# Novel III-V Compound Semiconductor Technologies for Low Power Digital Logic Applications

David Alan John Millar

Submitted in fulfilment of the requirements for the
Degree of Doctor of Philosophy

School of Engineering
College of Science and Engineering
University of Glasgow


June 2018

# Abstract

As silicon (Si) complementary metal oxide semiconductor (CMOS) technology continues to scale into the 10 nm node, chip power consumption is approaching 200 W/cm$^2$ and any further increase is unsustainable. Incorporating III-V compound semiconductor n-type devices into future CMOS generations could allow for the the reduction in supply voltage, and therefore, power consumption, while simultaneously improving on-state performance. The advanced state of Si CMOS places stringent demands on III-V devices, however: the current 14 nm Si tri-gate devices employ high aspect ratio, densely spaced fins which serve to significantly increase current per chip surface area.

III-V devices need to significantly out perform state of the art Si devices in order to merit their disruptive incorporation into the well established CMOS process. This necessitates that they too exploit the vertical dimension. To this end, this thesis reports on the fabrication, measurement and analysis of high aspect ratio junctionless InGaAs FinFETs.

The junctionless architecture was first demonstrated in 2010 and was shown to circumvent prohibitive fabrication challenges for devices with ultra short gate lengths. This work investigated the impact of fin width on both the on and off-state performance of 200 nm gate length devices, with nominal fin widths of 10, 15 and 20 nm.

Excellent subthreshold performance was demonstrated, with the narrowest fin width exhibiting a minimum subthreshold swing (SS) of 73 mV/Dec, and an average SS of 80 mV/Dec. over two decades of current. A maximum on-current, $I_{on}$, of 80.51 $\mu$A/cm$^2$ was measured at a gate overdrive of 0.5 V from an off-state current, $I_{off}$, of 100 nA/cm$^2$ and a drain voltage, $V_d$, of 0.5 V, with current normalised by gated perimeter. This is competitive with other III-V junctionless devices at similar gate lengths. With current normalised to base fin width, however, $I_{on}$ increases to 371.8 $\mu$A/cm$^2$, which is a record value among equivalently normalised non-planar III-V junctionless devices at any gate length. This technology, therefore, clearly demonstrates the feasibility of incorporating scaled, etched InGaAs fins into future logic generations.

Perhaps the greatest bottleneck to the incorporation of III-V compounds into future CMOS technology nodes, however, is the lack of a suitable III-V PMOS candidate: co-integrating different

material systems onto a common substate incurs great fabrication complexity, and therefore, cost. III-V antimonides, however, have recently emerged as promising candidates for III-V PMOS and exhibit the highest bulk electron mobility of all III-Vs in addition to a hole mobility second only to germanium.

InGaSb ternary compounds have been shown to offer the best combined performance for electrons and holes in the same material, and as such, have the potential to the enable the most simplistic incarnation of III-V CMOS; provided, of course, that is possible to form a gate stack to both device polarities with sufficient electrical properties. To date, however, there has been no investigation into the high-k dielectric interface to InGaSb. To this end, this thesis presents results of the first investigation into the impact of in-situ $H_2$ plasma exposure on the electrical properties of the p/n-$In_{0.3}Ga_{0.7}Sb$-$Al_2O_3$ interface.

The parameter space was explored systematically in terms of $H_2$ plasma power and exposure time, and further, the impact of impact of in-situ trimethylaluminium (TMA) pre-cleaning and annealing in forming gas was assessed. Metal oxide semiconductor capacitors (MOSCAPs) were fabricated subsequent to $H_2$ plasma processing and $Al_2O_3$ deposition, and the corresponding capacitance-voltage and conductance-voltage measurements were analysed both qualitatively and quantitatively via the simulation of an equivalent circuit model.

X-Ray photoelectron spectroscopy (XPS) analysis of samples processed as part of the plasma power series revealed a combination of ex-situ HCl cleaning and in-situ $H_2$ plasma exposure to completely remove In and Sb sub oxides, with the Ga-O content reduced to Ga-O:InGaSb <0.1. The optimal process, which included ex-situ HCl surface cleaning, in-situ $H_2$ plasma and TMA pre-cleaning, and a post gate metal forming gas anneal, was unequivocally demonstrated to yield a fully unpinnned MOS interface with both n and p-type MOSCAPs explicitly demonstrating a genuine minority carrier response. Interface state and border trap densities were extracted, with a minimum $D_{\text{it}}$ of $1.73\times\ 10^{12}$ $\text{cm}^{-2}$ $\text{eV}^{-1}$ located at $\sim$110 meV below the conduction band edge and peak border trap densities approximately aligned with the valence and conduction band edges of $3 \times 10^{19}$ $\text{cm}^{-3}$ $\text{eV}^{-1}$ and $6.5 \times 10^{19}$ $\text{cm}^{-3}$ $\text{eV}^{-1}$ respectively.

These results indicate that the optimal gate stack process is indeed applicable to both p and n-type InGaSb MOSFETs, and therefore, represent a critical advancement towards achieving high performance III-V CMOS.

# Contents

# Acknowledgements

While the last four years have undoubtedly been some of the most tiring, stressful and frustrating of my life, they have equally been some of the most stimulating, intellectually challenging and rewarding. Completing my PhD is an accomplishment of which I am extremely proud and one which I could not possibly have achieved without the help and support of several others.

First and foremost, I would like to give my sincere gratitude to my supervisor, Professor Iain Thayne, who often went above and beyond the call of duty, offered advice and assistance whenever needed and also afforded me the freedom to pursue areas of research which particularly grasped my interest.

My thanks also to all past and present members of our research group, narrow bandgap-ers and wide bandgap-ers alike. In particular, I would like to thank Dr Sankar Peralagu, Dr Olesya Ignatova, Dr Xu Li and Dr Yen-Chun Fu for training me on the multitude of different nanofabrication tools and measurement equipment, and for many a fruitful discussion. I would like to thank Dr Matthew Steer for growing all of the material used in this work which directly enabled the results achieved. My thanks also to my research collaborators at Tyndall National Institute, without whom much of this work would not have come to fruition. In particular, my thanks to Dr Ian Povey for depositing the gate dielectric on my FinFETs and to Professor Paul Hurley and Dr Karim Cherkaoui for their input and assistance with CV measurements on antimonide devices.

Beyond my immediate research group and collaborators, there are a variety of other colleagues and friends that I am hugely appreciative to. I would like to thank Dr Muhammad Mirza who offered extensive assistance with very specific issues associated with fabricating ultrascaled devices, including training me on the particularly awkward to use S900 microscope. My thanks to all the staff at the James Watt Nanofabrication Centre who keep the lab running like a well oiled machine and continue to provide the University of Glasgow with an unrivalled research facility. Thank you to Dr Scott Watson and Dr Kevin Gallagher who helped break up the monotony of days spent in the clean room, and to all my other friends for keeping me sane during a somewhat stressful time, in particular: Dr Robert White Dr, the soon to be Dr Konstantinos Floros, Gordon Nelson, Callum Malaney, Jonathan Baylis and Michael Idebolo.

A massive thank you to my brother, (Dr.) Ross (Millar), for the ever enlightening discussions on all things semicondcutor and for proof reading multiple chapters of this work at extremely short notice, and to my parents for having supported me not only throughout the particularly challenging period of the last 4 years, but for always having supported me in any endeavour.

Lastly, special thanks to my beautiful girlfriend, Dr Fiona Catterall (not PhD, merely medical), who has provided me with invaluable moral support, without which I would not have been able to achieve what I have.

# List of Tables

# List of Figures

# Acronyms

**AFM** atomic force microscopy

**ALD** atomic layer deposition

**APD** anti-phase domains

**ART** aspect ratio trapping

**BEF** beam error feedback

**BF** bright field

**BOX** buried oxide layer

**BSS** beam step size

**BTBT** band to band tunneling

**BTI** bias temperature instability

**CD** critical dimension

**CELO** confined epitaxial lateral overgrowth

**CMOS** complementary metal oxide semiconductor

**CV** capacitance-voltage

**CVD** chemical vapour deposition

**DAC** digital to analogue converter

**DC** direct current

**DE** digital etch(ing)

**DFT** density functional theory

**DIBL** drain induced barrier lowering

**DOS** density of states

**DWB** direct wafer bonding

**EBL** electron-beam lithography

**EDX** energy dispersive X-ray spectroscopy

**EOT** equivalent oxide thickness

**EWF** extra wide field

**FGA** forming gas anneal

**FIB** focussed ion beam

**FTIR** Fourier-transform infrared spectroscopy

**GAA** gate all around

**HEMT** high electron mobility transistor

**HH** heavy hole

**HSQ** hydrogen silsesquioxane

**IC** integrated circuit

**ICP** inductively coupled plasma

**ICP-RIE** inductively coupled plasma reactive ion etching

**IL** interfacial dipole layer

**IM** inversion mode

**ITRS** international technology roadmap for semiconductors

**JWNC** James Watt Nanofabrication centre

**LER** line edge roughness

**MBE** molecular beam epitaxy

**MOCVD** metal organic chemical vapour deposition

**MOS** metal oxide semiconductor

**MOSCAP** metal oxide semiconductor capacitor

**MOSFET** metal oxide semiconductor field effect transistor

**NMOS** n-channel metal oxide semiconductor

**PBE** parasitic bi-polar effect

**PEB** post exposure bake

**PL** photoluminescence

**PMMA** poly(methyl methacrylate)

**PMOS** p-channel metal oxide semiconductor

**PSF** point spread function

**RDF** random dopant fluctuations

**RIE** reactive ion etching

**RMS** root mean square

**RTA** rapid thermal anneal(er)

**RTN** random telegraph noise

**S/D** source/drain

**SB** Schottky barrier

**SCCM** standard cubic centimetres per minute

**SCEs** short channel effects

**SI** semi-insulating

**SOI** silicon on insulator

**SS** subthreshold swing

**TDD** threading dislocation densities

**TEM** transmission electron microscope

**TEMAHf** tetrakis-ethylmethylaminohafnium

**TLM** transmission line method

**TMA** Trimethylaluminium

**TNI** Tyndall National Institute

**UHR** ultra high resolution

**UTB** ultra thin body

**VLSI** very-large-scale integration

**VRU** variable resolution unit

**XOI** III-V on insulator

**XPS** X-ray photoelectron spectroscopy

# List of symbols

$\alpha_{\Gamma}$ Non-parabolicity factor

$C_{\mathbf{d}}$ Depletion capacitance

$C_{\mathbf{D,min}}$ Minimum depletion capacitance

$C_{\mathbf{eff}}$ Effective switching capacitance

$C_{\mathbf{g}}$ Gate capacitance

$C_{\mathbf{HF,a}}$ Measured capacitance at high frequency in accumulation

$C_{\mathbf{inv}}$ Inversion capacitance

$C_{\mathbf{LF,a}}$ Measured capacitance at low frequency in accumulation

$C_{\mathbf{m}}$ Measured capacitance

$C_{\mathbf{max,HF}}$ Maximum measured capacitance at high frequency

$C_{\mathbf{min,HF}}$ Minmimum measured capacitance at low frequency

$C_{\mathbf{min,T}}$ Total capacitance of an MOS structure at maximum depletion width

$C_{\mathbf{mod}}$ Capacitance modulation

$c_{\mathbf{n}}$ Capture probability for electrons

$C_{\mathbf{ox}}$ Oxide capacitance

$c_{\mathbf{p}}$ Capture probability for holes

$C_{\mathbf{s}}$ Semiconductor capacitance

$d_{\mathbf{f}}$ Electron beam diameter

$D_{\mathbf{it}}$ Density of interface traps

$D_{\mathbf{it}}^{+}$ Density of positively charged interface traps

$D_{\mathbf{it}}^{0}$ Density of electrically neutral interface traps

$\mathscr{E}$ Electric field

$E$ Energy

$E_{\mathbf{A}}$ Activation energy

$E_{\mathbf{F}}$ Fermi level

$E_{\mathbf{g}}$ Band gap

$E_{\mathbf{g},\Gamma}$ Bandgap energy between the HH maxima and $\Gamma$ valley minima

$m_{\Gamma}^{*}$ Effective mass of the $\Gamma$ valley

$E_{min,i}$ Energy of the conduction minima for the $i$th valley

$e_{\mathbf{n}}$ Emission probability for electrons

$e_{\mathbf{p}}$ Emission probability for holes

$\varepsilon_{\mathbf{r}}$ Relative permitivity

$E_{\mathbf{s}}$ Electric field at the semiconductor surface

$E_{\mathbf{T}}$ Trap energy

$E_{\mathbf{v}}$ Valence band minimum

$f$ Switching frequency

$F$ Fermi-Dirac distribution function

$F_{\mathbf{T}}$ Fermi-Dirac distribution as a function of trap energy

$g_i$ Density of states of the $i$th valley, where $i = \Gamma$L,X,HH

$G_{\mathbf{I}}$ Equivalent parallel conductance in inversion

$g_{\mathbf{m}}$ Transconductance

$G_{\mathbf{m}}$ Measured conductance

$g_{\mathbf{m,sat}}$ Peak transconductance in saturation

$h$ Planck's constant

$I_{\mathbf{d}}$ Drain current

$I_{\mathbf{off}}$ Off-current

$I_{\mathbf{on}}$ On-current

$\kappa$ Dennard scaling factor

$L_{\mathbf{g}}$ Gate length

$L_{\mathbf{perim}}$ gated perimeter length

$L_{\mathbf{s}}$ Parasitic series inductance

$m_0$ Electron rest mass

$N$ Number of decades

$n$ Electron concentration

$N_{\mathbf{A}}$ Acceptor doping density

$N_{\mathbf{BT}}$ Density of border traps per volume per energy

$N_{\mathbf{D}}$ Donor doping density

$n_{\mathbf{i}}$ Intrinsic carrier concentration

$n_{\mathbf{s}}$ Electron density at the surface

$\rho_{\mathbf{c}}$ Specific contact resitivity

$P_{\mathbf{D}}$ Dynamic power dissipation

$P_{\mathbf{S}}$ Static power dissipation

$p_{\mathbf{s}}$ Hole density at the surface

$q$ electronic charge

$Q_{\mathbf{d}}$ Depletion charge

$Q_{\mathbf{inv}}$ Inversion charge

$Q_{\mathbf{s}}$ Semiconductor charge

$R_{\mathbf{ch}}$ Channel resistance

$r_{\mathbf{c,n}}$ Capture rate for electrons

$r_{\mathbf{c,p}}$ Capture rate for holes

$R_{\mathbf{f}}$ Resist thickness

$r_{\mathbf{j}}$ Junction depth

$R_{\mathbf{on}}$ On-resistance

$r_{\mathbf{q}}$ RMS surface roughness

$R_{\mathbf{s}}$ MOSCAP series resistance

$\sigma$ Conductivity

$R_{\mathbf{sd}}$ Source/drain resistance

$\sigma_{\mathbf{n}}$ Electron capture cross section

$SS_{\mathbf{sat}}$ Minimum subthreshold swing in saturation

$\tau$ CV/I gate delay

$T$ Temperature

$t_{\mathbf{ch}}$ Channel thickness

$t_{\mathbf{ox}}$ Oxide thickness

$\mu$ Bulk mobility

$\mu_{\mathbf{n}}$ Bulk electron mobility

$\mu_{\mathbf{p}}$ Bulk hole mobility

$V_{\mathbf{b}}$ beam voltage

$V_{\mathbf{d}}$ Drain voltage

$V_{\mathbf{DD}}$ Supply Voltage

$V_{\mathbf{g}}$ Gate voltage

$v_{\mathbf{inj}}$ Injection Velocity

$V_{\mathbf{t}}$ Threshold voltage

$v_{\mathbf{th,n}}$ Thermal velocity of electrons

$W_{\mathbf{D}}$ Depletion width

$W_{\mathbf{D,max}}$ Maximum depletion width

$W_{\mathbf{fin}}$ beam voltage

$W_{\mathbf{g}}$ Gate width

$x$ Depth into the semiconductor, referenced from the sample surface

$\psi$ Electrostatic potential

$\psi_{\mathbf{s}}$ Surface potential

$\omega$ Angular frequency

# Associated Publications

- D. A. J. Millar, U. Peralagu, X. Li, M. J. Steer,Y.-C. Fu, P. K. Hurley and I. G. Thayne. Demonstration of Genuine Surface Inversion for the p/n $In_{0.3}Ga_{0.7}Sb$-$Al_2O_3$ MOS System with in-situ $H_2$ Plasma Cleaning. Submitted to Applied Physics Letters.

