, no. 8, pp. 1161–1164, 2020.

[16] P. Ye, G. Wilk, J. Kwo, B. Yang, H.-J. Gossmann, M. Frei, S. Chu, J. Mannaerts, M. Sergent, M. Hong *et al.*, "GaAs MOSFET with oxide gate dielectric grown by atomic layer deposition," *IEEE Electron Device Letters*, vol. 24, no. 4, pp. 209–211, 2003.

[17] J. Choi, Y. Mao, and J. Chang, "Development of hafnium based high-k materials-A review," *Materials Science and Engineering: R: Reports*, vol. 72, no. 6, pp. 97–136, 2011.

[18] J. Choi, V. Va, N. Gonzalez-Prelcic, R. Daniels, C. R. Bhat, and R. W. Heath, "Millimeter-Wave Vehicular Communication to Support Massive Automotive Sensing," *IEEE Communications Magazine*, vol. 54, no. 12, pp. 160–167, 2016. doi: 10.1109/MCOM.2016.1600071CM

[19] J. Hasch, E. Topak, R. Schnabel, T. Zwick, R. Weigel, and C. Waldschmidt, "Millimeter-wave technology for automotive radar sensors in the 77 GHz frequency band," *IEEE transactions on microwave theory and techniques*, vol. 60, no. 3, pp. 845–860, 2012.

[20] R. Lai, X. B. Mei, W. R. Deal, W. Yoshida, Y. M. Kim, P. H. Liu, J. Lee, J. Uyeda, V. Radisic, M. Lange *et al.*, "Sub 50 nm InP HEMT Device with F_{max} Greater than 1 THz," in *2007 IEEE International Electron Devices Meeting*, 2007. doi: 10.1109/IEDM.2007.4419013 pp. 609–611.

[21] X. Mei, W. Yoshida, M. Lange, J. Lee, J. Zhou, P.-H. Liu, K. Leong, A. Zamora, J. Padilla, S. Sarkozy *et al.*, "First Demonstration of Amplification at 1 THz using 25-nm InP High Electron Mobility Transistor Process," *IEEE Electron Device Letters*, vol. 36, no. 4, pp. 327–329, 2015. doi: 10.1109/LED.2015.2407193

[22] H.-B. Jo, S.-W. Yun, J.-G. Kim, D.-Y. Yun, I.-G. Lee, D.-H. Kim, T.-W. Kim, S.-K. Kim, J. Yun, T. Kim *et al.*, "L_g= 19 nm $In_{0.8}Ga_{0.2}As$ composite-channel HEMTs with f_T= 738 GHz and f_{max}= 492 GHz," in *2020 IEEE International Electron Devices Meeting (IEDM)*. IEEE, 2020, pp. 8–4.

[23] D.-H. Kim, B. Brar, and J. A. del Alamo, "f_T = 688 GHz and f_{max} = 800 GHz in L_g = 40 nm $In_{0.7}Ga_{0.3}As$ mHEMTs with $g_{m,max}$ > 2.7 ms/μm," in *2011 International Electron Devices Meeting*, 2011. doi: 10.1109/IEDM.2011.6131548 pp. 13.6.1–13.6.4.

[24] D.-H. Kim and J. A. Del Alamo, "Scalability of sub-100 nm InAs HEMTs on InP substrate for future logic applications," *IEEE transactions on electron devices*, vol. 57, no. 7, pp. 1504–1511, 2010.

[25] C.-H. Jan, M. Agostinelli, H. Deshpande, M. El-Tanani, W. Hafez, U. Jalan, L. Janbay, M. Kang, H. Lakdawala, J. Lin *et al.*, "RF CMOS technology scaling in high-k/metal gate era for RF SoC (system-on-chip) applications," in *2010 International Electron Devices Meeting*. IEEE, 2010, pp. 27–2.

[26] R. Carter, J. Mazurier, L. Pirro, J. Sachse, P. Baars, J. Faul, C. Grass, G. Grasshoff, P. Javorka, T. Kammler *et al.*, "22nm FDSOI technology for emerging mobile, Internet-of-Things, and RF applications," in *2016 IEEE International Electron Devices Meeting (IEDM)*. IEEE, 2016, pp. 2–2.

[27] B. K. Esfeh, V. Kilchytska, V. Barral, N. Planes, M. Haond, D. Flandre, and J.-P. Raskin, "Assessment of 28 nm UTBB FD-SOI technology platform for RF applications: Figures of merit and effect of parasitic elements," *Solid-State Electronics*, vol. 117, pp. 130–137, 2016.

[28] B. Sell, B. Bigwood, S. Cha, Z. Chen, P. Dhage, P. Fan, M. Giraud-Carrier, A. Kar, E. Karl, C.-J. Ku *et al.*, "22FFL: A high performance and ultra low power FinFET technology for mobile and RF applications," in *2017 IEEE International Electron Devices Meeting (IEDM)*. IEEE, 2017, pp. 29–4.

[29] J. Singh, J. Ciavatti, K. Sundaram, J. Wong, A. Bandyopadhyay, X. Zhang, S. Li, A. Bellaouar, J. Watts, J. Lee *et al.*, "14-nm FinFET technology for analog and RF applications," *IEEE Transactions on electron devices*, vol. 65, no. 1, pp. 31–37, 2017.

[30] C. Zota, F. Lindelöw, L.-E. Wernersson, and E. Lind, "High-frequency InGaAs tri-gate MOSFETs with f_{max} of 400 GHz," *Electronics Letters*, vol. 52, no. 22, pp. 1869–1871, 2016.

[31] S. Johansson, E. Memisevic, L.-E. Wernersson, and E. Lind, "RF characterization of vertical InAs nanowire MOSFETs with f_t and f_{max} above 140 GHz," in *26th International Conference on Indium Phosphide and Related Materials (IPRM)*. IEEE, 2014, pp. 1–2.