- D. A. J. Millar, X. Li, U. Peralagu, M. J. Steer, I. M. Povey, G. Gaspar, M. Schmidt, P. K. Hurley and I. G. Thayne, "High Aspect Ratio Junctionless InGaAs FinFETs Fabricated Using a Top-Down Approach" in: *76th Device Research Conference (DRC)*, 2018, pp. 1-2.

- D. A. J. Millar, S. Supardan, U. Peralagu, M. Sousa, X. Li, V.R. Dhanak, Y.C. Fu, M. Steer, H. Schmid, I.Z. Mitrovic and I.G. Thayne, "Electrical and Chemical Analysis of the in-situ $H_2$ Plasma Cleaned InGaSb-$Al_2O_3$ Interface" in: *48th IEEE Semiconductor Interface Specialists Conference (SISC)*, 2017.

- D. A. J. Millar, U. Peralagu, X. Li, Y.-C. Fu, G. Gaspar, P. K. Hurley and I. G. Thayne, "Improving the electrical properties of the $In_{0.3}Ga_{0.7}Sb$-$Al_2O_3$ interface via in-situ $H_2$ plasma and TMA exposure" in: *20th Conference on Insulating Films on Semiconductors (INFOS)*, 2017.

- D. A. J. Millar, U. Peralagu, Y.-C.Fu, X. Li, M. J. Steer and I. G. Thayne, "Initial Investigation on the Impact of In-Situ Hydrogen Plasma Exposure to the Interface Between Molecular Beam Epitaxially Grown p-$Ga_{0.7}In_{0.3}Sb$ (100) and Thermal Atomic Layer Deposited (ALD) $Al_2O_3$" in: *19th Workshop on Dielectrics in Microelectronics (WoDIM)*, 2016.

# Declaration

With the exception of chapters 1, 2 and 3, which contain introductory material, all work in this thesis was carried out by the author unless otherwise explicitly stated.

# Chapter 1

# Introduction

## 1.1 A Brief History of Si CMOS Scaling

Since the inception of the integrated circuit (IC) in the late 1950's [1], the number of transistors per IC has approximately doubled every two years: increasing from 1 in 1959 [1], to 2,300 in 1971 (the first microprocessor) [2], to over 10 billion in the present day [3]. This has been responsible for the exponential increase in computational power over the last 50 years as well as the accompanying social-economic benefits afforded by the microelectronic revolution. Such an exponential increase in transistor count was first predicted by Gordon Moore in 1965 based only on the limited data set of the preceding six years [4]. While at the time this was merely an empirical observation, the sustained validity of this growth rate, now known as 'Moore's law', lies in the economics of semiconductor manufacture [5]. Increasing the number of transistors per IC increases chip functionality, which in turn improves the performance to cost ratio for the consumer and induces exponential growth of the semiconductor market. This in turn promotes investment which funds the research required to develop the technology for the next generation, and the cycle repeats [6]. A critical enabler in perpetuating this cycle has been the progressive miniaturisation of the individual transistors which comprise modern digital logic circuits: the metal oxide semiconductor field effect transistor (MOSFET). In 1974, Dennard *et al*. published a seminal paper demonstrating how reducing the dimensions of MOSFETs could not only yield increased transistor densities but simultaneously improve their switching speed and power dissipation provided that certain design rules were followed [7]. The premise of Dennard scaling is to maintain the internal electric field, $\mathscr{E}$, of long-channel devices by reducing the supply voltage, $V_{\mathrm{DD}}$, and all device dimensions by the same scaling factor, $\kappa$. This includes the gate length, $L_{\mathrm{g}}$, the gate width, $W_{\mathrm{g}}$, the source/drain (S/D) junction depths, $r_{\mathrm{j}}$, and the gate oxide thickness, $t_{\mathrm{ox}}$, as illustrated in Fig. 1.1. Adhering strictly to Dennard scaling results in the reduction in the footprint of a MOSFET by $1/\kappa^2$ and a reduction in its power dissipation by the same factor. This

therefore allows for an increase in transistor density by $\kappa^2$, which increases functionality, yet with no increase in power density per chip area. Furthermore, the $CV/I$ gate delay[1] reduces by a factor of $\kappa$, meaning that an IC comprising Dennard scaled MOSFETs can operate at increased speed. In practice, however, there are limitations to scaling in this manner as not all parameters can be arbitrarily scaled. Implicit in Dennard scaling is that that threshold voltage, $V_t$, must also be reduced by the same factor, however, the subthreshold swing (SS) cannot be scaled beyond the Boltzmann limit and therefore reducing $V_t$ increases the off-state current, $I_{off}$ [8]. The doping density is required to be increased by $\kappa$ in order to reduce the depletion widths, $W_D$, by approximately $\kappa$, however, the doping concentration can only be increased up until a certain point, beyond which there is significant mobility degradation in the channel due to increased ionised impurity scattering, and ultimately junction breakdown occurs [9]. The gate oxide thickness cannot be arbitrarily scaled due to the quantum mechanical tunnelling of carriers through the oxide which increases the leakage current [9]. Furthermore, defect densities increase in extremely thin $SiO_2$ films which degrade their electrical properties [9]. Table 1.1 summaries these limitations and the deviations they impose on ideal Dennard scaling.

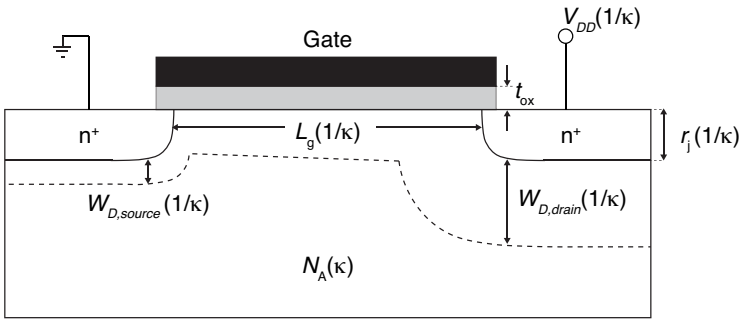


Figure 1.1: Schematic diagram of a metal oxide semiconductor field effect transistor (MOSFET) illustrating ideal geometric Dennard scaling.

Fig.1.2 shows the actual industry scaling trends from 1971 to 2015 in terms of transistor count, clock speed and total power dissipation. Up until the early 90's, scaling proceeded unabated and the total transistor count per processor increased at a rate close to Moore's law. This was achieved via a combination of increasing chip sizes and reducing device dimensions. During this time, $V_{DD}$ had only been reduced by a factor of $\sim$2 due to a reluctance to deviate from widely accepted industry standards. As a result, power consumption increased yet remained unproblematic. From the early 1990's to $\sim$2003, the trend of increasing chip size slowed due to cost constraints and the focus shifted to doubling the transistor density every two years as opposed to the previous three. Over this period, a precedence was placed on increased perfor-

[1] A useful metric for determining the switching speed of logic transistors is the inverter delay: the time taken to propagate a transition through a single inverter driving a second, identical inverter. It has been found empirically that the delay time calculated from $C_g V_{DD}/I_{on}$, where $C_g$ is the gate capacitance, $V_{DD}$ the supply voltage and $I_{on}$ the on-current, correlates closely to actual inverter delays. This therefore is an often used metric for logic transistors and is more commonly written as just $CV/I$ [8].

| Parameter | Dennard Scaling Factor | Actual Scaling factor | Limitation |
|---|---|---|---|
| $L_g$ | $1/\kappa$ | - | - |
| $t_{ox}$ | $1/\kappa$ | $>1/\kappa$ | Tunneling, defects |
| $r_j$ | $1/\kappa$ | $>1/\kappa$ | Resistance |
| $N_A$ | $\kappa$ | $<\kappa$ | Breakdown, $\mu$ degradation |
| $V_t$ | $1/\kappa$ | $\gg 1/\kappa$ | Off-current |
| $V_{DD}$ | $\kappa$ | $\gg 1/\kappa$ | $V_t$ |
| $\mathscr{E}$ | 1 | >1 | - |

Table 1.1: Summary of the ideal Dennard scaling parameters in comparison to those used in practice due to the limitations given.

mance in spite of increased power consumption and $L_g$ was scaled faster than other minimum feature sizes, while $V_{DD}$ continued to be sub-scaled. The result was that operating frequencies reached > 3 GHz significantly ahead of the trend, however, power consumption increased by a factor of 50 to > 100 $\mathrm{Wcm}^{-2}$. This was approaching the limit that could be tolerated without cooling costs becoming prohibitively expensive [5].

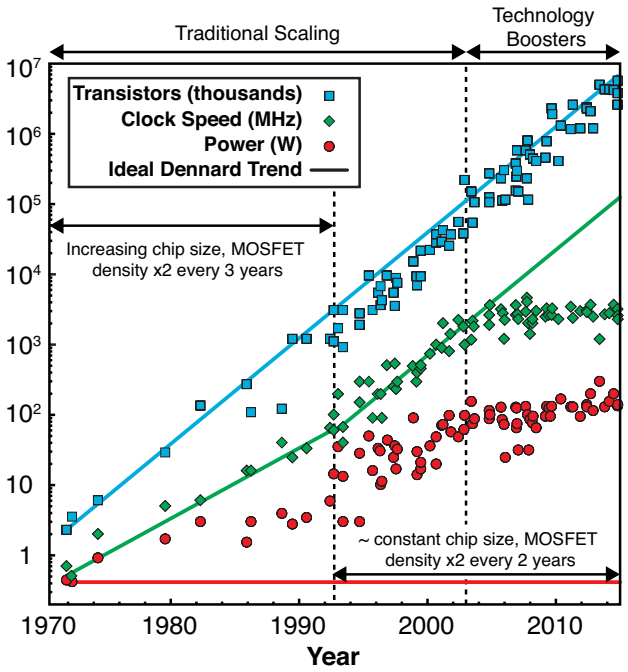


Figure 1.2: Comparison between ideal Dennard scaling (solid lines) and the actual industry scaling trends (symbols) between 1971 and 2015, in terms of number of transistors per integrated circuit (IC), clock speed, and total power dissipation. Data extracted from Ref. [10].

The total power consumption of a MOSFET comprises dynamic power, $P_D$, and static power, $P_S$, which result from switching and parasitic leakage currents respectively [8]. Historically, $P_D$ has dominated while $P_S$ has been negligible [8]. Dynamic power can be calculated from

$$P_D = C_{eff} V_{DD}^2 f \tag{1.1}$$

where $C_{eff}$ is the effective total capacitance being switched at frequency $f$ [8]. Clearly it is desirable to reduce $V_{DD}$ in order to reduce $P_D$, however, doing so increases $P_S$ due to the inability to scale $V_t$, as discussed above. As such, as $L_g$ has been scaled, $P_S$ has played an ever increasing role. This has been exacerbated by what are known as short channel effects (SCEs): for short channel lengths, the electric field emanating from the drain encroaches on the source, decreasing the source-channel barrier and inducing a negative $V_t$ shift. This shift, known as drain induced barrier lowering (DIBL), further increases the off-state leakage. The dynamic power and static power industry trends as a function of $L_g$ are shown in Fig. 1.3.

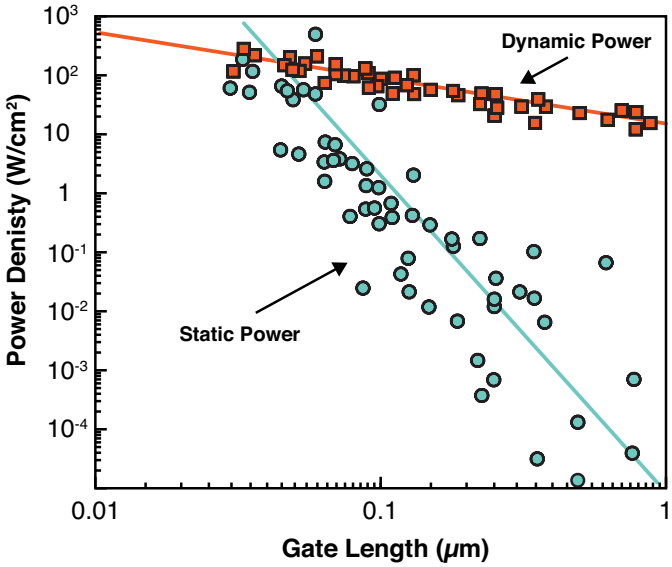


Figure 1.3: Dynamic and static power dissipation as a function of MOSFET gate length. Data extracted from Ref. [8].

To circumvent these issues, a variety non-traditional scaling 'technology boosters' have had to be progressively implemented beyond the 130 nm node. At the 90 nm node, compressive and tensile strained channels were employed to enhance electron and hole mobility ($\mu_n/\mu_p$) respectively [11]. This mitigated the effect of mobility degradation due to channel doping [11]. At the 45 nm node, the poly-silicon/$SiO_2$ gate stack was replaced by a metal gate and high-k dielectric which removed the poly depletion effect and allowed for a thicker physical oxide thickness to reduce the tunnelling current [11]. In 2011, the non-planar tri-gate architecture was introduced which reduced the SS from ∼100 mV/Dec. for planar devices to ∼70 mV/Dec. [12]. This allowed for a reduction in $V_{DD}$ and $V_t$ while maintaining gate overdrive and off-state performance. Furthermore, this architecture offered superior immunity to SCEs, which reduced DIBL. In 2014, the 14 nm node was put into production which further refined the 22 nm devices

by improving the fin sidewall verticality and increasing their height from 34 to 42 nm [13]. The former improved the electrostatic integrity allowing for shorter gate lengths and the latter, in combination with reduced fin spacing, increased the on-current per chip surface area [13]. This technology is expected to survive the 10 nm node, however, at this stage chip power densities have increased to $\sim$200 Wcm$^{-2}$ and consequently any further scaling mandates the reduction of $V_{\mathrm{DD}}$. This is not possible without comprising performance if Si is retained as the channel material.

## 1.2 The Advantage of III-V Compound Semiconductors

III-V Compound semiconductors offer superior electron transport properties in comparison to Si, and therefore their implementation in future n-channel metal oxide semiconductor (NMOS) devices could allow for the reduction of $V_{DD}$ while simultaneously improving on-state performance without incurring further leakage [5, 14]. The advantage that III-Vs offer is often argued by invoking their excellent bulk electron mobilities [5]; the direct implication of which is that in a diffusive transport regime, carriers will drift to higher velocities than in Si and therefore yield increased currents. This is an oversimplification however given the extremely scaled gate lengths that Si complementary metal oxide semiconductor (CMOS) technology has achieved. Should III-Vs ever be adopted, they will now do so at gate lengths shorter than that corresponding to the 10 nm node, and as such, will operate in the ballistic/quasi-basllistic regime, not diffusive. When discussing ballistic transport an important figure of merit is the injection velocity, $v_{inj}$ [5, 15]: the velocity with which a carrier is injected into the channel. Given that no scattering occurs in the channel of a ballistic device, it is $v_{inj}$ that limits the maximum carrier velocity. In this regard, III-V compounds offer a significant advantage over Si: Fig.1.4 compares experimentally extracted injection velocities as a function of gate length for In(Ga)As high electron mobility transistors (HEMTs), and Si/strained Si MOSFETs [5, 16]. The values in this figure for the III-V devices were extracted at $V_{DD}$ = 0.5 V, whereas the data for Si was extracted at $V_{DD}$ = 1-1.3 V. Despite this discrepancy, $v_{inj}$ for the III-V devices is more than twice that of Si/strained Si at scaled gate lengths [5, 16].

The maximum velocity of carriers is not the only consideration, however. Current is the product of the charge density in the channel and the average velocity with which carriers travel [15], and it has been argued that in the ballistic limit, the low density of states (DOS) exhibited by III-Vs negates their performance advantage in terms of transport properties [17, 18]. Further, it has been stated that the narrow band gap of III-Vs will increases leakage currents due to increased band to band tunneling (BTBT) [19]. Addressing the latter point first: given the sub 10 nm insertion point, III-V devices will have to embody non-planar architectures which geometrically confine the channel [5]. Quantisation effects in such a device will serve to effectively widen the bandgap and therefore reduce the proposed BTBT issues [5]. Moreover, the plethora of lattice matched III-V compounds allows great flexibility for heterostructure engineering, which can be utilised for example to fabricate a device with a narrow band gap, high $v_{inj}$ source, and wider band gap, low-leakage drain [20]. The former argument with regards to on-state performance fuels an interesting discussion, and recent results have shown the above assertion to not reflect the reality. Firstly, the effective mass in the channel of III-V devices is significantly higher than the bulk value, which is a result of strong non-parabolicty of the conduction band coupled with electron quantisation in confined channels [5]. Consequently, the degradation to on-state performance due to a low inversion capacitance is not as substantial as proposed [5]. Secondly,

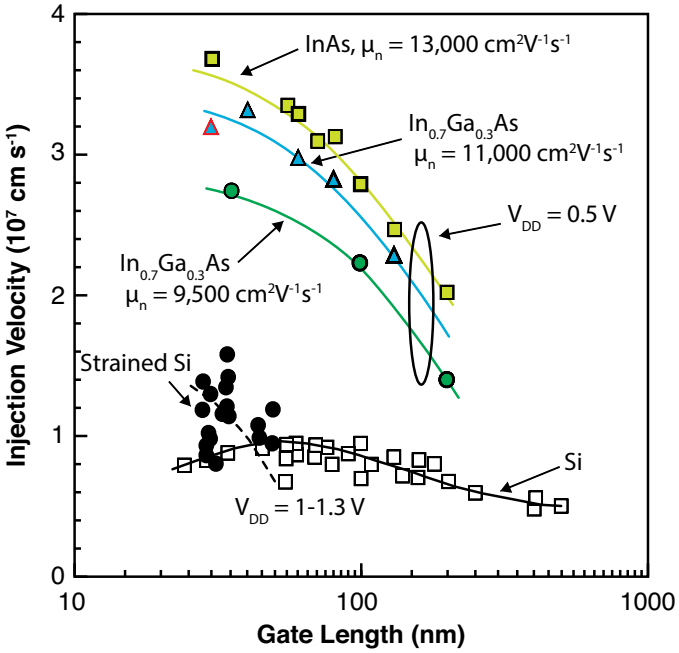


Figure 1.4: Comparison of the electron injection velocity, $v_{\text{inj}}$, as a function of gate length, for In(Ga)As HEMTs and Si/strained Si MOSFETs. The red highlighted data point corresponds precisely to the predicted ballistic $v_{\text{inj}}$ of an $In_{0.7}Ga_{0.3}As$ HEMT, as calculated by Monte Carlo simulation, indicating fully ballistic transport. Data extracted from [5, 16].

the above does not consider the extent to which the different material systems can operate at their ballistic limit [18]. An anomalous degradation in the mobility of Si MOSFETs was observed as gate lengths were scaled to less than 100 nm, an effect which continued to worsen with further scaling of $L_{\text{g}}$ [21, 22]. The same effect was observed in GaAs HEMTs which led Shur *et al*. to propose the concept of a (non-physical) channel length dependant ballistic mobility[2] [23]. This mobility combines with the bulk value using Matthiessen's rule [23] and dominates at short gate lengths [23]. Shur's model accurately predicted the results observed in III-V HEMTs and found them to operate close to their ballistic limit [23]. This is a result which has subsequently been repeated for other III-V HEMTs [19, 24] as well as InGaAs [25] and InAs FETs [26]. Indeed, the reduction in $v_{\text{inj}}$ observed in Fig. 1.4 for the data point highlighted in red of $In_{0.7}Ga_{0.3}As$ has been shown to correspond closely to fully ballistic transport [19]. Conversely, scaled state of the art Si MOSFETs appear to operate at only $\sim 50-60\%$ of their ballistic limit (the ratio of experimental on-current, $I_{\text{on}}$, to the theoretical ballistic current) [18, 27] and therefore suffer greater mobility degradation at short gate lengths compared to III-Vs. Increased electron-phonon

[2]The notion of a ballistic mobility is somewhat convoluted as the basis of mobility itself is intrinsically related to scattering: the mobility of a semiconductor arises in a diffusive transport regime - where scattering is present - and differs from material to material due to differing mean free path lengths between scattering events [9]. The term mobility therefore loses its physical meaning when used in reference to ballistic transport, where no scattering occurs. Nonetheless, while not yet having an explicit physical origin, the ballistic mobility of Shur *et al*. is a useful way of identifying a disparity in observed transport properties between Si and III-Vs.

scattering [28] in Si as well as backscattering of carriers at the drain contact [29] have been cited as potential causes of this degradation, however, for the latter it is not known why this would be worse in Si than for III-Vs [29, 30]. Further investigation is required to fully elucidate the discrepancy between the ballistic efficiency of Si and III-Vs, however, it appears that in this regard, III-Vs hold the advantage. Combining this with the above points suggest that III-Vs do indeed offer superior on-state performance than Si, even in the ultimately scaled device. Indeed, in recent years there has been exceptional progress made with III-V MOSFETs and the performance achieved certainly merits their continued investigation.

## 1.3 The Requirements for III-V MOS Integration

For III-V devices to be integrated into future CMOS ICs, there are two fundamental requirements which must be met: they must significantly outperform Si in order to merit their inclusion - which will pose disruption to the well established Si CMOS process - and they must do so while being economical to mass produce. This places stringent requirements on a variety of technological modules, such as gate stack and self aligned contact technologies. Rapid progress has been made in recent years in this regard and a comprehensive review of technologies that pertain to III-V MOSFETs, and indeed state of the art III-V MOSFETs themselves, is given in Chapter 2. While there are a variety of III-V technologies in terms of channel material, architecture etc. that could ultimately meet the demands required, whatever the specifics of the device itself, it must utilise the following:

- **A Non-Planar Architecture.** III-Vs are more susceptible to SCEs than Si due to their higher permittivity [5], and the improved electrostatic integrity afforded by non-planar architectures must be utilised to minimise off-state leakage. This is also important with regards to on-state performance. The high aspect ratio of 14 nm tri-gate Si technology means that while III-Vs may well hold the advantage in terms of intrinsic performance, unless they too exploit the vertical dimension, they will not compete in terms of on-current per chip area. This can be seen explicitly by comparing the on-state performance of the world leading III-V device (in terms of $I_{on}$) to 14 nm Si technology. The 75 nm gate length InGaAs FinFET demonstrated by Zota *et al*. in 2017 exhibits an on-current at $V_{DD} = 0.5$ V which is $\sim 215$ % higher than Si at $V_{DD} = 0.8$ V, when $I_{on}$ is normalised by gated perimeter. If $I_{on}$ is re-normalised by fin pitch, however, which corresponds to the chip surface area used, the high aspect ratio Si devices retain the advantage, as shown in Fig. 1.5.

- **Si CMOS Compatible Process Modules.** Vast sums of money have been invested in Si CMOS foundries and it is therefore critical that III-V devices can be fabricated with minimum disruption to this environment; their fabrication must be compatible with the existing toolset. This places a variety of limitations on III-V processing, all of which are discussed in the following chapter, however, one such limitation is that non-planar devices must be fabricated using a 'top-down' approach. Non-planar III-V devices exist in two flavours: grown 'bottom-up' and etched 'top-down'. Bottom-up processes either require non-CMOS compatible catalysts and or utilise processes which induce significant complexity and therefore cost. Conversely, dry etch is commonly used in Si CMOS and a top-down dry etched III-V device could readily be incorporated into a CMOS process. This is non-trivial, however, as III-Vs are highly susceptible to dry etch damage and the anisotropic etching of In-containing compounds is notoriously difficult.

Should a III-V NMOS device meet the above requirements, it may still not guarantee its imple-

mentation. Perhaps the greatest bottleneck to III-V metal oxide semiconductor (MOS) technology is the lack of equivalent performance for p-type devices. If this were to remain the case, III-V NMOS devices would have to be co-integrated with a different material system (likely Ge which, conversely, exhibits excellent p-channel metal oxide semiconductor (PMOS) performance but not NMOS) onto a common substrate. This could prove prohibitive to the economic viability, as co-integrating different material systems and concurrently fabricating complementary devices, where each device polarity requires different processing chemistries and has a different thermal budget, incurs great fabrication complexity and therefore cost. A far more desirable solution would be the implementation of p and n-type devices fabricated from the same material system, and more desirable still, the same material.

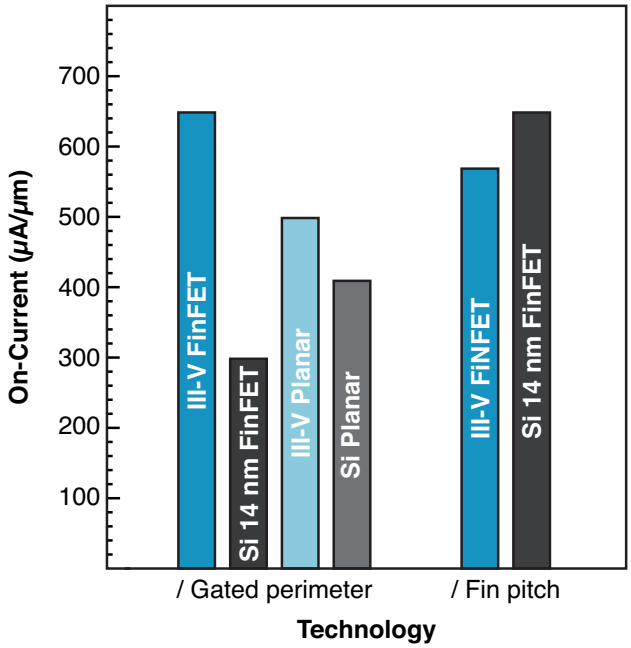


Figure 1.5: Comparison between the on-state performance of state of the art III-V and Si MOSFETs when on-current, $I_{on}$, is normalised by gated perimeter and fin pitch. Data extracted from [31].

## 1.4 Thesis Outline

This thesis presents two novel III-V compound semiconductor technologies which pertain to the key challenges addressed in the previous section:

- **A Top-Down, High Aspect Ratio, Junctionless InGaAs FinFET**. The junctionless architecture was first demonstrated in 2010 and has since gained significant traction as it circumvents prohibitively expensive fabrication complexities which arise when scaling gate lengths beyond the current state of the art [32]. The architecture utilises a uniform doping profile throughout the source, drain, and channel regions of the device, and consequently, it offers an even greater simplification to the fabrication of scaled III-V devices than it does Si: removing the need to form highly doped, self-aligned S/D regions by complex re-growth processes [33]. Given the extremely scaled insertion point that III-Vs are facing, in addition to the critical requirement of cost-effective fabrication, the junctionless architecture is a promising candidate for future technology nodes. InGaAs was chosen as the material to investigate this architecture given its excellent electron transport properties and relatively mature gate stack and contact technology.

- **In-situ $H_2$ Plasma Cleaning of the InGaSb-$Al_2O_3$ Interface**. Antimony based compound semiconductors are unique among III-V compounds and exhibit excellent transport properties for both electrons and holes: InSb exhibits the highest $\mu_n$ of any semiconductor [5], while the $\mu_p$ of GaSb is second only to Ge [5]. Of course, given the discussion in Section 1.2, bulk mobility ($\mu$) does not fundamentally determine the performance at scaled gate lengths, however, $v_{inj}$ is, to the first order, proportional to $\sqrt{\mu}$ [34], and therefore $\mu$ serves as valid basis for investigation. InGaSb ternary compounds exhibit the optimal transport properties for electrons and holes in the same material, and could offer a solution for III-V CMOS devices with a common channel material [35]. To date, however, there has been no investigation into the electrical properties of the high-k-InGaSb interface. As such, this thesis presents the first investigation into the impact of in-situ $H_2$ plasma cleaning on the electrical properties of the InGaSb-$Al_2O_3$ interface.

While the work undertaken coherently pertains to III-V CMOS, the specifics of the results chapters are somewhat diverse. Accordingly, they have been written such that they are entirely self contained, each including an in depth discussion of the motivation behind their investigation, as well as a specific review of the corresponding literature, discussion of the relevant theory, and conclusion on the importance of their findings. Chapter 2 provides a more general review of the field as a whole, with the specific requirements of III-V technologies elaborated on, and state of the art devices compared to Si. Chapter 3 discusses the various fabrication techniques used in this work, with emphasis placed on the techniques that proved the most challenging for device realisation. Chapters 4 and 5 are the results chapters which are structured as discussed

above. Finally, Chapter 6 concludes on all of the work presented, re-iterating the key findings and discussing the ramifications.

# Bibliography


[1] J. S. Kilby, “Invention of the integrated circuit,” *IEEE Transactions on Electron Devices*, vol. 23, no. 7, pp. 648–654, 1976.

[2] "The Story of the Intel 4004", Intel, 2010. [Online]. Available: https://www.intel.co.uk/content/www/uk/en/history/museum-story-of-intel-4004.html (Accessed: June 2018)

[3] "Sparc M7 Tops 10 Billion Transistors", The Linley Group, 2014. [Online]. Available: http://www.linleygroup.com/newsletters/newsletter_detail.php?num=5219 (Accessed: June 2018)

[4] G. E. Moore, “Cramming more components onto integrated circuits,” *Proceedings of the IEEE*, vol. 86, no. 1, pp. 82–85, 1965.

[5] J. A. del Alamo, “Nanometre-scale electronics with III-V compound semiconductors,” *Nature*, vol. 479, no. 7373, pp. 317–323, 2011.

[6] "More-than-Moore White Paper", ITRS, 2014. [Online]. Available: http://www.itrs2.net/uploads/4/9/7/7/49775221/irc-itrs-mtm-v2_3.pdf (Accessed: June 2018)

[7] R. H. Dennard, F. H. Gaensslen, V. L. Rideout, E. Bassous, and A. R. LeBlanc, “Design of Ion-Implanted MOSFET’s With Very Small Physical Dimensions,” *IEEE Journal of Solid-State Circuits*, vol. 9, no. 5, pp. 256–268, 1974.

[8] E. J. Nowak, “Maintaining the benefits of cmos scaling when scaling bogs down,” *IBM Journal of Research and Development*, vol. 46, no. 2.3, pp. 169–180, 2002.

[9] S. M. Sze and K. K. NG, *Physics of Semiconductor Devices*, 3rd ed. Wiley, 2006.

[10] "40 years of microprocessor trend data", K. Rupp, 2015. [Online]. Available: https://www.karlrupp.net/2015/06/40-years-of-microprocessor-trend-data/ (Accessed: June 2018)

[11] K. J. Kuhn, "Moore's Law Past 32nm: Future Challenges in Device Scaling," in *13th International Workshop on Computational Electronics*, 2009, pp. 1–6.

[12] C. Auth, C. Allen, A. Blattner, D. Bergstrom, M. Brazier, M. Bost, M. Buehler, V. Chikarmane, T. Ghani, T. Glassman, R. Grover, W. Han, D. Hanken, M. Hattendorf, P. Hentges, R. Heussner, J. Hicks, D. Ingerly, P. Jain, S. Jaloviar, R. James, D. Jones, J. Jopling, S. Joshi, C. Kenyon, H. Liu, R. McFadden, B. McIntyre, J. Neirynck, C. Parker, L. Pipes, I. Post, S. Pradhan, M. Prince, S. Ramey, T. Reynolds, J. Roesler, J. Sandford, J. Seiple, P. Smith, C. Thomas, D. Towner, T. Troeger, C. Weber, P. Yashar, K. Zawadzki, and K. Mistry, "A 22nm high performance and low-power CMOS technology featuring fully-depleted tri-gate transistors, self-aligned contacts and high density MIM capacitors," in *Symposium on VLSI Technology*, 2012, pp. 131–132.

[13] S. Natarajan, M. Agostinelli, S. Akbar, M. Bost, A. Bowonder, V. Chikarmane, S. Chouksey, A. Dasgupta, K. Fischer, Q. Fu, T. Ghani, M. Giles, S. Govindaraju, R. Grover, W. Han, D. Hanken, E. Haralson, M. Haran, M. Heckscher, R. Heussner, P. Jain, R. James, R. Jhaveri, I. Jin, H. Kam, E. Karl, C. Kenyon, M. Liu, Y. Luo, R. Mehandru, S. Morarka, L. Neiberg, P. Packan, A. Paliwal, C. Parker, P. Patel, R. Patel, C. Pelto, L. Pipes, P. Plekhanov, M. Prince, S. Rajamani, J. Sandford, B. Sell, S. Sivakumar, P. Smith, B. Song, K. Tone, T. Troeger, J. Wiedemer, M. Yang, and K. Zhang, "A 14nm logic technology featuring 2nd-generation FinFET, air-gapped interconnects, self-aligned double patterning and a 0.0588 ţm2SRAM cell size," in *IEEE International Electron Devices Meeting (IEDM)*, 2014, pp. 3.7.1–3.7.3.

[14] I. Thayne, R. Hill, M. Holland, X. Li, H. Zhou, D. Macintyre, S. Thoms, K. Kalna, C. Stanley, A. Asenov, R. Droopad, and M. Passlack, "Review of Current Status of III-V MOSFETs," *ECS Transactions*, vol. 19, no. 5, pp. 275–286, 2009.

[15] K. Natori, "Ballistic metal-oxide-semiconductor field effect transistor," *Journal of Applied Physics*, vol. 76, no. 8, pp. 4879–4890, 1994.