[32] J. Wu, Y. Fang, B. Markman, H.-Y. Tseng, and M. J. W. Rodwell, "$L_g = 30$ nm InAs Channel MOSFETs Exhibiting $f_{max} = 410$ GHz and $f_t = 357$ GHz," *IEEE Electron Device Letters*, vol. 39, no. 4, pp. 472–475, 2018. doi: 10.1109/LED.2018.2803786

[33] C. B. Zota, C. Convertino, M. Sousa, D. Caimi, K. Moselund, and L. Czornomaz, "High-Frequency Quantum Well InGaAs-on-Si MOSFETs With Scaled Gate Lengths," *IEEE Electron Device Letters*, vol. 40, no. 4, pp. 538–541, 2019. doi: 10.1109/LED.2019.2902519

[34] B. Markman, S. T. Š. Brunelli, A. Goswami, M. Guidry, and M. J. Rodwell, "$In_{0.53}Ga_{0.47}As$/InAs Composite Channel MOS-HEMT Exhibiting 511 GHz f_τ and 256 GHz f_{max}," *IEEE Journal of the Electron Devices Society*, vol. 8, pp. 930–934, 2020.

[35] D. Han, D. C. Ruiz, G. Bonomo, T. Saranovac, O. J. Ostinelli, and C. R. Bolognesi, "Low-noise microwave performance of 30 nm GaInAs MOS-HEMTs: Comparison to low-noise HEMTs," *IEEE Electron Device Letters*, vol. 41, no. 9, pp. 1320–1323, 2020.

[36] A. Tessmann, A. Leuther, F. Heinz, F. Bernhardt, L. John, H. Massler, L. Czornomaz, and T. Merkle, "20-nm $In_{0.8}Ga_{0.2}As$ MOSHEMT MMIC technology on silicon," *IEEE Journal of Solid-State Circuits*, vol. 54, no. 9, pp. 2411–2418, 2019.

[37] L. Whitaker, B. Markman, and M. J. Rodwell, "Self-Aligned InGaAs Channel MOS-HEMTs for High Frequency Applications," in *2023 Device Research Conference (DRC)*, 2023. doi: 10.1109/DRC58590.2023.10186996 pp. 1–2.

[38] C. J. Martínez Blanque *et al.*, "Calculation of the electronic structure and transport properties of semiconductor nanowires," 2016.

[39] Y. Peter and M. Cardona, *Fundamentals of semiconductors: physics and materials properties*. Springer Science & Business Media, 2010.

[40] G. Bastard, "Wave mechanics applied to semiconductor heterostructures," 1990.

[41] R. G. Veprek, "Computational modeling of semiconductor nanostructures for optoelectronics," Ph.D. dissertation, ETH Zurich, 2009.

[42] J. Bardeen, "An improved calculation of the energies of metallic Li and Na," *The Journal of Chemical Physics*, vol. 6, no. 7, pp. 367–371, 1938.

[43] F. Seitz and R. Johnson, "Modern theory of solids," *Uspekhi Fizicheskikh Nauk*, vol. 23, no. 3, pp. 293–315, 1940.

[44] C. Pryor and M.-E. Pistol, "Band-edge diagrams for strained III-V semiconductor quantum wells, wires, and dots," *Physical Review B*, vol. 72, no. 20, p. 205311, 2005.

[45] J. Luttinger, "Quantum theory of cyclotron resonance in semiconductors: General theory," *Physical review*, vol. 102, no. 4, p. 1030, 1956.

[46] E. O. Kane, "Band structure of indium antimonide," *Journal of Physics and Chemistry of Solids*, vol. 1, no. 4, pp. 249–261, 1957.

[47] G. Dresselhaus, "Spin-orbit coupling effects in zinc blende structures," *Physical Review*, vol. 100, no. 2, p. 580, 1955.

[48] G. Bir and G. Pikus, "Symmetry and Strain Induced Effects in Semiconductors (Wiley, New York, 1974)."

[49] J. Rychlewski, "On Hooke's law," *Journal of Applied Mathematics and Mechanics*, vol. 48, no. 3, pp. 303–314, 1984.

[50] J. Bardeen and W. Shockley, "Deformation potentials and mobilities in non-polar crystals," *Physical review*, vol. 80, no. 1, p. 72, 1950.

[51] C. Herring and E. Vogt, "Transport and deformation-potential theory for many-valley semiconductors with anisotropic scattering," *Physical review*, vol. 101, no. 3, p. 944, 1956.

[52] B. A. Foreman, "Analytical envelope-function theory of interface band mixing," *Physical review letters*, vol. 81, no. 2, p. 425, 1998.

[53] R. G. Veprek, S. Steiger, and B. Witzigmann, "Ellipticity and the spurious solution problem of $k \cdot p$ envelope equations," *Physical Review B*, vol. 76, no. 16, p. 165320, 2007.

[54] A. Rahman, J. Guo, S. Datta, and M. Lundstrom, "Theory of ballistic nanotransistors," *IEEE Transactions on Electron Devices*, vol. 50, no. 9, pp. 1853–1864, 2003. doi: 10.1109/TED.2003.815366

[55] E. Fermi, "Zur Quantelung des idealen einatomigen Gases," *Zeitschrift für Physik*, vol. 36, no. 11-12, pp. 902–912, 1926.

[56] P. A. M. Dirac, "On the theory of quantum mechanics," *Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character*, vol. 112, no. 762, pp. 661–677, 1926.

[57] Y. Taur, D. A. Buchanan, W. Chen, D. J. Frank, K. E. Ismail, S.-H. Lo, G. A. Sai-Halasz, R. G. Viswanathan, H.-J. Wann, S. J. Wind *et al.*, "CMOS scaling into the nanometer regime," *Proceedings of the IEEE*, vol. 85, no. 4, pp. 486–504, 1997.

[58] K. Suzuki, T. Tanaka, Y. Tosaka, H. Horie, and Y. Arimoto, "Scaling theory for double-gate SOI MOSFET's," *IEEE Transactions on Electron Devices*, vol. 40, no. 12, pp. 2326–2329, 1993.