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

[17] M. V. Fischetti and S. E. Laux, "Are GaAs MOSFETs worth building? A model-based comparison of Si and GaAs n-MOSFETs," in *IEEE International Electron Devices Meeting (IEDM)*, 1989, pp. 481–484.

[18] S. H. Park, Y. Liu, N. Kharche, M. Salmani Jelodar, G. Klimeck, M. S. Lundstrom, and M. Luisier, "Performance comparisons of III-V and strained-Si in planar FETs and nonpla-

nar FinFETs at ultrashort gate length (12 nm)," *IEEE Transactions on Electron Devices*, vol. 59, no. 8, pp. 2107–2114, 2012.

[19] D. K. Tejas Krishnamohan and K. C. Saraswat, *Fundamentals of III-V Semiconductor MOSFETs, eds S.Oktyabrsky and P.D. Ye*, 1st ed. Springer, 2010.

[20] O. Kilpi, J. Wu, J. Svensson, E. Lind, and L. Wernersson, "Vertical heterojunction InAs/InGaAs nanowire MOSFETs on Si with $I_{on}$= 330 $\mu$A/$\mu$m at $I_{off}$=100 nA/$\mu$m and $V_D$= 0.5 V," in *Symposium on VLSI Technology*, 2017, pp. T36–T37.

[21] A. Cros, K. Romanjek, D. Fleury, S. Harrison, R. Cerutti, P. Coronel, B. Dumont, A. Pouydebasque, R. Wacquez, B. Duriez, R. Gwoziecki, F. Boeuf, H. Brut, G. Ghibaudo, and T. Skotnicki, "Unexpected mobility degradation for very short devices : A new challenge for CMOS scaling," in *IEEE International Electron Devices Meeting (IEDM)*, 2006, pp. 1–4.

[22] K. Huet, J. Saint-Martin, A. Bournel, S. Galdin-Retailleau, P. Dollfus, G. Ghibaudo, and M. Mouis, "Monte carlo study of apparent mobility reduction in nano-mosfets," in *37th European Solid State Device Research Conference (ESSDERC)*, 2007, pp. 382–385.

[23] M. Shur, "Low ballistic mobility in submicron HEMTs," *IEEE Electron Device Letters*, vol. 23, no. 9, pp. 511–513, 2002.

[24] N. Neophytou, T. Rakshit, and M. S. Lundstrom, "Performance Analysis of 60-nm Gate-Length III-V InGaAs HEMTs: Simulations Versus Experiments," *IEEE Transactions on Electron Devices*, vol. 56, no. 7, pp. 1377–1387, 2009.

[25] J. Lin, Y. Wu, J. A. del Alamo, and D. A. Antoniadis, "Analysis of Resistance and Mobility in InGaAs Quantum-Well MOSFETs From Ballistic to Diffusive Regimes," *IEEE Transactions on Electron Devices*, vol. 63, no. 4, pp. 1464–1470, 2016.

[26] N. Kharche, G. Klimeck, D.-H. Kim, J. A. del Alamo, and M. Luisier, "Multiscale Metrology and Optimization of Ultra-Scaled InAs Quantum Well FETs," *IEEE Transactions on Electron Devices*, vol. 58, no. 7, pp. 1963–1971, 2011.

[27] A. Lochtefeld and D. A. Antoniadis, "On experimental determination of carrier velocity in deeply scaled NMOS: How close to the thermal limit?" *IEEE Electron Device Letters*, vol. 22, no. 2, pp. 95–97, 2001.

[28] M. Luisier, "Performance comparison of GaSb, strained Si, and InGaAs double-gate ultrathin-body n-FETs," *IEEE Electron Device Letters*, vol. 32, no. 12, pp. 1686–1688, 2011.

[29] K. Natori, H. Iwai, and K. Kakushima, "Anomalous degradation of low-field mobility in short-channel metal-oxide-semiconductor field-effect transistors," *Journal of Applied Physics*, vol. 118, no. 23, 2015.

[30] D. Antoniadis, "On Apparent Electron Mobility in Si nMOSFETs from Diffusive to Ballistic Regime," *IEEE Transactions on Electron Devices*, vol. 63, no. 7, pp. 2650–2656, 2016.

[31] C. B. Zota, F. Lindelow, L. Wernersson, and E. Lind, "InGaAs tri-gate MOSFETs with record on-current," in *IEEE International Electron Devices Meeting (IEDM)*, 2016, pp. 3.2.1–3.2.4.

[32] J.-P. Colinge, C.-W. Lee, A. Afzalian, N. D. Akhavan, R. Yan, I. Ferain, P. Razavi, B. O'Neill, A. Blake, M. White, A.-M. Kelleher, B. McCarthy, and R. Murphy, "Nanowire transistors without junctions," *Nature Nanotechnology*, vol. 5, no. 3, pp. 225–229, 2010.

[33] U. Singisetti, M. A. Wistey, G. J. Burek, A. K. Baraskar, B. J. Thibeault, A. C. Gossard, M. J. W. Rodwell, B. Shin, E. J. Kim, P. C. McIntyre, B. Yu, Y. Yuan, D. Wang, Y. Taur, P. Asbeck, and Y. Lee, "$In_{0.53}Ga_{0.47}As$ Channel MOSFETs With Self-Aligned InAs Source/Drain Formed by MEE Regrowth," *IEEE Electron Device Letters*, vol. 30, no. 11, pp. 1128–1130, 2009.

[34] Y. Liu, M. Luisier, A. Majumdar, D. A. Antoniadis, and M. S. Lundstrom, "On the Interpretation of Ballistic Injection Velocity in Deeply Scaled MOSFETs," *IEEE Transactions on Electron Devices*, vol. 59, no. 4, pp. 994–1001, 2012.

[35] Z. Yuan, A. Nainani, A. Kumar, X. Guan, B. R. Bennett, J. B. Boos, M. G. Ancona, and K. C. Saraswat, "InGaSb: Single channel solution for realizing III-V CMOS," in *Symposium on VLSI Technology*, 2012, pp. 185–186.

# Chapter 2

# III-V Compounds for Digital Logic: A Review

## 2.1 Overview

Historically, the biggest impediment to the progression of III-V MOSFET technology has been the unsuitability of any III-V's native oxide for use as a gate dielectric []. In stark contrast, the real assets of Si technology are the stability and excellent electrical properties of the Si-$SiO_2$ interface []. Achieving a suitable dielectric interface to a III-V compound is not the only challenge, however, and the difficulties faced are exacerbated by the advance state of Si CMOS. As gate lengths are scaled, for example, the S/D contact resistance becomes comparable to the channel resistance, and this places an increasing precedence on minimising contact resistivity with each successive technology generation. Achieving low contact resistances to III-V compounds, while utilising Si compatible process modules, is non-trivial, however. Other challenges include: forming non-planar device channels, also utilising Si CMOS compatible processes, and monolithically integrating III-V devices onto Si substrates, both of which must be achieved without damaging the III-V compound and consequently degrading its intrinsic performance advantage. A global research effort has been required to overcome the above issues, as well as others, in order for III-V MOSFETs to compete with state of the art Si devices. This chapter provides a detailed review of the progress made in III-V MOSFET technology, including a discussion of: the various surface cleaning and deposition techniques that have facilitated unpinned, high-k gate stacks to III-Vs; the rationale for the choice of III-V channel material and an overview of the various device architectures demonstrated thus far; the progression from optically aligned Au based contacts utilised in III-V HEMTs to self aligned, Au-free, low resistance contacts for logic devices; methods of monolithic integration of III-V materials onto Si substrates; and a summary of world leading device performance. The discussion is focussed on n-type technolo-

gies and device performance given that this is almost exclusively where significant progress has been made. The chapter concludes with a discussion of the need for a high performance III-V p-type MOSFET and provides an overview of the current state of the art.

## 2.2 Gate Stack Technology

### 2.2.1 Unpinning the III-V-Dielectric Interface

It is critical that the oxide of a MOSFET forms an interface to the semiconductor through which the electrostatic potential in the semiconductor can be freely modulated by an applied gate bias; this is referred to as an 'unpinned' interface. There is no III-V compound, however, whose interface to its native oxide is unpinned. Consequently, to realise a III-V MOSFET, said III-Vs defective, native oxide must be prevented from forming, or removed after it has formed, and an alternative oxide subsequently deposited. The first demonstration of an unpinned dielectric interface to a III-V compound was reported in 1995: Passlack *et al.* utilised in-situ molecular beam epitaxy (MBE) to grow $Ga_2O_3$ on GaAs which was shown to yield an interface that is thermodynamically stable [1, 2], has an extremely low mid gap $D_{it}$ of $\sim 3 \times 10^{10}$ cm$^{-2}$ eV$^{-1}$ [1, 3], and one which explicitly demonstrates a genuine inversion response for both p and n-type capacitors [2]. This was a significant breakthrough in III-V MOS technology and led to the first demonstration of a III-V MOSFET [4, 5].

In 2003, a further major advance was made when Ye *et al.* demonstrated a GaAs MOSFET which employed an $Al_2O_3$ gate oxide grown by atomic layer deposition (ALD) [6]. Utilising ALD to form the gate oxide is preferable to in-situ MBE growth as it is a more robust and scalable process, and one which is, and was at the time, commonly used throughout the Si industry [7]. This first demonstration exhibited surprisingly promising results given that the GaAs surface had been exposed to air. It was subsequently discovered that one of the precursors for $Al_2O_3$, Trimethylaluminium (TMA), etches the native oxides on the GaAs surface and passivates it such that they do not reform during subsequent ALD oxidation cycles; an effect deemed 'self cleaning'[6, 8, 9]. Given the success of these results, the ALD of high-k dielectrics on other III-Vs, including InGaAs [10, 11], InAs [12] and InP [13], quickly followed, and, as it transpired, produced yet superior results in terms of electrical performance (lower $D_{it}$, greater Fermi level movement [13, 14]) in comparison to the high-k-GaAs interface. density functional theory (DFT) subsequently elucidated the origin of this improvement. It was shown that on GaAs-$Al_2O_3$/$HfO_2$ interfaces, interface traps appear predominantly as a result of the As-As dimer and result in high $D_{it}$ located around mid gap, causing Fermi level pinning [15, 16]. DFT shows that in In containing compounds, the trap states due to group V dimers and group-III dangling bonds are located in energy well within the conduction band, reducing their impact on electrical performance [15].

The addition of pre-ALD surface cleaning processes, such as those using ex-situ $(NH_2)_4S$ [17], and in-situ $H_2$ and $N_2$ plasma [18], as well as post-ALD processing, such as annealing in forming gas [19], offered further improvements to the electrical properties of high-k interfaces to

In containing III-V compounds. Additionally, multi-layer stacks, such as the $Al_2O_3/HfO_2$ bi-layer [20], for example, were employed to exploit the advantageous interface properties of certain dielectrics ($Al_2O_3$ in this instance) while taking advantage of higher permittivity dielectrics away from the interface to reduce the total equivalent oxide thickness (EOT) of the stack. As a result, fully unpinned interfaces have been demonstrated where a genuine minority carrier response is observed for both p and n-type MOS capacitors, with mid gap $D_{it}$ as low as $\sim 3 \times 10^{11}$ cm$^{-2}$eV$^{-1}$ [20, 21] at EOTs as low as $\sim$1 nm [21].

### 2.2.2 Reliability

With the progress discussed above, the high-k-III-V interface no longer impedes the realisation of individual n-type III-V MOSFETs which outperform Si at equivalent gate lengths (summarised in Section 2.6). A more pressing issue is achieving III-V gate stacks with sufficient reliability that mass production is feasible [22]. A feature commonly observed in capacitance-voltage (CV) measurements of III-V metal oxide semiconductor capacitors (MOSCAPs) is substantial frequency dispersion in accumulation. The accepted explanation for this is the tunnelling of carriers from the semiconductor into border traps located throughout the depth of the oxide [23, 24]. This is prevalent in III-Vs as their conduction and valence band edges are aligned in energy with regions of particularly high densities of border traps in both $Al_2O_3$ and $HfO_2$ [25]. It has recently been shown that there is a correlation between carriers tunnelling into border traps in the oxide and bias temperature instability (BTI) [26], and it suggested that, as a result of this unfavourable energetic alignment, it is the reliability of III-V gate stacks that now poses the greatest impediment to the mass production of III-V MOSFETs [27].

In 2016, data from Imec was published which showed that the interaction between carriers in the semiconductor and border traps in the oxide could be reduced by the insertion of an (undisclosed) interfacial dipole layer (IL) which aligns the conduction band of InGaAs with a lower density of border traps in the oxide [28]. This work was furthered in 2017 and it was shown that a gate stack comprising the IL/$LaSiO_x$/$HfO_2$ could be realised which had an EOT of 1.15 nm, mid gap $D_{it}$ of $\sim 3 \times 10^{11}$ cm$^{-2}$eV$^{-1}$, and sufficient reliability as to exceed the international technology roadmap for semiconductors (ITRS) requirement of 10 years operation with $\Delta V_{fb} < 30$ mV [21]. This represents a true breakthrough in III-V MOS technology and demonstrates that it is now sufficiently mature for mass production.

## 2.3 Channel Material and Device Architecture

### 2.3.1 Channel Material

As discussed in Sections 1.2 and 1.4, bulk electron mobility is a logical starting point from which to base the choice of channel material (in the case of NMOS devices). Of III-V compounds, arsenides and antimonides have the highest bulk electron mobilities [29]. Antimonide MOS technology, however, is extremely immature due to challenges with material growth and device fabrication which are a consequence of the highly reactive nature of antimony containing compounds [30]. Indeed, as mentioned in Section 1.4, this is an area which is investigated in this work and the reader is referred to Chapter 5 for more details regarding the progression of III-V antimonide MOS technology. In contrast, arsenides are the most technologically mature III-V compounds, with extensive research having been conducted into gate stack formation and contact technology. Of the III-V arsenides, $In_xGa_{1-x}As$ and InAs are of the most interest. Of the two, InAs has the highest electron mobility ($\mu_n = 44,000$ cm$^2$/Vs [31]), however, the bandgap energy must also be considered in order to compromise between high on-state performance and low leakage off-state performance. The bandgap energy of InAs of 354 meV [31] is problematic in terms of meeting the ITRS targets for off-state performance [32]. The bandgap of $In_xGa_{1-x}As$ increases with increasing GaAs mole fraction, however, the corresponding electron mobility decreases [33], as summarised in Table 2.1. Accordingly, there is a trade-off to be made when selecting a III-V channel material in order to yield a device which significantly outperforms Si in the on-state, and consequently facilitates the reduction in $V_{DD}$, while simultaneously retaining a sufficiently low leakage current such that the power consumption benefit gained by a reduced $V_{DD}$ is not negated by an increased static power dissipation.

| Material | $In_{0.3}Ga_{0.7}As$ | $In_{0.53}Ga_{0.47}As$ | $In_{0.7}Ga_{0.3}$ | InAs | InSb |
|---|---|---|---|---|---|
| $\mu_n$ (cm$^2$/Vs) | ~7,600 | 12,000 | ~20,000 | 44,000 | 78,000 |
| $E_g$ (meV) | ~1,000 | 740 | ~590 | 354 | 170 |

Table 2.1: Room temperature bulk electron mobility and bandgap energies for InGaAs [33], InAs [31] and InSb [34].

$In_xGa_{1-x}$ with 53 % InAs mole fraction has been extensively investigated as it provides a desirable trade-off in this regard, and furthermore, it is lattice matched to InP, which is readily available in wafer diameters up to 3"; making it ideal for process development and benchmarking [29]. Other InAs mole fractions, and indeed InAs, have also been researched and can be grown pseudomorphically on either InP or GaAs substrates depending on the composition [29]. Additionally, devices have been demonstrated which utilise In(Ga)As channels with wider bandgap materials incorporated in the drain in order to exploit the advantageous transport properties of narrower band gap materials while reducing the leakage current due to tunnelling

mechanisms at the drain [32, 35, 36].

### 2.3.2 Architecture

A variety of different III-V MOSFET architectures have been demonstrated, including planar-quantum well [37] devices, lateral FinFETs [38], and lateral and vertical gate all around (GAA) nanowires [36, 39]. The primary considerations of the architecture of a III-V MOSFET for logic applications are twofold: it is required to provide low access resistance to the channel, and it is required to provide sufficient electrostatic control of the channel such that devices can be scaled to the ultra short gate lengths exhibited by Si CMOS.

#### Access Resistance

In order to provide low access resistance to the channel, a high doping concentration is required in the S/D regions to facilitate low resistance, tunnelling ohmic contacts [40]. In Si devices, this is achieved by selectively doping the S/D regions using ion implantation, however, this method is not suitable for III-V MOSFETs as it yields comparatively low maximum doping densities due to poor activation efficiency, can result in a loss of stoichiometry due to preferential group V evaporation, and can cause amorphization due to sputtering damage which cannot be recovered by high temperature annealing [41]. The two most commonly adopted architectures for III-V MOSFETs which achieve highly doped S/D regions without the need for ion implantation are: the re-grown S/D arhicture [41], and the recessed gate architecture [29]. These two schemes are depicted in Fig. 2.1. Re-grown S/D architectures utilise metal organic chemical vapour deposition (MOCVD) to selectively grow highly doped S/D regions during the device fabrication process, subsequent to the initial epitaxial growth of the channel material and buffer layers. Conversely, recessed gate architectures include a highly doped cap layer during epitaxial growth, which is subsequently selectively etched in the gate region of the device to reveal the channel material. Both of these strategies have proved successful in fabricating state of the art devices [38, 42], with their main drawbacks being additional complexity, and therefore cost, to that of the standard Si CMOS process. This is another area which results presented in this thesis address, and the reader is referred to Chapter 4, which presents the fabrication, measurement and analysis of a novel architecture - the junctionless FinFET - which circumvents the requirement to form doping concentration gradients between the S/D and channel regions, and offers an architecture which is inherently scalable with a greatly simplified fabrication process.

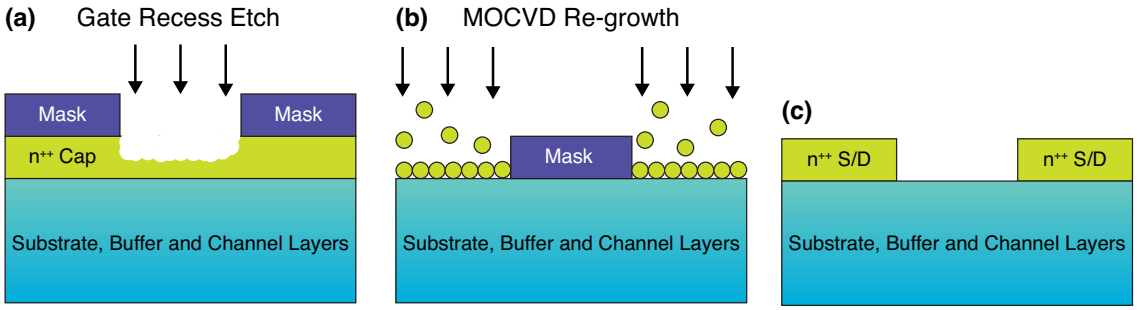


Figure 2.1: Comparison between (a) recessed gate and (b) re-grown source/drain process flows, both of which yield highly doped, raised source/drain regions as depicted in (c).

### Electrostatic Integrity

With regards to electrostatic integrity, it is pivotal that the device architecture provides sufficient control of the channel to mitigate SCEs, however, it is equally important that in doing so, the architecture does not degrade the carrier transport such that the advantage in comparison to Si is lost. Early endeavours into III-V MOSFETs found that at high electric fields, devices with surface channels - where the gate oxide is formed directly on the channel material - exhibited lower experimental mobilities than Si [29]. This was a consequence of increased scattering in the channel resulting from its close proximity to the gate stack [29]. Buried channel devices circumvent this issue by including a wide bandgap buffer which spatially separates the channel from the gate [29]. This architecture is problematic for scaling, however, as the oxide capacitance is limited by the inclusion of the widebandgap buffer. Both architectures are depicted in Fig. 2.2. As III-V MOS technology has progressed, the surface roughness at the III-V-high-k interface has reduced, and consequently, employing a surface channel architecture is no longer problematic: experimental mobilities as high as 3,500 $cm^2/Vs$ have been demonstrated in surface channel devices [21] and indeed the world leading III-V MOSFET utilises a surface channel [38] (discussed further in Section 2.6).

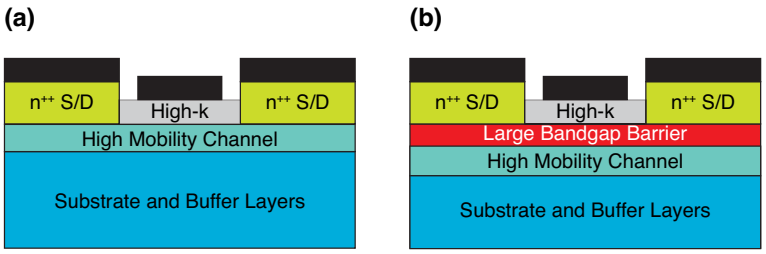


Figure 2.2: Comparison between (a) surface channel and (b) buried channel III-V mosfet architectures.

The electrostatic integrity of a planar surface channel device is no longer sufficient to compete with Si CMOS, however. Non-planar architectures are required to sufficiently suppress SCEs and have been implemented in Si CMOS since the 22 nm node [43, 44]. Doing so in a III-

V device, using a Si CMOS compatible top-down process, is non trivial, however: III-Vs are more susceptible to dry etch damage which cannot be subsequently recovered by annealing [7, 29]. Furthermore, it has been shown that sidewall verticality is important in order to suppress leakage currents in FinFETs [45], and achieving a low damage vertical etch process with In-containing compounds is notoriously difficult [46]. Chapter 4.5 provides a more in depth review of low damage, anisotropic etching of In-containing III-Vs. A critical enabling technology in achieving high performance, non-planar III-V MOSFETs is what is known as digital etch(ing) (DE) [47]. DE comprises alternating cycles of oxidation of the semiconductor surface followed by the subsequent selective etching of the (self limited) oxide formed. In this way, damaged layers of semiconductor can be controllably removed.

Fig. 2.3 summarises the various non-planar architectures demonstrated thus far. Architectures where the gate wraps entirely around the channel (b,c), known as GAA, offer improved immunity to SCEs in comparison to FinFETs, where the channel is only gated on three of it's four sides. GAA devices have been demonstrated in both lateral (b) and vertical orientations (c); vertical devices offer a unique advantage in that they decouple the gate length from the device footprint and therefore allow for increased density scaling.

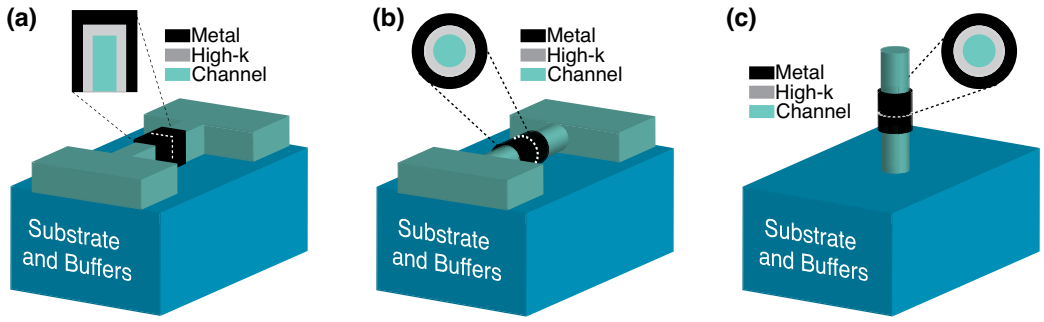


Figure 2.3: Illustration of (a) FinFET, (b) lateral gate all around and (c) vertical gate all around nanowire architectures.

## 2.4 Contact Technology

The contacts to scaled III-V MOSFETs must be self-aligned, as is the case for Si CMOS devices [48], in order to minimise the gate-contact distance and therefore minimise the access resistance to the channel as well as the footprint of the device. Further, in order to adhere to Si compatible processing, contacts must be Au free: Au readily diffuses and induces deep trap levels in the Si bandgap and therefore is banned from Si CMOS processes [41]. Moreover, specific contact resistances of 0.5 $\Omega\mu m^2$ must be achieved while adhering to the above constraints in order to meet the requirements for their insertion beyond the 10 nm node [49].

In stark contrast to the above requirements, historically, III-V HEMTs have relied upon optically aligned, non-scaled (on the order of micrometeres), annealed Ni-Au-Ge contacts to provide low contact resistances [50]. This contact scheme degenerately dopes the semiconductor locally with Ge to yield a narrow Schottky barrier (SB) which carriers can tunnel through [50]. This is required for contacts to GaAs as surface states pin the Fermi level resulting in a large SB height irrespective of the metal work function [51], known as the Bardeen limit [52]. Narrower band gap In-containing III-Vs offer an advantage in comparison to GaAs due to the alignment of surface states at the metal-semiconductor interface: this causes the Fermi level to pin near the conduction band of InAs, which results in a low SB height for n-type contacts [53].

Au-free self aligned silicde like contact schemes have been demonstrated on InGaAs based on Ni [54], Co [55] and Pd [56], however the most promising contact technology is with non-alloyed refractory metals [49]. Refractory metals are desirable as they are highly thermally stable and can be easily dry etched allowing for simple integration into a device process flow [49].

Large scale in-situ sputtered Mo contacts have been demonstrated to yield $\rho_c$ as lows as 0.5 $\pm$0.3 $\Omega\mu m^2$ to InAs [57] and 1.3$\pm$0.5 $\Omega\mu m^2$ to InGaAs [57] doped to the mid $10^{19}$ $cm^{-3}$, which is the range required for scaled device integration and comparable with the best contacts to Si [7]. More recently, an ex-situ variant of this contact scheme to $1\times10^{19}$ $cm^{-3}$ doped $In_{0.53}Ga_{0.47}As$ was analysed using a nano transmission line method (TLM) structure, which utilise TLM gap spaces on the order of tens of nanometre's, and found to yield $\rho_c$=0.69 $\pm$0.3 $\Omega\mu m^2$ [49]. Furthermore it was subsequently integrated into self aligned InGaAs FinFET process which yielded state of the art performance [58].

## 2.5 III-V on Si

Given the high cost of III-V materials, it does not make economical sense to employ them for mechanical use as the substrate. For the cost effective mass production of III-V devices, it will be required to selectively form III-V regions on a Si substrate. This is problematic however given the difference in lattice constant and thermal expansion coefficient between III-Vs and Si, in addition to the fact that III-Vs are polar materials whereas Si is non-polar [59]. When films grown on a lattice mismatched substrate grow past their pseudomorphic critical thickness, the overgrown material relaxes with misfit dislocations forming at the interface and threading dislocations propagating through the overgrown layer [59]. In the case of III-V arsenides and phosphides, threading dislocations are orientated along the $\{111\}$ plane and propagate at an angle of 60° [59]. Threading dislocations act as recombination and scattering centres which degrade carrier transport [59]. One method to minimise threading dislocation densities (TDD) is to employ a thick, compositionally graded buffer to bridge the gap in lattice constant between the substrate and the device layer [59]. The maximum thickness of a buffer layer is limited however by the mismatch in thermal expansion coefficients, as cracks can form upon temperature changes, such as when the wafer cools post growth [59]. Another complication is the formation of anti-phase domains (APD), which is inherent to polar on no-polar growth [59]. In reality, the (001) Si substrate surface is not perfectly flat but rather comprises monoatomic steps. When III-Vs are grown on a surface with monolayer (or an odd number of) steps, anti-phase boundaries (APB) are formed where the two III-V domains are separated by a plane of similar bonds: either III-III or V-V [59]. This does not happen with double layer (or an even number of) steps. APBs are electrically charged defects which degrade the mobility of III-Vs [59]. It was discovered that growing on off cut Si surfaces with a 4-6° tilt towards the [110] resulted in the single atomic steps reorganising into energetically more stable double atomic steps under high temperature annealing conditions [59]. APBs can be minimised in this manner however the use of an off cut substrate is undesirable for Si CMOS compatible mass production, where the standard is (001) orientated Si substrates which are exactly orientated within a tolerance of $\pm$ 0.5° [60].

A more favourable solution to forming low defect density III-V regions on Si substrates is the use of aspect ratio trapping (ART). ART grows the III-V layer within a high aspect ratio trench, as shown in Fig. 2.4, which 'traps' threading dislocations propagating orthogonally to the direction of the trench in the sidewall, preventing them from reaching the III-V surface [61]. This can be done on the nanometer scale meaning that shallow trench depths can be used, negating any cracking issues due to thermal expansion. Furthermore, ART can utilise (001) substrates: prior to III-V growth, the Si substrate at the base of the trench can be wet etched to from a (111) v-grooved surface localised to the trench only, leaving a (001) Si surface elsewhere on the wafer for the fabrication of alternative devices if desired [59].

A direct extension of the ART technique is confined epitaxial lateral overgrowth (CELO) [62]. While ART only geometrically filters defects which are in the direction of the high aspect ratio, CELO filters defects in 2 dimensions. This is achieved by growing the III-V layer initially within a confined seed region and subsequently overgrowing into a cavity, which is formed by selectively etching a sacrificial layer. The change in growth direction from vertical in the seed region to lateral in the cavity further aids defect filtering [62]. This process is shown in Fig. 2.5. transmission electron microscope (TEM) inspection of InP grown on Si using this method revealed a perfectly crystalline structure with 8% mismatch to the Si substrate, corresponding to fully relaxed InP [62].

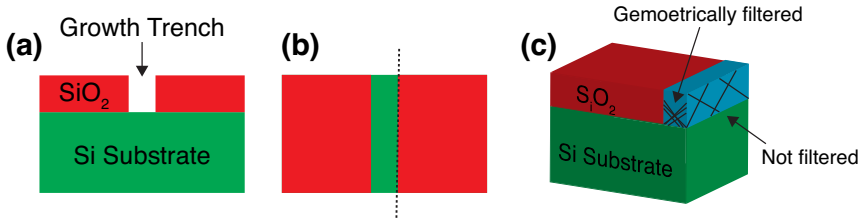


Figure 2.4: Illustration of the aspect ratio trapping (ART) technique. (a,b) show cross sectional and plan view schematics of the Si substrate and high aspect ratio trench prior to growth. (c) illustrates how threading dislocations are geometrically filtered, post III-V growth, with the view shown sliced along the dotted line in (b).

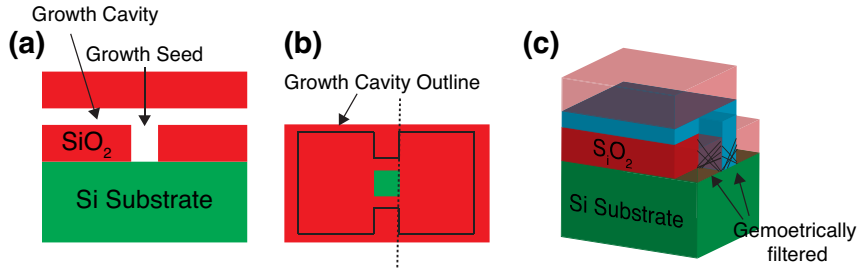


Figure 2.5: Illustration of the confined epitaxial lateral overgrowth (CELO) technique. (a,b) show cross sectional and plan view schematics of the Si substrate, seed region, and confined cavity. (c) illustrates how threading dislocations are geometrically filtered in 2 dimensions, post III-V growth, with the view shown sliced along the dotted line in (b).

A solution to yield III-V on insulator (XOI) has been demonstrated using direct wafer bonding (DWB) [59]. DWB bonds a III-V wafer to a Si substrate which includes a buried oxide layer (BOX) by bringing the two, cleaned, surfaces into intimate contact and annealing [59]. Subsequently the back of the III-V substrate can either be etched [63] or split [64] from the III-V device layers; yielding a III-V on insulator on Si structure as shown in Fig. 2.6. Using the splitting technique is advantageous as the III-V substrate can be recycled for repeating the process [64]. Growing the III-V layer on a III-V substrate limits the wafer size that can be bonded however. A solution to this is to utilise the techniques discussed above to grow the III-V layer

on a Si donor wafer, which can then be used to transfer the III-V layer to an silicon on insulator (SOI) wafer [59].

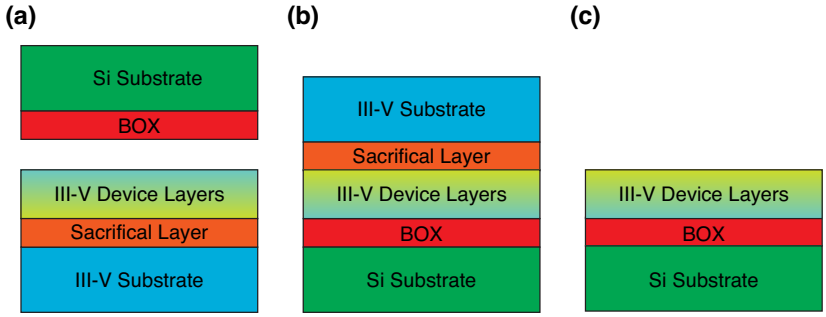


Figure 2.6: Illustration of the direct wafer bonding process (DWB). (a) shows the III-V and Si substrates with the device layers and BOX grown, prior to being brought into contact. (b) depicts the total layer structure subsequent to the bonding process. The III-V substrate can then be removed by etching or splitting from the device layers to yield the layer structure shown in (c).