[59] M.-D. Ko, C.-W. Sohn, C.-K. Baek, and Y.-H. Jeong, "Study on a scaling length model for tapered tri-gate FinFET based on 3-D simulation and analytical analysis," *IEEE transactions on electron devices*, vol. 60, no. 9, pp. 2721–2727, 2013.

[60] T.-S. Chen, "Determination of the Capacitance, Inductance, and Characteristic Impedance of Rectangular Lines," *IRE Transactions on Microwave Theory and Techniques*, vol. 8, no. 5, pp. 510–519, 1960. doi: 10.1109/TMTT.1960.1124779

[61] D. John, L. Castro, and D. Pulfrey, "Quantum capacitance in nanoscale device modeling," *Journal of Applied Physics*, vol. 96, no. 9, pp. 5180–5184, 2004.

[62] N. Neophytou, A. Paul, M. S. Lundstrom, and G. Klimeck, "Bandstructure effects in silicon nanowire electron transport," *IEEE Transactions on Electron Devices*, vol. 55, no. 6, pp. 1286–1297, 2008.

[63] A. Afzalian, C.-W. Lee, N. Dehdashti Akhavan, R. Yan, I. Ferain, and J.-P. Colinge, "Quantum Confinement Effects in Capacitance Behavior of Multigate Silicon Nanowire MOSFETs," *IEEE Transactions on Nanotechnology*, vol. 10, no. 2, pp. 300–309, 2011. doi: 10.1109/TNANO.2009.2039800

[64] J. Knoch, W. Riess, and J. Appenzeller, "Outperforming the Conventional Scaling Rules in the Quantum-Capacitance Limit," *IEEE Electron Device Letters*, vol. 29, no. 4, pp. 372–374, 2008. doi: 10.1109/LED.2008.917816

[65] F. Urbach, "The long-wavelength edge of photographic sensitivity and of the electronic absorption of solids," *Physical Review*, vol. 92, no. 5, p. 1324, 1953.

[66] H. Bohuslavskyi, A. Jansen, S. Barraud, V. Barral, M. Cassé, L. Le Guevel, X. Jehl, L. Hutin, B. Bertrand, G. Billiot *et al.*, "Cryogenic subthreshold swing saturation in FD-SOI MOSFETs described with band broadening," *IEEE Electron Device Letters*, vol. 40, no. 5, pp. 784–787, 2019.

[67] P. Rhodes, "Fermi-Dirac functions of integral order," *Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences*, vol. 204, no. 1078, pp. 396–405, 1950.

[68] S. M. Sze and M.-K. Lee, *Semiconductor Devices: Physics and Technology: Physics and Technology*. Wiley Global Education, 2012.

[69] P. Landsberg, "D grad ν or grad (D ν)?" *Journal of applied physics*, vol. 56, no. 4, pp. 1119–1122, 1984.

[70] G. Baum and H. Beneking, "Drift velocity saturation in MOS transistors," *IEEE Transactions on Electron Devices*, vol. 17, no. 6, pp. 481–482, 1970.

[71] R. Landauer, "Spatial variation of currents and fields due to localized scatterers in metallic conduction," *IBM Journal of Research and Development*, vol. 32, no. 3, pp. 306–316, 1988.

[72] Y. Imry, "Physics of Mesoscopic Systems: Directions in Condensed Matter Physics, edited by G. Grinstein and G. Mazenco," 1986.

[73] M. Lundstrom and J. Guo, *Nanoscale transistors: device physics, modeling and simulation*. Springer Science & Business Media, 2006.

[74] W. Kruppa and K. Sodomsky, "An Explicit Solution for the Scattering Parameters of a Linear Two-Port Measured with an Imperfect Test Set (Correspondence)," *IEEE Transactions on Microwave Theory and Techniques*, vol. 19, no. 1, pp. 122–123, 1971. doi: 10.1109/TMTT.1971.1127466

[75] A. Davidson, K. Jones, and E. Strid, "LRM and LRRM calibrations with automatic determination of load inductance," in *36th ARFTG conference digest*, vol. 18. IEEE, 1990, pp. 57–63.

[76] M. Koolen, J. Geelen, and M. Versleijen, "An improved de-embedding technique for on-wafer high-frequency characterization," in *Proc. Bipolar Circuits Technol. Meeting*. Minneapolis, MN, 1991, pp. 188–191.

[77] C. B. Zota, G. Roll, L.-E. Wernersson, and E. Lind, "Radio-frequency characterization of selectively regrown InGaAs lateral nanowire MOSFETs," *IEEE Transactions on Electron Devices*, vol. 61, no. 12, pp. 4078–4083, 2014.

[78] M. Hellenbrand, E. Lind, O.-P. Kilpi, and L.-E. Wernersson, "Effects of traps in the gate stack on the small-signal RF response of III-V nanowire MOSFETs," *Solid-State Electronics*, vol. 171, p. 107840, 2020.

[79] I. Kwon, M. Je, K. Lee, and H. Shin, "A simple and analytical parameter-extraction method of a microwave MOSFET," *IEEE Transactions on Microwave Theory and Techniques*, vol. 50, no. 6, pp. 1503–1509, 2002.

[80] Y.-L. Lai and K.-H. Hsu, "A new pinched-off cold-FET method to determine parasitic capacitances of FET equivalent circuits," *IEEE Transactions on Microwave Theory and Techniques*, vol. 49, no. 8, pp. 1410–1418, 2001. doi: 10.1109/22.939921

[81] W. Liu, "Fundamentals of III-V devices: HBTs, MESFETs, and HFETs/HEMTs," *(No Title)*, 1999.

[82] Y. Yuan, L. Wang, B. Yu, B. Shin, J. Ahn, P. C. McIntyre, P. M. Asbeck, M. J. Rodwell, and Y. Taur, "A Distributed Model for Border Traps in Al_2O_3-InGaAs MOS Devices," *IEEE Electron Device Letters*, vol. 32, no. 4, pp. 485–487, 2011.

[83] M. Isler and K. Schunemann, "Impact-ionization effects on the high-frequency behavior of HFETs," *IEEE transactions on microwave theory and techniques*, vol. 52, no. 3, pp. 858–863, 2004.

[84] M. Chertouk, H. Heiss, D. Xu, S. Kraus, W. Klein, G. Bohm, G. Trankle, and G. Weimann, "Metamorphic InAlAs/InGaAs HEMTs on GaAs substrates with a novel composite channels design," *IEEE Electron Device Letters*, vol. 17, no. 6, pp. 273–275, 1996. doi: 10.1109/55.496455

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

[86] G. C. DeSalvo, C. A. Bozada, J. L. Ebel, D. C. Look, J. P. Barrette, C. L. A. Cerny, R. W. Dettmer, J. K. Gillespie, C. K. Havasy, T. J. Jenkins *et al.*, "Wet Chemical Digital Etching of GaAs at Room Temperature," *Journal of The Electrochemical Society*, vol. 143, no. 11, p. 3652, nov 1996. doi: 10.1149/1.1837266. [Online]. Available: https://dx.doi.org/10.1149/1.1837266

[87] W. Lu, X. Zhao, D. Choi, S. El Kazzi, and J. A. del Alamo, "Alcohol-Based Digital Etch for III-V Vertical Nanowires With Sub-10 nm Diameter," *IEEE Electron Device Letters*, vol. 38, no. 5, pp. 548–551, 2017. doi: 10.1109/LED.2017.2690598

[88] J. Lin, X. Zhao, D. A. Antoniadis, and J. A. del Alamo, "A Novel Digital Etch Technique for Deeply Scaled III-V MOSFETs," *IEEE Electron Device Letters*, vol. 35, no. 4, pp. 440–442, 2014. doi: 10.1109/LED.2014.2305668

[89] E. Memisevic, J. Svensson, M. Hellenbrand, E. Lind, and L.-E. Wernersson, "Vertical InAs/GaAsSb/GaSb tunneling field-effect transistor on Si with s = 48 mv/decade and i_{ON} = 10 μa/μm for i_{OFF} = 1 na/μm at v_{DS} = 0.3 v," in *2016 IEEE International Electron Devices Meeting (IEDM)*, 2016. doi: 10.1109/IEDM.2016.7838450 pp. 19.1.1–19.1.4.

[90] P. Olausson, L. Södergren, M. Borg, and E. Lind, "Optimization of Near-Surface Quantum Well Processing," *physica status solidi (a)*, vol. 218, no. 7, p. 2000720, 2021. doi: https://doi.org/10.1002/pssa.202000720. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.1002/pssa.202000720

[91] É. O'Connor, S. Monaghan, K. Cherkaoui, I. M. Povey, and P. K. Hurley, "Analysis of the minority carrier response of n-type and p-type Au/Ni/Al_2O_3/$In_{0.53}Ga_{0.47}As$/InP capacitors following an optimized $(NH_4)_2S$ treatment," *Applied Physics Letters*, vol. 99, no. 21, p. 212901, 2011.

[92] H. Trinh, E. Y. Chang, P.-W. Wu, Y. Wong, C. Chang, Y. Hsieh, C. Yu, H. Nguyen, Y. Lin, K. Lin *et al.*, "The influences of surface treatment and gas annealing conditions on the inversion behaviors of the atomic-layer-deposition Al_2O_3/n-$In_{0.53}Ga_{0.47}As$ metal-oxide-semiconductor capacitor," *Applied Physics Letters*, vol. 97, no. 4, p. 042903, 2010.

[93] P. Rodriguez, L. Toselli, E. Ghegin, N. Chevalier, N. Rochat, E. Martinez, and F. Nemouchi, "In situ cleaning of InGaAs surfaces prior to low contact resistance metallization," *Microelectronic Engineering*, vol. 156, pp. 91–96, 2016.

[94] T. Hoshii, S. Lee, R. Suzuki, N. Taoka, M. Yokoyama, H. Yamada, M. Hata, T. Yasuda, M. Takenaka, and S. Takagi, "Reduction in interface state density of Al_2O_3/InGaAs metal-oxide-semiconductor interfaces by InGaAs surface nitridation," *Journal of Applied Physics*, vol. 112, no. 7, p. 073702, 2012.

[95] R. Suzuki, N. Taoka, M. Yokoyama, S.-H. Kim, T. Hoshii, T. Maeda, T. Yasuda, O. Ichikawa, N. Fukuhara, M. Hata *et al.*, "Impact of atomic layer deposition temperature on HfO_2/InGaAs metal-oxide-semiconductor interface properties," *Journal of Applied Physics*, vol. 112, no. 8, p. 084103, 2012.

[96] R. Suzuki, N. Taoka, M. Yokoyama, S. Lee, S. Kim, T. Hoshii, T. Yasuda, W. Jevasuwan, T. Maeda, O. Ichikawa *et al.*, "1-nm-capacitance-equivalent-thickness HfO_2/Al_2O_3/InGaAs metal-oxide-semiconductor structure with low interface trap density and low gate leakage current density," *Applied Physics Letters*, vol. 100, no. 13, p. 132906, 2012.

[97] O. Ceballos-Sanchez, E. Martinez, C. Guedj, M. Veillerot, and A. Herrera-Gomez, "Indium and gallium diffusion through zirconia in the TiN/ZrO_2/InGaAs stack," *Applied Physics Letters*, vol. 106, no. 22, p. 221602, 2015.

[98] J. Huang, N. Goel, H. Zhao, C. Kang, K. Min, G. Bersuker, S. Oktyabrsky, C. Gaspe, M. Santos, P. Majhi *et al.*, "InGaAs MOSFET performance and reliability improvement by simultaneous reduction of oxide and interface charge in ALD (La) AlO_x/ZrO_2 gate stack," in *2009 IEEE International Electron Devices Meeting (IEDM)*. IEEE, 2009, pp. 1–4.