## 2.6 State of the Art III-V NMOS

Figures 2.7 and 2.8 summarise the current state of the art III-V NMOS devices and compares them to Si NMOS in terms of $I_{on}$ and $Q$: $Q$ is equal to $g_m$/SS and simultaneously conveys the electrostatic control and transport properties of a device [65]. For all III-V data points shown, $I_{on}$ is measured at $V_g$ = 0.5 V from $I_{off}$, where $I_{off}$ = 100 nA/cm$^2$ and $V_d$ = 0.5 V. The values for Si devices have been scaled to $V_{DD}$ = 0.5 V as per References [7, 38]. Both planar and non-planar devices are shown, including those in 'bottom-up' and 'top-down' flavours, in addition to devices on both III-V and Si substrates. For non-planar devices, all currents are normalised by the gated perimeter. Selected key devices are highlighted and elaborated on below.

Rapid progress has been made it recent years and, as facilitated by the advancements in the technology modules discussed in the previous Sections, there are now multiple III-V devices which significantly outperform Si at equivalent gate lengths. The first III-V MOSFET to do so was demonstrated by Radosavljecvic *et al.* [37] in 2009. This device utilises a buried $In_{0.7}Ga_{0.3}As$ channel in a planar quantum well, recessed gate architecture and had a 75 nm gate length. The structure was grown on an off cut Si substrate utilising a thick, compositionally graded buffer. The performance of this device in terms of $I_{on}$ in fact remains the highest reported of any III-V MOSFET on Si, however it's very-large-scale integration (VLSI) potential is limited given the buried channel architecture and method of Si integration posing problems in terms of scaling and CMOS compatibility respectively, as discussed in Sections 2.3.2 and 2.5.

The first fully VLSI compatible, scalable device, fabricated on 300 mm Si (001) substrates with well behaved transistor characteristics was demonstrated by Waldron *et al.* in 2014 [61]. The devices utilised the ART technique to grow individual, 55 nm wide InGaAs Fins on top of an InP buffer and achieved $I_{on}$, $g_{m,sat}$ and $SS_{sat}$ values of 29.7 $\mu$A/$\mu$m, 558 $\mu$S/$\mu$m and 190 mV/Dec. respectively. The process was further refined to remove the InP beneath fin, creating a lateral GAA nanowire [66]. Subsequent improvements to the gate stack process via DE, sulfur passivation and the insertion of an undisclosed interfacial layer yielded a minimum mid gap $D_{it}$ as low as $9 \times 10^{11}$ cm$^{-2}$eV$^{-1}$ [39]. The optimised GAA device exhibited $I_{on}$, $g_{m,sat}$ and $SS_{sat}$ values of 175 $\mu$A/$\mu$m, 2190 $\mu$S/$\mu$m and 90mV/Dec. respectively. The Improved off-state performance in comparison to the equivalent FinFET was attributed to the improved electrostatic control afforded by the GAA architecture and the lower $D_{it}$ gate stack. The improved on-state performance was attributed to the removal of Mg in the channel, which was previously required to p-dope the InP buffer in order to reduce leakage, and unintentionally diffused into the InGaAs layer [39].

Sun *et al.* [73] demonstrated an InGaAs FinFET on Si in 2017, also using the ART technqiue, with record performance in terms of $I_{on}$ at scaled gate lengths for a VLSI compatible device. The parasitic bi-polar effect (PBE) is particularly problematic in scaled III-V devices due to

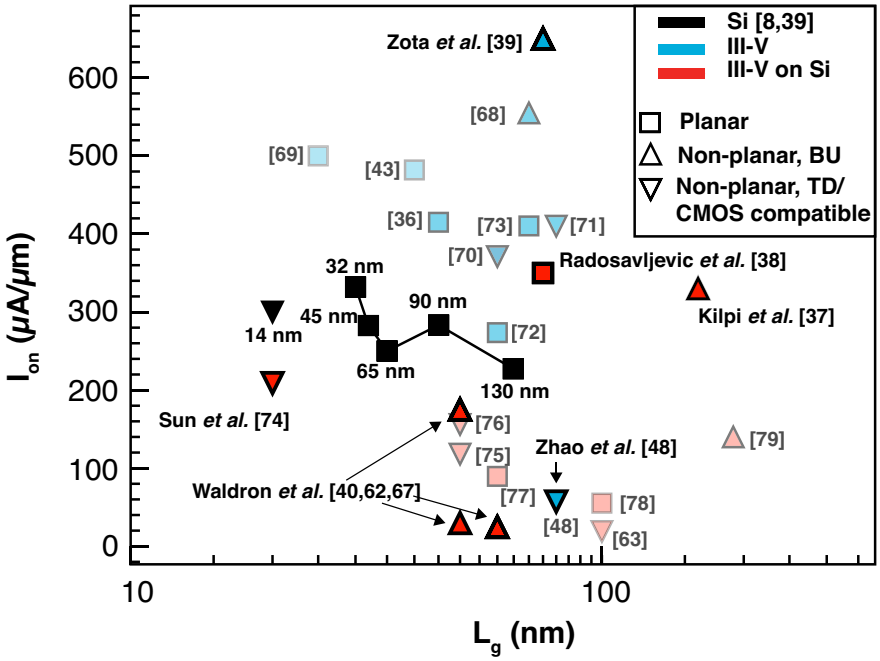


Figure 2.7: Summary of the state of the art III-V NMOS in terms of $I_{\text{on}}$ compared to Si. The values for Si were scaled to $V_{DD} = 0.5$ V as per References [7, 38]. The data for III-V devices is from the following References: [35–39, 42, 47, 61, 62, 66–78]. The data points which are discussed in the text are highlighted.

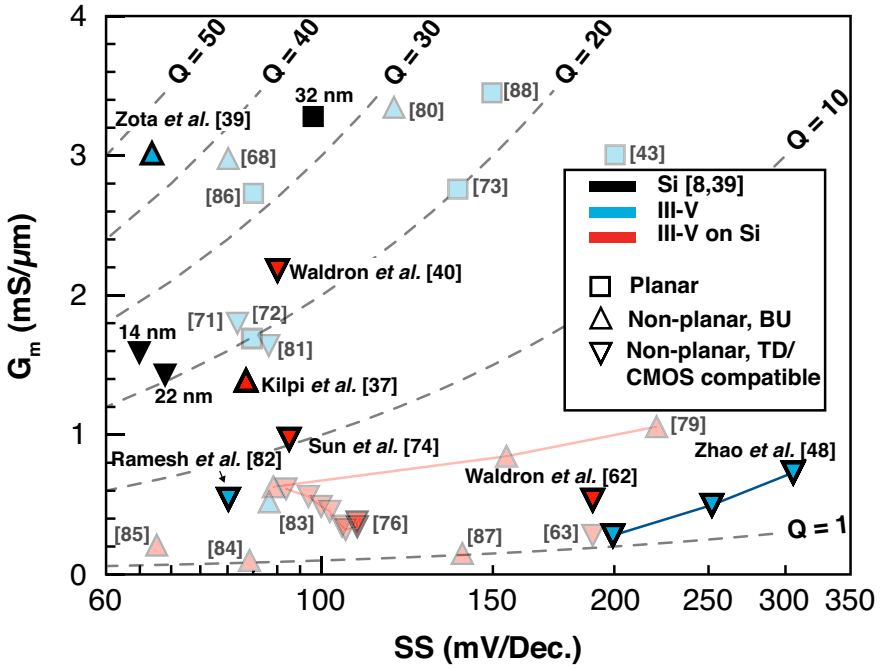


Figure 2.8: Summary of the state of the art III-V NMOS in terms of $Q$ factor compared to Si. The data for III-V devices is from the following References: [36, 38, 39, 42, 47, 61–63, 67, 70–73, 75, 78–86]. The data points which are discussed in the text are highlighted.

their narrow bandgaps [87]. This issue is further exacerbated by the fact the PBE increases with architectures that are required at scaled gate lengths in order to retain electrostatic integrity, such as XOI and GAA devices [87]. Consequently, demonstrating scaled III-V devices on Si can be problematic. By optimising their ART process, Sun et al improved the quality of their InP buffer, allowing for a more efficient conduction path for holes generated by BTBT, and limiting parasitic bi-polar action. This facilitated 32 nm gate length devices with $I_{on}$, $g_{m,sat}$ and $SS_{sat}$ values of 250 $\mu$A/$\mu$m, 971 $\mu$S/$\mu$m and $\sim$ 93 mV/Dec. respectively.

Unsurprisingly, the best performing III-V MOSFET demonstrated thus far was fabricated on a III-V substrate: despite the advances made in III-V on Si technology, it is inevitable that the defect density will be higher in III-Vs on Si compared to those grown on native substrates. In 2017, Zota *et al.* demonstrated an $In_{0.85}Ga_{0.15}As$ FinFET with the highest $I_{on}$ of any MOSFET demonstrated thus far [38]. The fin was formed by selective area growth using MOCVD on an InP substrate. The contact scheme used InGaAs regrowth, also by MOCVD. The gate stack process employed DE, sulufur passivation and utilised a bi-layer of $Al_2O_3$ and $HfO_2$ which yielded an EOT of $\sim$ 1 nm. The device exhibited of $I_{on}$, $g_{m,sat}$ and $SS_{sat}$ values of 650 $\mu$A/$\mu$m, 3 mS/cm$^2$ and $\sim$ 67 mV/Dec. respectively which correspond to the highest $I_{on}$ and $Q$ (45) of any MOSFET.

The first top-down III-V nanowire was demonstrated by Zhao *et al.* in 2013 [47]. The current world leading top-down device was demonstrated by Ramesh *et al.* in 2017 [81] which was the first to utilise arrays for etched wires as oppose to single wire devices and exhibited excellent SS performance ($SS_{sat}$=80mV/Dec.) and a competitive $g_{m,sat}$ of 500 $\mu S/\mu$m.

Bottom-up vertical III-V devices are more mature, and high performance has been demonstrated since 2008 [82]. A particular current bottom-up vertical device of note was demonstrated by Kilipi et al in 2017 [36]. It was grown on a Si substrate, and utilises an InAs channel with an InGaAs drain to minimise off-state leakage. In doing so it achieves a particularly large $I_{on}$ of 330 $\mu$A/$\mu$m given it's comparatively long gate length of 130 nm. This is facilitated by a large $g_{m,sat}$ (1.4 $\mu S/\mu$m) coupled with impressive off-state performance ($SS_{sat}$=85 mV/Dec., Q=16).

## 2.7 III-V PMOS

From a CMOS circuit design point of view, it is favourable to have complementary devices which exhibit comparable on-state performance. PMOS devices however are generally worse than NMOS due to a lower mobility of holes in comparison to electrons. For Si devices, this yields a disparity in $I_{on}$ between PMOS and NMOS of $\sim 3\times$, which circuit designers have been able to facilitate. For future CMOS technologies, it would be undesirable for this gap in performance to widen any further. This poses a problem for an all III-V CMOS technology as bulk holes mobilities are orders of magnitude lower than bulk electron mobilities for the majority of III-V materials; particularly those which are technologically mature. Currently, the most mature, high mobility material suitable for PMOS devices is Ge, however co-integrating different materials systems incurs great fabrication complexity and therefore cost. Antimonides have recently emerged as promising candidates for III-V PMOS devices owing to their excellent bulk hole transport properties however, as previously mentioned, this is an extremely immature material system with regards to MOSFET technology, with comparatively little published on gate stack or contact optimisation. Antimonide devices have been demonstrated however most employ long channels, have been measured at differing supply voltages, and only report certain noteworthy metrics; making benchmarking their performance in the same manner as in the previous Chapter difficult. Rather, the following highlights some key devices and certain enabling technologies which have demonstrated promising results.

In 2008, Radosaljevic *et al.* demonstrated a quantum well FET which utilised a biaxial compressively strained (1-2 %) InSb channel with gate lengths as scaled as 40 nm [88]. While this device was described as a FET, it was in fact HEMT-like in that the channel was gated through an AlInSb wide bandgap buffer as opposed to a dielectric. Nonetheless, it's impressive performance in terms of $g_{m,sat}$ (510 $\mu$S/$\mu$m), owing to an experimental effective hole velocity which was extracted as $>2\times$ that of Si, was the first device that highlighted the potential for antimonide PMOS devices for logic applications.

In 2010, Nainani *et al.* demonstrated the first genuine antimonide MOSFET [89]. The device utilised a biaxial compressively strained InGaSb channel (0.7%-1.7%) with a 5 $\mu$m gate length. Both surface channel and buried channel devices were demonstrated. A 10 nm thick $Al_2O_3$ layer was used as the gate dielectric which was deposited using ALD subsequent to cleaning the antimonide surface in HCl. This process, in addition to a post gate metal forming gas anneal (FGA), was shown to yield a mid gap $D_{it}$ as low as $3 \times 10^{11}\text{cm}^{-2}\text{eV}^{-1}$ on GaSb [89]. The resulting MOSFETs exhibited a minimum SS of 120 mV/Dec., $g_{m,sat}$ of 140 $\mu$S/$\mu$m and exhibited a peak hole mobility of 910 $\text{cm}^2$/Vs: offering a 400 % peak improvement over that demonstrated by Ge [89].

In 2012, the first antimonide PMOSFET on Si was demonstrated by Takei *et al.* [90]. The

devices were formed on Si substrates using a processes known as epitaxial layer transfer (ELT), which is a slight variant of the DWB technique discussed previously. An InGaSb channel was employed with an InAs cap through which the devices were gated. Both bottom gated - through the BOX - and top gated devices were demonstrated. The top gate devices utilised $ZrO_2$ as the gate dielectric which was deposited by ALD. The best performing devices exhibited an minimum SS of 130 mV/Dec., a maximum $g_{m,sat}$ of 36 $\mu$S/$\mu$m), and a peak effective hole mobility of 480 cm$^2$/Vs at $V_{DD}$=0.5 V [90].

In 2015, Lu *et al.* demonstrated the first InGaSb p-channel FinFET [91]. The devices were fabricated using a top-down approach which was facilitated by the development of a low damage antimonide dry etch process which yielded fins as narrow as 15 nm [91]. Additionally, extremely low resistance, Si compatible, Ni based ohmic contacts were developed which exhibited $\rho_c$ of $3.5\times10^{-8}$ $\Omega$cm$^2$. These developments yielded device performance with a maximum $g_{m,sat}$ of 122 $\mu$S/$\mu$m for 100 nm $L_g$. This work was improved upon in 2017 by incorporating an antimonde compatible digital etch [92]. DE processes for other III-V usually use acids which are buffered in water to selectively etch the oxide formed [93]. It was found that water in fact etches antimonides however and therefore the DE process was not self limited [92]. This was remedied by using a solution of HCl buffered in IPA which enabled the fabrication of devices with fin widths as small as 10 nm and gate lengths of 20 nm [92]. These devices demonstrated a record high $g_{m,sat}$ of 160 $\mu$S/$\mu$m however were plagued by poor off-state performance, with a minimum SS of only 290 mV/Dec. for devices with long gate lengths (1 $\mu$m) [92].

## 2.8 Chapter Summary

This chapter has reviewed the progress made in terms of III-V MOS modules which have facilitated the demonstration of n-type devices which now significantly outperform Si at equivalent gate lengths, in addition to discussing the current state of III-V PMOS. A literature review is also included in each result chapter that relates more specifically to the corresponding work presented.

# Bibliography


[1] M. Passlack, M. Hong, J. P. Mannaerts, S. N. G. Chu, R. L. Opila, and N. Moriya, "In-situ $Ga_2O_3$ process for GaAs inversion/accumulation device and surface passivation applications," in *IEEE International Electron Devices Meeting (IEDM)*, 1995, pp. 383–386.

[2] M. Passlack, M. Hong, and J. P. Mannaerts, "Quasistatic and high frequency capacitance-voltage characterization of $Ga_2O_3$-GaAs structures fabricated by in situ molecular beam epitaxy," *Applied Physics Letters*, vol. 68, no. 8, pp. 1099–1101, 1996.