[99] C.-Y. Chang, O. Ichikawa, T. Osada, M. Hata, H. Yamada, M. Takenaka, and S. Takagi, "Impact of La_2O_3 interfacial layers on InGaAs metal-oxide-semiconductor interface properties in Al_2O_3/La_2O_3/InGaAs gate stacks deposited by atomic-layer-deposition," *Journal of Applied Physics*, vol. 118, no. 8, p. 085309, 2015.

[100] T.-W. Kim, D.-H. Kim, D. Koh, H. Kwon, R. Baek, D. Veksler, C. Huffman, K. Matthews, S. Oktyabrsky, A. Greene *et al.*, "Sub-100 nm InGaAs quantum-well (QW) tri-gate MOSFETs with Al_2O_3/HfO_2 (EOT< 1 nm) for low-power logic applications," in *2013 IEEE International Electron Devices Meeting*. IEEE, 2013, pp. 16–3.

[101] O.-P. Kilpi, J. Svensson, E. Lind, and L.-E. Wernersson, "Electrical properties of vertical InAs/InGaAs heterostructure MOSFETs," *IEEE Journal of the Electron Devices Society*, vol. 7, pp. 70–75, 2018.

[102] J. Lin, X. Cai, Y. Wu, D. A. Antoniadis, and J. A. del Alamo, "Record Maximum Transconductance of 3.45 ms/ μm for III-V FETs," *IEEE Electron Device Letters*, vol. 37, no. 4, pp. 381–384, 2016. doi: 10.1109/LED.2016.2529653

[103] O.-P. Kilpi, M. Hellenbrand, J. Svensson, A. R. Persson, R. Wallenberg, E. Lind, and L.-E. Wernersson, "High-Performance Vertical III-V Nanowire MOSFETs on Si With $g_m > 3$ mS/μm," *IEEE Electron Device Letters*, vol. 41, no. 8, pp. 1161–1164, 2020. doi: 10.1109/LED.2020.3004716

[104] F. Lindelöw, N. S. Garigapati, L. Södergren, M. Borg, and E. Lind, "III-V nanowire MOSFETs with novel self-limiting $\wedge$-ridge spacers for RF applications," *Semiconductor Science and Technology*, vol. 35, no. 6, p. 065015, 2020.

[105] C. B. Zota, F. Lindelöw, L.-E. Wernersson, and E. Lind, "InGaAs tri-gate MOSFETs with record on-current," in *2016 IEEE International Electron Devices Meeting (IEDM)*. IEEE, 2016, pp. 3–2.

[106] K. Tan, R. Dia, D. Streit, T. Lin, T. Trinh, A. Han, P. Liu, P.-M. Chow, and H. Yen, "94-GHz 0.1-μm T-gate low-noise pseudomorphic InGaAs HEMTs," *IEEE Electron Device Letters*, vol. 11, no. 12, pp. 585–587, 1990. doi: 10.1109/55.63047

[107] M. Aust, H. Wang, M. Biedenbender, R. Lai, D. Streit, P. Liu, G. Dow, and B. Allen, "A 94-GHz monolithic balanced power amplifier using 0.1-μm gate GaAs-based HEMT MMIC production process technology," *IEEE microwave and guided wave letters*, vol. 5, no. 1, pp. 12–14, 1995.

[108] N. S. Garigapati, F. Lindelöw, L. Södergren, and E. Lind, "Capacitance Scaling in $In_{0.71}Ga_{0.29}As$/InP MOSFETs With Self-Aligned a:Si Spacers," *IEEE Transactions on Electron Devices*, vol. 68, no. 8, pp. 3762–3767, 2021. doi: 10.1109/TED.2021.3092299

[109] O. Stier, "Electronic and Optical Properties of Quantum Dots and Wires," 2005.

[110] I. Vurgaftman, J. á. Meyer, and L. R. Ram-Mohan, "Band parameters for III-V compound semiconductors and their alloys," *Journal of applied physics*, vol. 89, no. 11, pp. 5815–5875, 2001.

APPENDICES

A

$\mathbf{k} \cdot \mathbf{p}$ Theory

Parameter	$In_xGa_{1-x}As$	InP
E_g (eV)	$x\left(0.417 - 0.276 \times 10^{-3}\frac{T^2}{T+93}\right) + (1 - x)\left(1.519 - 0.5405 \times 10^{-3}\frac{T^2}{T+204}\right) - 0.477x(1-x)$	1.354
Δ_{SO} (eV)	$0.34 - 0.093x + 0.15x^2$	0.108
m^* (kg)	$(0.0667 - 0.0429x - 0.0091x(1-x))m_0$	$0.079m_0$
E_p (eV)	$17x + 23(1-x) - 4.97x(1-x)$	16
a_L (nm)	$x\left(0.60583 - 2.74 \times 10^{-6}(T-300)\right) + (1 - x)\left(0.565325 - 3.88 \times 10^{-6}(T-300)\right)$	0.58697
E_v' (eV)	$-\frac{\Delta_{SO}}{3} + 0.07x + 0.34$	-0.036
E_c (eV)	$E_g + 0.07x + 0.34$	1.354

Table A.1: $In_xGa_{1-x}As$ and InP material parameters expressions in terms of indium composition and temperature. All parameters are taken from [109, 110] except for E_p.

$In_xGa_{1-x}As$ and InP material parameters that are used in simulations are given in Table A.1 and $\mathbf{k} \cdot \mathbf{p}$ matrix elements are provided in Table A.2. The $N_- = M - \frac{\hbar^2}{2m_0}$ and $N_+ = N - N_-$ are calculated for the respective material. The $\mathbf{k} \cdot \mathbf{p}$ and strain matrices of the bulk semiconductor are given in equations (A.1) and (A.2), respectively.

Parameter	$In_xGa_{1-x}As$	InP
A_c (eV.m^2)	$\frac{\hbar^2}{2m_0}\left(\frac{1}{m^*}-\frac{E_p}{E_g}\left(\frac{E_g+\frac{2}{3}\Delta_{so}}{E_g+\Delta_{SO}}\right)\right)$	3.88×10^{-20}
P (eV.m)	$\sqrt{\frac{\hbar^2}{2m_0}E_p}$	7.8081×10^{-10}
γ_1	$\frac{1}{\left(\frac{(1-x)}{7.1}+\frac{x}{19.7}\right)}$	4.95
γ_2	$\frac{1}{\left(\frac{(1-x)}{2.02}+\frac{x}{8.4}\right)}$	1.65
γ_3	$\frac{1}{\left(\frac{(1-x)}{2.91}+\frac{x}{9.3}\right)}$	2.35
L (eV.m^2)	$\frac{P^2}{E_g}-\frac{\hbar^2}{2m_0}(1+\gamma_1+4\gamma_2)$	-2.39×10^{-20}
M (eV.m^2)	$-\frac{\hbar^2}{2m_0}(1+\gamma_1-2\gamma_2)$	-1.097×10^{-19}
N (eV.m^2)	$\frac{P^2}{E_g}-\frac{3\hbar^2}{m_0}\gamma_3$	-1.4254×10^{-20}
a_c (eV)	-8.013 + 2.933x - 2.61$x(1-x)$	-5.35
b_v (eV)	-1.824 + 0.024x	-2
a_g (eV)	-8.233 + 2.153x	-6.35
d_v (eV)	-5.062 + 1.462x	-4.2
l (eV)	$2b_v+a_c-a_g$	-3
m (eV)	$a_c-a_g-b_v$	3
n (eV)	$\sqrt{3}d_v$	-7.2746
C_{11} (GPa)	118.8-8.5x	101.1
C_{12} (GPa)	53.8-8.5x	56.1
C_{44} (GPa)	59.4-19.8x	45.6

Table A.2: The Hamiltonian parameters include Kane, Luttinger, and strain elements of $In_xGa_{1-x}As$ and InP expressions are adapted from [38, 109].

$$H_{kp,4} = \begin{bmatrix} E_c + A_c\left(k_x^2 + k_y^2 + k_z^2\right) & iPk_x & iPk_y & iPk_z \\ -iPk_x & E_v^{'} + Lk_x^2 + M\left(k_y^2 + k_z^2\right) & Nk_xk_y & Nk_xk_z \\ -iPk_y & Nk_yk_x & E_v^{'} + Lk_y^2 + M\left(k_x^2 + k_z^2\right) & Nk_yk_z \\ -iPk_z & Nk_zk_x & Nk_zk_y & E_v^{'} + Lk_z^2 + M\left(k_x^2 + k_y^2\right) \end{bmatrix} \tag{A.1}$$

$$H_{st,4} = \begin{bmatrix} a_c\left(\epsilon_{11} + \epsilon_{22} + \epsilon_{33}\right) & -iP\sum_j \epsilon_{xj}k_j & -iP\sum_j \epsilon_{yj}k_j & -iP\sum_j \epsilon_{zj}k_j \\ iP\sum_j \epsilon_{xj}k_j & l\epsilon_{11} + m\left(\epsilon_{22} + \epsilon_{33}\right) & n\epsilon_{12} & n\epsilon_{13} \\ iP\sum_j \epsilon_{yj}k_j & n\epsilon_{21} & l\epsilon_{22} + m\left(\epsilon_{11} + \epsilon_{33}\right) & n\epsilon_{23} \\ iP\sum_j \epsilon_{zj}k_j & n\epsilon_{31} & n\epsilon_{32} & l\epsilon_{33} + m\left(\epsilon_{11} + \epsilon_{22}\right) \end{bmatrix} \tag{A.2}$$

The $\mathbf{k}\cdot\mathbf{p}$ and strain matrices for 2D quantum well, assuming it is confined in x-direction are given in equations (A.3) and (A.4), respectively.

$$H_{kp,4}^{'} = \begin{bmatrix} E_c - \partial_x A_c \partial_x + A_c\left(k_y^2 + k_z^2\right) & P\partial_x & iPk_y & iPk_z \\ -P\partial_x & E_v^{'} - \partial_x L\partial_x + M\left(k_y^2 + k_z^2\right) & -i\partial_x N_+ k_y - ik_y N_- \partial_x & -i\partial_x N_+ k_z - ik_z N_- \partial_x \\ -iPk_y & i\partial_x N_+ k_y + ik_y N_- \partial_x & E_v^{'} - \partial_x M\partial_x + \left(Lk_y^2 + Mk_z^2\right) & Nk_yk_z \\ -iPk_z & i\partial_x N_+ k_z + ik_z N_- \partial_x & Nk_zk_y & E_v^{'} - \partial_x M\partial_x + \left(Mk_y^2 + Lk_z^2\right) \end{bmatrix} \tag{A.3}$$

and

$$H'_{st,A} = \begin{bmatrix} a_c\left(\epsilon_{11}+\epsilon_{22}+\epsilon_{33}\right) & -P\epsilon_{11}\partial_x & -iP\epsilon_{22}k_y & -iP\epsilon_{33}k_z \\ P\epsilon_{11}\partial_x & l\epsilon_{11}+m\left(\epsilon_{22}+\epsilon_{33}\right) & 0 & 0 \\ iP\epsilon_{22}k_y & 0 & l\epsilon_{22}+m\left(\epsilon_{11}+\epsilon_{33}\right) & 0 \\ iP\epsilon_{33}k_z & 0 & 0 & l\epsilon_{33}+m\left(\epsilon_{11}+\epsilon_{22}\right) \end{bmatrix} \tag{A.4}$$