[3] M. Passlack, M. Hong, J. P. Mannaerts, R. L. Opila, S. N. G. Chu, N. Moriya, F. Ren, and J. R. Kwo, "Low D Thermodynamically Stable $Ga_2O_3$-GasAs Interfaces: Fabrication, Characterization, and Modeling," *IEEE Transactions on Electron Devices*, vol. 44, no. 2, pp. 214–225, 1997.

[4] F. Ren, M. W. Hong, W. S. Hobson, J. M. Kuo, J. R. Lothian, J. P. Mannaerts, J. Kwo, Y. K. Chen, and A. Y. Cho, "Enhancement-mode p-channel GaAs MOSFETs on semi-insulating substrates," in *IEEE International Electron Devices Meeting (IEDM)*, 1996, pp. 943–945.

[5] F. Ren, M. Hong, W. Hobson, J. Kuo, J. Lothian, J. Mannaerts, J. Kwo, S. Chu, Y. Chen, and A. Cho, "Demonstration of enhancement-mode p- and n-channel gaas mosfets with $ga_2o_3(gd_2o_3)$ as gate oxide," *Solid-State Electronics*, vol. 41, no. 11, pp. 1751 – 1753, 1997.

[6] P. D. Ye, G. D. Wilk, J. Kwo, B. Yang, H. J. L. Gossmann, M. Frei, S. N. G. Chu, J. P. Mannaerts, M. Sergent, M. Hong, K. K. Ng, and J. Bude, "GaAs MOSFET with oxide gate dielectric grown by atomic layer deposition," *IEEE Electron Device Letters*, vol. 24, no. 4, pp. 209–211, 2003.

[7] J. A. del Alamo, "Nanometre-scale electronics with IIIâĂŞV compound semiconductors," *Nature*, vol. 479, no. 7373, pp. 317–323, 2011.

[8] M. M. Frank, G. D. Wilk, D. Starodub, T. Gustafsson, E. Garfunkel, Y. J. Chabal, J. Grazul, and D. A. Muller, "$HfO_2$ and $Al_2O_3$ gate dielectrics on GaAs grown by atomic layer deposition," *Applied Physics Letters*, vol. 86, no. 15, p. 152904, 2005.

[9] M. L. Huang, Y. C. Chang, C. H. Chang, Y. J. Lee, P. Chang, J. Kwo, T. B. Wu, and M. Hong, "Surface passivation of III-V compound semiconductors using atomic-layer-deposition-grown $Al_2O_3$," *Applied Physics Letters*, vol. 87, no. 25, pp. 1–3, 2005.

[10] P. D. Ye, G. D. Wilk, B. Yang, J. Kwo, H. J. L. Gossmann, M. Hong, K. K. Ng, and J. Bude, "Depletion-mode InGaAs metal-oxide-semiconductor field-effect transistor with oxide gate dielectric grown by atomic-layer deposition," *Applied Physics Letters*, vol. 84, no. 3, pp. 434–436, 2004.

[11] Y. Xuan, H. C. Lin, P. D. Ye, and G. D. Wilk, "Capacitance-voltage studies on enhancement-mode InGaAs metal-oxide-semiconductor field-effect transistor using atomic-layer-deposited $Al_2O_3$ gate dielectric," *Applied Physics Letters*, vol. 88, no. 26, p. 263518, 2006.

[12] N. Li, E. S. Harmon, J. Hyland, D. B. Salzman, T. P. Ma, Y. Xuan, and P. D. Ye, "Properties of InAs metal-oxide-semiconductor structures with atomic-layer-deposited $Al_2O_3$ Dielectric," *Applied Physics Letters*, vol. 92, no. 14, pp. 1–4, 2008.

[13] Y. Q. Wu, Y. Xuan, P. D. Ye, Z. Cheng, and A. Lochtefeld, "Inversion-type enhancement-mode InP MOSFETs with ALD $Al_2O_3$, $HfO_2$ and HfAlO nanolaminates as high-k gate dielectrics," in *65th Annual Device Research Conference (DRC)*, 2007, pp. 117–118.

[14] P. Ye, Y. Xuan, Y. Wu, and M. Xu, "Inversion-Mode $In_xGa_{1-x}As$ MOSFETs (X=0.53,0.65,0.75) with Atomic-Layer deposited High-K Dielectrics," *ECS Transactions*, vol. 19, no. 2, pp. 605–614, 2009.

[15] L. Lin and J. Robertson, "Defect states at III-V semiconductor oxide interfaces," *Applied Physics Letters*, vol. 98, no. 8, pp. 2009–2012, 2011.

[16] W. Wang, K. Xiong, R. M. Wallace, and K. Cho, "Impact of interfacial oxygen content on bonding, stability, band offsets, and interface states of GaAs:$HfO_2$ interfaces," *The Journal of Physical Chemistry C*, vol. 114, no. 51, pp. 22 610–22 618, 2010.

[17] É. O'Connor, B. Brennan, V. Djara, K. Cherkaoui, S. Monaghan, S. B. Newcomb, R. Contreras, M. Milojevic, G. Hughes, M. E. Pemble, R. M. Wallace, and P. K. Hurley, "A systematic study of $(NH_4)_2S$ passivation (22%, 10%, 5%, or 1%) on the interface properties of the $Al_2O_3/In_{0.53}Ga_{0.47}As$/InP system for n-type and p-type $In_{0.53}Ga_{0.47}As$ epitaxial layers," *Journal of Applied Physics*, vol. 109, no. 2, pp. 0–10, 2011.

[18] V. Chobpattana, J. Son, J. J. Law, R. Engel-Herbert, C. Y. Huang, and S. Stemmer, "Nitrogen-passivated dielectric/InGaAs interfaces with sub-nm equivalent oxide thickness and low interface trap densities," *Applied Physics Letters*, vol. 102, no. 2, pp. 2–5, 2013.

[19] H. D. Trinh, E. Y. Chang, P. W. Wu, Y. Y. Wong, C. T. Chang, Y. F. Hsieh, C. C. Yu, H. Q. Nguyen, Y. C. Lin, K. L. Lin, and M. K. Hudait, "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, pp. 53–56, 2010.

[20] E. 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.

[21] S. Sioncke, J. Franco, A. Vais, V. Putcha, L. Nyns, A. Sibaja-Hernandez, R. Rooyackers, S. C. Ardila, V. Spampinato, A. Franquet, J. W. Maes, Q. Xie, M. Givens, F. Tang, X. Jiang, M. Heyns, D. Linten, J. Mitard, A. Thean, D. Mocuta, and N. Collaert, "First demonstration of $\sim$3500 $cm^2$/V-s electron mobility and sufficient BTI reliability (max $V_{ov}$ up to 0.6V) $In_{0.53}Ga_{0.47}As$ nFET using an IL/$LaSiO_x$/$HfO_2$ gate stack ," in *Symposium on VLSI Technology*, 2017, pp. T38–T39.

[22] J. Franco, V. Putcha, A. Vais, S. Sioncke, N. Waldron, D. Zhou, G. Rzepa, P. J. Roussel, G. Groeseneken, M. Heyns, N. Collaert, D. Linten, T. Grasser, and B. Kaczer, "Characterization of oxide defects in InGaAs MOS gate stacks for high-mobility n-channel MOSFETs (invited)," in *IEEE International Electron Devices Meeting (IEDM)*, 2017, pp. 7.5.1–7.5.4.

[23] F. Heiman and G. Warfield, "The effects of oxide traps on the MOS capacitance," *IEEE Transactions on Electron Devices*, vol. 12, no. 4, pp. 167–178, 1965.

[24] Y. Yuan, B. Yu, J. Ahn, P. C. McIntyre, P. M. Asbeck, M. J. W. Rodwell, and Y. Taur, "A distributed bulk-oxide trap model for $Al_2O_3$ InGaAs MOS devices," *IEEE Transactions on Electron Devices*, vol. 59, no. 8, pp. 2100–2106, 2012.

[25] A. Vais, J. Franco, D. Lin, V. Putcha, S. Sioncke, A. Mocuta, N. Collaert, A. Thean, and K. De Meyer, "On the distribution of oxide defect levels in $Al_2O_3$ and $HfO_2$ high-k dielectrics deposited on InGaAs metal-oxide-semiconductor devices studied by capacitance-voltage hysteresis," *Journal of Applied Physics*, vol. 121, no. 14, pp. 3–10, 2017.

[26] T. Grasser, B. Kaczer, W. Goes, H. Reisinger, T. Aichinger, P. Hehenberger, P. Wagner, F. Schanovsky, J. Franco, M. T. Luque, and M. Nelhiebel, "The paradigm shift in understanding the bias temperature instability: From reactionâĂŞdiffusion to switching oxide traps," *IEEE Transactions on Electron Devices*, vol. 58, no. 11, 2011.

[27] J. Franco, B. Kaczer, A. Vais, A. Alian, H. Arimura, V. Putcha, S. Sioncke, N. Waldron, D. Zhou, L. Nyns, and et al., "Bias temperature instability (bti) in high-mobility channel

devices with high-k dielectric stacks: Sige, ge, and ingaas," *MRS Advances*, vol. 1, no. 49, p. 3329âĂŞ3340, 2016.

[28] A. Vais, A. Alian, L. Nyns, J. Franco, S. Sioncke, V. Putcha, H. Yu, Y. Mols, R. Rooyackers, D. Lin, J. W. Maes, Q. Xie, M. Givens, F. Tang, X. Jiang, A. Mocuta, N. Collaert, K. D. Meyer, and A. Thean, "Record mobility ($\mu_{eff}$ ~3100 $cm^2$/V-s) and reliability performance ($V_{ov}$ ~0.5V for 10yr operation) of $In_{0.53}Ga_{0.47}As$ MOS devices using improved surface preparation and a novel interfacial layer," in *IEEE Symposium on VLSI Technology*, 2016, pp. 1–2.

[29] I. Thayne, R. Hill, M. Holland, X. Li, H. Zhou, D. Macintyre, S. Thoms, K. Kalna, C. Stanley, A. Asenov, R. Droopad, and M. Passlack, "Review of Current Status of III-V MOSFETs," *ECS Transactions*, vol. 19, no. 5, pp. 275–286, 2009.

[30] W. Lu, I. P. Roh, D. . Geum, S. . Kim, J. D. Song, L. Kong, and J. A. del Alamo, "10-nm Fin-width InGaSb p-channel self-aligned FinFETs using antimonide-compatible digital etch," in *IEEE International Electron Devices Meeting (IEDM)*, 2017, pp. 17.7.1–17.7.4.

[31] "InAs Electrical Properties", IOFFE, 2018. [Online]. Available: http://www.ioffe.ru/SVA/NSM/Semicond/InAs/electric.html

[32] J. Mo, E. Lind, G. Roll, and L. E. Wernersson, "Reduction of off-state drain leakage in InGaAs-based metal-oxide-semiconductor field-effect transistors," *Applied Physics Letters*, vol. 105, no. 3, pp. 1–5, 2014.

[33] "InGaAs Electrical Properties", IOFFE, 2018. [Online]. Available: http://www.ioffe.ru/SVA/NSM/Semicond/GaInAs/electric.html

[34] "InSb Electrical Properties", IOFFE, 2018. [Online]. Available: http://www.ioffe.ru/SVA/NSM/Semicond/InSb/electric.html

[35] C. Y. Huang, S. Lee, V. Chobpattana, S. Stemmer, A. C. Gossard, B. Thibeault, W. Mitchell, and M. Rodwell, "Low Power IIIâĂŞV InGaAs MOSFETs featuring InP recessed source/drain spacers with $I_{on}$=120 $\mu$A/$\mu$m at $I_{off}$=1 nA/$\mu$m and $V_{DS}$=0.5 V," in *IEEE International Electron Devices Meeting (IEDM)*, 2014, pp. 25.4.1–25.4.4.

[36] O. Kilpi, J. Wu, J. Svensson, E. Lind, and L. Wernersson, "Vertical heterojunction InAs/InGaAs nanowire MOSFETs on Si with $I_{on}$= 330 ÎijA/Îijm at $I_{off}$=100 nA/Îijm and $V_D$= 0.5 V," in *Symposium on VLSI Technology*, 2017, pp. T36–T37.

[37] M. Radosavljevic, B. Chu-Kung, S. Corcoran, G. Dewey, M. K. Hudait, J. M. Fastenau, J. Kavalieros, W. K. Liu, D. Lubyshev, M. Metz, K. Millard, N. Mukherjee, W. Rachmady, U. Shah, and R. Chau, "Advanced high-K gate dielectric for high-performance

short-channel $In_{0.7}Ga_{0.3}As$ quantum well field effect transistors on silicon substrate for low power logic applications," in *IEEE International Electron Devices Meeting (IEDM)*, 2009, pp. 1–4.

[38] C. B. Zota, F. Lindelow, L. Wernersson, and E. Lind, "InGaAs tri-gate MOSFETs with record on-current," in *IEEE International Electron Devices Meeting (IEDM)*, 2016, pp. 3.2.1–3.2.4.

[39] N. Waldron, S. Sioncke, J. Franco, L. Nyns, A. Vais, X. Zhou, H. C. Lin, G. Boccardi, J. W. Maes, Q. Xie, M. Givens, F. Tang, X. Jiang, E. Chiu, A. Opdebeeck, C. Merckling, F. Sebaai, D. van Dorp, L. Teugels, A. S. Hernandez, K. D. Meyer, K. Barla, N. Collaert, and Y. Thean, "Gate-all-around InGaAs nanowire FETS with peak transconductance of 2200$\mu$S/Îijm at 50nm $L_g$ using a replacement Fin RMG flow," in *IEEE International Electron Devices Meeting (IEDM)*, 2015, pp. 31.1.1–31.1.4.

[40] S. M. Sze and K. K. NG, *Physics of Semiconductor Devices*, 3rd ed. Wiley, 2006.

[41] R. J. Hill, J. Oh, C. Park, J. Barnett, J. Price, J. Huang, N. Goel, W. Y. Loh, P. Kirsch, P. Majhi, and R. Jammy, "CMOS Scaling with III-V Channels for Improved Performance and Low Power," *ECS Transactions*, vol. 35, no. 3, pp. 335–344, 2011.

[42] S. Lee, C. Huang, D. Cohen-Elias, B. J. Thibeault, W. Mitchell, V. Chobpattana, S. Stemmer, A. C. Gossard, and M. J. W. Rodwell, "Highly Scalable Raised Source/Drain InAs Quantum Well MOSFETs Exhibiting $I_{on}$=482 uA/$\mu$m at $I_{off}$=100 nA/$\mu$m and $V_{DD}$=0.5 V," *IEEE Electron Device Letters*, vol. 35, no. 6, pp. 621–623, 2014.

[43] C. Auth, C. Allen, A. Blattner, D. Bergstrom, M. Brazier, M. Bost, M. Buehler, V. Chikarmane, T. Ghani, T. Glassman, R. Grover, W. Han, D. Hanken, M. Hattendorf, P. Hentges, R. Heussner, J. Hicks, D. Ingerly, P. Jain, S. Jaloviar, R. James, D. Jones, J. Jopling, S. Joshi, C. Kenyon, H. Liu, R. McFadden, B. McIntyre, J. Neirynck, C. Parker, L. Pipes, I. Post, S. Pradhan, M. Prince, S. Ramey, T. Reynolds, J. Roesler, J. Sandford, J. Seiple, P. Smith, C. Thomas, D. Towner, T. Troeger, C. Weber, P. Yashar, K. Zawadzki, and K. Mistry, "A 22nm high performance and low-power CMOS technology featuring fully-depleted tri-gate transistors, self-aligned contacts and high density MIM capacitors," in *Symposium on VLSI Technology*, 2012, pp. 131–132.

[44] S. Natarajan, M. Agostinelli, S. Akbar, M. Bost, A. Bowonder, V. Chikarmane, S. Chouksey, A. Dasgupta, K. Fischer, Q. Fu, T. Ghani, M. Giles, S. Govindaraju, R. Grover, W. Han, D. Hanken, E. Haralson, M. Haran, M. Heckscher, R. Heussner, P. Jain, R. James, R. Jhaveri, I. Jin, H. Kam, E. Karl, C. Kenyon, M. Liu, Y. Luo, R. Mehandru, S. Morarka, L. Neiberg, P. Packan, A. Paliwal, C. Parker, P. Patel, R. Patel, C. Pelto, L. Pipes, P. Plekhanov, M. Prince, S. Rajamani, J. Sandford, B. Sell, S. Sivakumar, P. Smith, B. Song,

K. Tone, T. Troeger, J. Wiedemer, M. Yang, and K. Zhang, "A 14nm logic technology featuring 2nd-generation FinFET, air-gapped interconnects, self-aligned double patterning and a 0.0588 $\mu$m2SRAM cell size," in *IEEE International Electron Devices Meeting (IEDM)*, 2014, pp. 3.7.1–3.7.3.