In the similar way, $\mathbf{k}\cdot\mathbf{p}$ and strain matrices for 1D nanowire,which is confined in both x and y direction are given here:

$$H'_{kp,A} = \begin{pmatrix} E_c-\partial_x A_c\partial_x-\partial_y A_c\partial_y+A_c k_z^2 & P\partial_x & P\partial_y & iPk_z \\ -P\partial_x & E'_v-\partial_x L\partial_x-\partial_y M\partial_y+Mk_z^2 & -\partial_x N_+\partial_y-\partial_y N_-\partial_x & -i\partial_x N_+k_z-ik_z N_-\partial_x \\ -P\partial_y & \partial_x N_+\partial_y+\partial_y N_-\partial_x & E'_v-\partial_x M\partial_x-\partial_y L\partial_y+Mk_z^2 & -i\partial_y N_+k_z-ik_z N_-\partial_y \\ -iPk_z & i\partial_x N_+k_z+ik_z N_-\partial_x & i\partial_y N_+k_z+ik_z N_-\partial_y & E'_v-\partial_x M\partial_x-\partial_y M\partial_y+Lk_z^2 \end{pmatrix} \tag{A.5}$$

and

$$H'_{st,A} = \begin{bmatrix} a_c\left(\epsilon_{11}+\epsilon_{22}+\epsilon_{33}\right) & -P\epsilon_{11}\partial_x-P\epsilon_{12}\partial_y & -P\epsilon_{12}\partial_x-P\epsilon_{22}\partial_y & -iP\epsilon_{33}k_z \\ P\epsilon_{11}\partial_x+P\epsilon_{12}\partial_y & l\epsilon_{11}+m\left(\epsilon_{22}+\epsilon_{33}\right) & n\epsilon_{12} & 0 \\ P\epsilon_{12}\partial_x+P\epsilon_{22}\partial_y & n\epsilon_{12} & l\epsilon_{22}+m\left(\epsilon_{11}+\epsilon_{33}\right) & 0 \\ iP\epsilon_{33}k_z & 0 & 0 & l\epsilon_{33}+m\left(\epsilon_{11}+\epsilon_{22}\right) \end{bmatrix} \tag{A.6}$$

B

Fabrication Details

The fabrication process of InGaAs composite channel quantum well MOSFETs is discussed in detail in this chapter.

SELF-ALIGNED INGAAS CC QUANTUM WELL MOSFET FABRICATION PROCESS

1. **n^+ Etching or L_{sd} Definition**
   - **Sample Preparation**
     Chip 10 ×12 mm is prepared from 4″ MBE wafer
     Organic cleaning: 5 min in hot acetone (65°C), 2 min in Isopropyl alcohol (IPA)
     Ozone cleaning: 8 min in ozone cleaner at 500 sccm O_2 flow and at room temperature.
   - **Al_2O_3/SiN Mask Preparation**
     Mask Deposition
     Deposition of thin approx. 5 nm (50 cycles) Al_2O_3 using ALD at 100°C. TMAl and H_2O are used as precursor gases.
     40 nm (*approx. 4 cyc*) thick SiN deposition using PECVD at 250°C and SiH_4, N_2 are precursor gases.
     Resist Deposition
     Demoisturizing the sample at 200°C for 5 min.
     Spin coat ARP 6200.09 at 4000 rpm for 55 s with acceleration of 2500 rpms. The approximate resist thickness is around 200 nm.
     Post baking the resist for 1 min at 160°C

EBL 1: GRE-I Lithography
EBL (50 kV), area exposure with a base dose of 125 μC/cm^2, 40 μm aperture size, and 6 nm step size.
Resist Development
1 min in Amyl acetate, 1 min rinsing in DIW.
Resist ashing in oxygen plasma for 30 s. *This step removes the undeveloped resist particles in the openings. Note that this ashing will also thin down the resist thickness.*

- **Etching of Al_2O_3/SiN Stack**
  SiN etch: (CHF_3:O_2) (50/5 sccm), RF power: 75 W, 20 mT, at T = 20°C for 80 s. *The etch rate is approx 32 nm/min.*
  Alumina etch: HF (1:400) for 15 s, 1 min in DIW and N_2 blow dry.
  Resist removal in oxygen plasma for 5 min.
- **Etching of highly doped contacts**
  Freshly prepared H_3PO_4:H_2O_2:H_2O (1:1:25) for 45 s. *Using a thin alumina etch stop layer is used to minimize the undercut in n^+ cap layer.*
- **Al_2O_3/SiN Mask Removal**
  HF (48%) for 2 min, 1 min rinsing in DIW and N_2 blow dry. *This is a sacrificial mask used to etch n^+. Another option could be to use dry etch SiN and wet etching of Alumina.*
- **1-Digital etch**
  8 min in Ozone cleaner at room temperature, 500 sccm O_2 flow, 15 s etch in HCl (1:10), 30 s in DIW and N_2 blow dry. *This additional digital etch removes any residues in the active area.*

2. **Contacts Lithography**

- **Resist mask preparation**
  Resist deposition
  Demoisturizing the sample at 200°C for 5 min.
  Spin coat the sample with ma-N 1420 at 6000 rpm for 45 s with acceleration of 1500 rpms. The approximate resist thickness is 1.4 μm.
  Post baking the resist at 95°C for 3 min.
  Photolithography
  Laser source with a wavelength of 375 nm. The base dose is 700 mJ/cm^2, and the defocus is 0.
  Resist Development
  ma-D 533/S for 70 s, 1 min DIW and N_2 blow dry.
  Resist ashing in oxygen plasma for 1 min.

- **Metal Evaporation and Lift-off**
  Ti/Pd/Au (5/5/40 nm) metal evaporation using e-beam evaporator.
  Lift-off in hot acetone 65°C for 30 min, 30 s in IPA and N_2 blow dry.
  Resist ashing using oxygen plasma for 1 min.

3. **Device Isolation**
   - **MESA Etch Mask Preparation**
     Resist Deposition
     Demoisturizing the sample at 180°C for 5 min.
     Spin coat S1805 at 4000 rpm for 60 s with acceleration of 1000 rpms.
     Post baking at 115°C for 90 s.
     Photolithography
     Laser source with a wavelength of 405 nm. The base area dose is 220 mJ/cm^2, and the defocus is 0.
     Resist development
     MF 319 for 90 s, 1 min DIW and N_2 blow dry.
     Resist ashing in oxygen plasma for 1 min.
     Baking the developed resist mask at 120°C for 10 min. *This baking helps to improve the resist adhesion to the underneath layer, thereby decreasing the horizontal etch rate.*
   - **MESA Etching**
     n^+ cap layer: H_3PO_4:H_2O_2:H_2O (1:1:25) for 45 s or in Citric acid:H_2O_2 (1:1) for 55 s.
     n-InP barrier layer: HCl (1:1) for 4 to 5 s, *InP etch rate is approximately 4 nm/s.*
     InGaAs CC channel: Citric acid:H_2O_2 (1:1) for 14 s, *This particular acid solution has etch selectivity between InGaAs and InAlAs.*
   - **MESA Etch Mask Removal**
     5 min in acetone, 30 s IPA, followed by resist ashing in oxygen plasma for 1 min.
4. **Gate Recess Etch Step - II**
   - **SiN Mask Preparation**
     Mask Deposition
     Approxmately 40 nm (*approx. 4 cyc*) thick SiN deposition using PECVD at 250°C and SiH_4, N_2 are precursor gases.

Resist Deposition

Demoisturizing the sample at 200°C for 5 min.

Spin coat ARP 6200.09 at 4000 rpm for 55 s with acceleration of 2500 rpms.

Post baking of the resist for 1 min at 160°C.

EBL 2: GRE-II Lithography

EBL (50 kV) with an area base dose of 125 μC/cm^2, 40 μm aperture size and 6 nm step size.

Resist Development

1 min in amyl acetate, 1 min rinsing in DIW, and N_2 blow dry.

Resist ashing in oxygen plasma for 30 s.

Gate Length Definition

SiN etch using (CHF_3:O_2) (50/5 sccm), 20 mT, RF (75 W) at T = 20°C for 80 s.

- **InP recess etch**

  3 digital etches are performed to remove the 7 nm thick n-InP layer in the channel region. One digital etch cycle consists of 8 min in ozone cleaner at O_2 500 sccm at room temperature to create a thin oxide layer and its selective etching in HCl (1:10) for 15 s. The etch rate is around 2 nm/cycle.

  Resist Removal

  5 to 7 min of resist ashing using oxygen plasma.

5. **Passivation and High-K Gate Oxide Deposition**
   - Sulfur passivation in Ammonium sulfide:H_2O (1:1) for 20 min, 20 s DIW.
   - Al_2O_3/HfO_2 (5/37 cycles) bilayer oxide deposition using ALD. The approximate EOT is 1 nm.
6. **Gate Contact Lithography and Metallization**
   - **Resist Mask Preparation**

     Resist Deposition

     Demoisturizing the sample at 180°C for 5 min.

     Spin coat PMMA A6 at 4500 rpm for 45 s with acceleration of 1500 rpms. The approximate thickness of the resist is close to 450 nm.

     Post baking the resist at 180°C for 2 min.

     EBL 2:Gate Contact Lithography

     EBL (50 kV) with the area base dose of 835 μC/cm^2 and 40 μm aperture size and 6 nm step size.

Resist development
90 s in MIBK:IPA (1:3), 30 s in IPA and N_2 blow dry.
Resist ashing in oxygen plasma for 45 s.

- **Metal Evaporation and Lift-off**
  Ti/Pd/Au (2/5/200 nm) metal stack evaporation using e-beam evaporator.
  Lift-off in hot acetone at 65°C for 30 min.
  Resist ashing using oxygen plasma for 1 min. *This oxygen plasma ashing removed the extra resist residues post lift-off.*

7. **Etching of SiN on contacts**

   - **Resist mask preparation**
     Resist Deposition
     Demoisturizing the sample at 180°C for 5 min.
     Spin coat S1805 at 4000 rpm for 60 s with acceleration of 1000 rpms.
     Post baking at 115°C for 90 s.
     Photolithography
     Laser source with a wavelength of 405 nm. The base area dose is 220 mJ/cm^2, and the defocus is 0.
     Resist development
     MF 319 for 90 s, 1 min DIW and N_2 blow dry.
     Resist ashing in oxygen plasma for 1 min.
     Baking the developed resist mask at 120°C for 10 min. *This baking helps to improve the resist adhesion to the underneath layer, thereby decreasing the horizontal etch rate.*
   - **Etching of gate oxide and SiN**
     Bi-layer gate oxide etch in BOE(10:1) for 2 min 30 s, 1 min in DIW and N_2 blow dry.
     SiN etch: (CHF_3:O_2) (50/5 sccm), RF power: 75 W, 20 mT, at T = 20°C for 80 s.
   - **Resist Removal**
     5 to 7 min of resist ashing using oxygen plasma.

8. **Pads lithography**
   - **Resist mask preparation**
     Resist deposition
     Demoisturizing the sample at 200°C for 5 min.
     Spin coat the sample with ma-N 440 at 6000 rpm for 45 s with acceleration of 1500 rpms. The approximate resist thickness is 4.4 μm.

Post baking the resist at 95°C for 3 min.

Photolithography

Laser source with a wavelength of 375 nm. The area base dose is 2000 mJ/cm^2, and a defocus is 0.

Resist Development

ma-D 532/S for 1 min 45 s, 1 min DIW and N_2 blow dry.

Resist ashing using oxygen plasma for 1 min.

- **Metal Evaporation and Lift-off**

  Ti/Au (5/130 nm) metal evaporation using e-beam evaporator.

  Lift-off using hot acetone 65°C for 30 min.

  Resist ashing using oxygen plasma for 1 min.