[45] N. Agrawal, Y. Kimura, R. Arghavani, and S. Datta, "Impact of Transistor Architecture (Bulk Planar, Trigate on Bulk, Ultrathin-Body Planar SOI) and Material (Silicon or III-V Semiconductor) on Variation for Logic and SRAM Applications," *IEEE Transactions on Electron Devices*, vol. 60, no. 10, pp. 3298–3304, 2013.

[46] X. Zhao and J. A. Del Alamo, "Nanometer-scale vertical-sidewall reactive ion etching of InGaAs for 3-D III-V MOSFETs," *IEEE Electron Device Letters*, vol. 35, no. 5, pp. 521–523, 2014.

[47] X. Zhao, J. Lin, C. Heidelberger, E. A. Fitzgerald, and J. A. del Alamo, "Vertical nanowire InGaAs MOSFETs fabricated by a top-down approach," in *IEEE International Electron Devices Meeting (IEDM)*, 2013, pp. 28.4.1–28.4.4.

[48] R. Baker, I. of Electrical, E. Engineers, and I. S.-S. C. Society, *CMOS: Circuit Design, Layout, and Simulation*, ser. IEEE Press Series on Microelectronic Systems. Wiley, 2008, no. v. 1. [Online]. Available: https://books.google.co.uk/books?id=N0XgLh2d2pkC

[49] W. Lu, A. Guo, A. Vardi, and J. A. del Alamo, "A Test Structure to Characterize Nano-Scale Ohmic Contacts in III-V MOSFETs," *IEEE Electron Device Letters*, vol. 35, no. 2, pp. 178–180, 2014.

[50] A. Dimoulas, A. Toriumi, and S. E. Mohney, "Source and Drain Contacts for Germanium and IIIâĂŞV FETs for Digital Logic," *MRS Bull.*, vol. 34, no. 07, pp. 522–529, 2009.

[51] M. Wittmer, R. Pretorius, J. Mayer, and M.-A. Nicolet, "Investigation of the au-ge-ni system used for alloyed contacts to gaas," *Solid-State Electronics*, vol. 20, no. 5, pp. 433 – 439, 1977.

[52] J. Bardeen, "Surface states and rectification at a metal semi-conductor contact," *Physical Review*, vol. 71, pp. 717–727, 1947.

[53] H. H. Wieder, "Surface and interface barriers of $In_xGa_{1x}As$ binary and ternary alloys," *Journal of Vacuum Science & Technology B*, vol. 21, no. 4, pp. 1915–1919, 2003.

[54] S. Kim, M. Yokoyama, N. Taoka, R. Iida, S. Lee, R. Nakane, Y. Urabe, N. Miyata, T. Yasuda, H. Yamada, N. Fukuhara, M. Hata, M. Takenaka, and S. Takagi, "Self-Aligned Metal Source/Drain $In_xGa_{1-x}As$ n-MetalâĂŞOxideâĂŞSemiconductor Field-Effect Transistors Using NiâĂŞInGaAs Alloy," *Applied Physics Express*, vol. 4, no. 2, p. 024201, 2011.

[55] S. Kim, M. Yokoyama, N. Taoka, R. Nakane, T. Yasuda, O. Ichikawa, N. Fukuhara, M. Hata, M. Takenaka, and S. Takagi, "$In_{0.53}Ga_{0.47}As$ metal-oxide-semiconductor field-effect transistors with self-aligned metal source/drain using Co-InGaAs alloys," *Applied Physics Letters*, vol. 100, no. 7, p. 073504, 2012.

[56] E. Y.-J. Kong, Ivana, X. Zhang, Q. Zhou, J. Pan, Z. Zhang, and Y.-C. Yeo, "Investigation of pdâĂŞingaas for the formation of self-aligned source/drain contacts in ingaas metalâĂŞoxideâĂŞsemiconductor field-effect transistors," *Solid-State Electronics*, vol. 85, pp. 36 – 42, 2013.

[57] U. Singisetti, M. A. Wistey, J. D. Zimmerman, B. J. Thibeault, M. J. W. Rodwell, A. C. Gossard, and S. R. Bank, "Ultralow resistance in situ Ohmic contacts to InGaAs/InP," *Applied Physics Letters*, vol. 93, no. 18, pp. 2–4, 2008.

[58] A. Vardi, L. Kong, W. Lu, X. Cai, X. Zhao, J. Grajal, and J. A. del Alamo, "Self-aligned InGaAs FinFETs with 5-nm fin-width and 5-nm gate-contact separation," in *IEEE International Electron Devices Meeting (IEDM)*, 2017, pp. 17.6.1–17.6.4.

[59] Q. Li and K. M. Lau, "Epitaxial growth of highly mismatched iii-v materials on (001) silicon for electronics and optoelectronics," *Progress in Crystal Growth and Characterization of Materials*, vol. 63, no. 4, pp. 105 – 120, 2017.

[60] K. Volz, A. Beyer, W. Witte, J. Ohlmann, I. Nmeth, B. Kunert, and W. Stolz, "GaP-nucleation on exact Si (001) substrates for III/V device integration," *Journal of Crystal Growth*, vol. 315, no. 1, pp. 37–47, 2011.

[61] N. Waldron, C. Merckling, W. Guo, P. Ong, L. Teugels, S. Ansar, D. Tsvetanova, F. Sebaai, D. H. van Dorp, A. Milenin, D. Lin, L. Nyns, J. Mitard, A. Pourghaderi, B. Douhard, O. Richard, H. Bender, G. Boccardi, M. Caymax, M. Heyns, W. Vandervorst, K. Barla, N. Collaert, and A. V. Thean, "An InGaAs/InP quantum well finfet using the replacement fin process integrated in an RMG flow on 300mm Si substrates," in *Symposium on VLSI Technology*, 2014, pp. 1–2.

[62] L. Czornomaz, E. Uccelli, M. Sousa, V. Deshpande, V. Djara, D. Caimi, M. D. Rossell, R. Erni, and J. Fompeyrine, "Confined Epitaxial Lateral Overgrowth (CELO): A novel concept for scalable integration of CMOS-compatible InGaAs-on-insulator MOSFETs on large-area Si substrates," in *Symposium on VLSI Technology*, 2015, pp. T172–T173.

[63] J. Lin, L. Czornomaz, N. Daix, D. A. Antoniadis, and J. A. del Alamo, "Ultrathin Body InGaAs MOSFETs on III-V-On-Insulator Integrated With Silicon Active Substrate (III-V-OIAS)," *IEEE Transactions on Electron Devices*, vol. 63, no. 8, pp. 3088–3095, 2016.

[64] L. Czornomaz, N. Daix, D. Caimi, M. Sousa, R. Erni, M. D. Rossell, M. El-Kazzi, C. Rossel, C. Marchiori, E. Uccelli, M. Richter, H. Siegwart, and J. Fompeyrine, "An integration path for gate-first UTB III-V-on-insulator MOSFETs with silicon, using direct wafer bonding and donor wafer recycling," in *International Electron Devices Meeting (IEDM)*, 2012, pp. 23.4.1–23.4.4.

[65] G. Doornbos and M. Passlack, "Benchmarking of IIIV n-MOSFET maturity and feasibility for future CMOS," *IEEE Electron Device Letters*, vol. 31, no. 10, pp. 1110–1112, 2010.

[66] N. Waldron, C. Merckling, L. Teugels, P. Ong, S. A. U. Ibrahim, F. Sebaai, A. Pourghaderi, K. Barla, N. Collaert, and A. V. Thean, "InGaAs Gate-All-Around Nanowire Devices on 300mm Si Substrates," *IEEE Electron Device Letters*, vol. 35, no. 11, pp. 1097–1099, 2014.

[67] C. B. Zota, F. LindelÃűw, L. Wernersson, and E. Lind, "InGaAs nanowire MOSFETs with $I_{ON}$= 555 $\mu$A/$\mu$m at $I_{OFF}$= 100 nA/$\mu$m and $V_{DD}$= 0.5 V," in *IEEE Symposium on VLSI Technology*, 2016, pp. 1–2.

[68] S. Lee, V. Chobpattana, C. . Huang, B. J. Thibeault, W. Mitchell, S. Stemmer, A. C. Gossard, and M. J. W. Rodwell, "Record $I_{on}$(0.50 mA/$\mu$m at $V_{DD}$= 0.5 V and $I_{off}$= 100 nA/$\mu$m) 25 nm-gate-length $ZrO_2$/InAs/InAlAs MOSFETs," in *Symposium on VLSI Technology*, 2014, pp. 1–2.

[69] T. . Kim, D. . Kim, D. H. Koh, H. M. Kwon, R. H. Baek, D. Veksler, C. Huffman, K. Matthews, S. Oktyabrsky, A. Greene, Y. Ohsawa, A. Ko, H. Nakajima, M. Takahashi, T. Nishizuka, H. Ohtake, S. K. Banerjee, S. H. Shin, D. . Ko, C. Kang, D. Gilmer, R. J. W. Hill, W. Maszara, C. Hobbs, and P. D. Kirsch, "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 *IEEE International Electron Devices Meeting (IEDM)*, 2013, pp. 16.3.1–16.3.4.

[70] T.-W. Kim, D.-h. Koh, C.-s. Shin, W.-k. Park, T. Orzali, C. Hobbs, W. P. Maszara, and D.-h. Kim, "Lg=80-nm Trigate Quantum-Well $In_{0.53}Ga_{0.47}As$ Metal Oxide Semiconductor Field-Effect Transistors With $Al_2O_3$," *IEEE Electron Device Letters*, vol. 36, no. 3, pp. 223–225, 2015.

[71] M. Radosavljevic, G. Dewey, D. Basu, J. Boardman, B. Chu-Kung, J. M. Fastenau, S. Kabehie, J. Kavalieros, V. Le, W. K. Liu, D. Lubyshev, M. Metz, K. Millard, N. Mukherjee, L. Pan, R. Pillarisetty, W. Rachmady, U. Shah, H. W. Then, and R. Chau, "Electrostatics improvement in 3-D tri-gate over ultra-thin body planar InGaAs quantum well field effect transistors with high-K gate dielectric and scaled gate-to-drain/gate-to-source separation," in *IEEE International Electron Devices Meeting (IEDM)*, 2011, pp. 33.1.1–33.1.4.

[72] J. Lin, X. Zhao, T. Yu, D. A. Antoniadis, and J. A. del Alamo, "A new self-aligned quantum-well MOSFET architecture fabricated by a scalable tight-pitch process," in *IEEE International Electron Devices Meeting (IEDM)*, 2013, pp. 16.2.1–16.2.4.

[73] X. Sun, C. D'Emic, C. . Cheng, A. Majumdar, Y. Sun, E. Cartier, R. L. Bruce, M. Frank, H. Miyazoe, K. . Shiu, S. Lee, J. Rozen, J. Patel, T. Ando, W. . Song, M. Lofaro, M. Krishnan, B. Obrodovic, K. . Lee, H. Tsai, W. . Wang, W. Spratt, K. Chan, S. Lee, J. . Yau, P. Hashemi, M. Khojasteh, M. Cantoro, J. Ott, T. Rakshit, Y. Zhu, D. Sadana, C. . Yeh, V. Narayanan, R. T. Mo, Y. . Heo, D. . Kim, M. S. Rodder, and E. Leobandung, "High performance and low leakage current InGaAs-on-silicon FinFETs with 20 nm gate length," in *2017 Symposium on VLSI Technology*, 2017, pp. T40–T41.

[74] V. Djara, V. Deshpande, E. Uccelli, N. Daix, D. Caimi, C. Rossel, M. Sousa, H. Siegwart, C. Marchiori, J. M. Hartmann, K. . Shiu, C. . Weng, M. Krishnan, M. Lofaro, R. Steiner, D. Sadana, D. Lubyshev, A. Liu, L. Czornomaz, and J. Fompeyrine, "An InGaAs on Si platform for CMOS with 200 mm InGaAs-OI substrate, gate-first, replacement gate planar and FinFETs down to 120 nm contact pitch," in *Symposium on VLSI Technology*, 2015, pp. T176–T177.

[75] V. Djara, V. Deshpande, M. Sousa, D. Caimi, L. Czornomaz, and J. Fompeyrine, "CMOS-Compatible Replacement Metal Gate InGaAs-OI FinFET With $I_{ON} = 156\ \mu\text{A}/\mu\text{mat}V_{DD} =$ 0.5V and$I_{OFF} = 100\text{nA}/\mu\text{m}$," *IEEE Electron Device Letters*, vol. 37, no. 2, pp. 169–172, 2016.

[76] J. Lin, L. Czornomaz, N. Daix, D. A. Antoniadis, and J. A. del Alamo, "Ultra-thin-body self-aligned InGaAs MOSFETs on insulator (III-V-O-I) by a tight-pitch process," in *72nd Device Research Conference (DRC)*, 2014, pp. 217–218.

[77] S. H. Kim, M. Yokoyama, R. Nakane, O. Ichikawa, T. Osada, M. Hata, M. Takenaka, and S. Takagi, "High performance extremely-thin body InAs-on-insulator MOSFETs on Si with Ni-InGaAs metal S/D by contact resistance reduction technology," in *Symposium on VLSI Technology*, 2013, pp. T52–T53.

[78] M. Berg, K. Persson, O. Kilpi, J. Svensson, E. Lind, and L. Wernersson, "Self-aligned, gate-last process for vertical InAs nanowire MOSFETs on Si," in *IEEE International Electron Devices Meeting (IEDM)*, 2015, pp. 31.2.1–31.2.4.

[79] C. B. Zota, L. Wernersson, and E. Lind, "Single suspended InGaAs nanowire MOSFETs," in *IEEE International Electron Devices Meeting (IEDM)*, 2015, pp. 31.4.1–31.4.4.

[80] J. J. Gu, X. W. Wang, J. Shao, A. T. Neal, M. J. Manfra, R. G. Gordon, and P. D. Ye, "III-V gate-all-around nanowire MOSFET process technology: From 3D to 4D," in *International Electron Devices Meeting (IEDM)*, 2012, pp. 23.7.1–23.7.4.

[81] S. Ramesh, T. Ivanov, E. Camerotto, N. Sun, J. Franco, A. Sibaja-Hernandez, R. Rooyackers, A. Alian, J. Loo, A. Veloso, A. Milenin, D. Lin, P. Favia, H. Bender, N. Collaert, A. V. Thean, and K. D. Meyer, "Top-down ingaas nanowire and fin vertical fets with record performance," in *IEEE Symposium on VLSI Technology*, 2016, pp. 1–2.

[82] C. Thelander, L. E. FröbergFroberg, C. Rehnstedt, L. Samuelson, and L. E. Wernersson, "Vertical enhancement-mode InAs nanowire field-effect transistor with 50-nm wrap gate," *IEEE Electron Device Letters*, vol. 29, no. 3, pp. 206–208, 2008.

[83] K. Tomioka, M. Yoshimura, and T. Fukui, "A III-V nanowire channel on silicon for high-performance vertical transistors," *Nature*, vol. 488, no. 7410, pp. 189–192, 2012.

[84] K. Tomioka, J. Motohisa, and T. Fukui, "III-V nanowire channel on Si: From high-performance vertical FET to steep-slope devices," in *International Symposium on VLSI Technology*, 2015, pp. 1–2.

[85] S. W. Chang, X. Li, R. Oxland, S. W. Wang, C. H. Wang, R. Contreras-Guerrero, K. K. Bhuwalka, G. Doornbos, T. Vasen, M. C. Holland, G. Vellianitis, M. J. H. van Dal, B. Duriez, M. Edirisooriya, J. S. Rojas-Ramirez, P. Ramvall, S. Thoms, U. Peralagu, C. H. Hsieh, Y. S. Chang, K. M. Yin, E. Lind, L. . Wernersson, R. Droopad, I. Thayne, M. Passlack, and C. H. Diaz, "InAs N-MOSFETs with record performance of $I_{on}$= 600 $\mu$A/$\mu$m at $I_{off}$= 100 nA/$\mu$m ($V_d$= 0.5 V)," in *IEEE International Electron Devices Meeting (IEDM)*, 2013, pp. 16.1.1–16.1.4.

[86] K. M. Persson, M. Berg, M. B. Borg, J. Wu, S. Johansson, J. Svensson, K. Jansson, E. Lind, and L. E. Wernersson, "Extrinsic and intrinsic performance of vertical InAs nanowire MOSFETs on Si substrates," *IEEE Transactions Electron Devices*, vol. 60, no. 9, pp. 2761–2767, 2013.

[87] B. Obradovic, T. Rakshit, W.-E. Wang, D. Lin, N. Waldron, N. Collaert, and M. S. Rodder, "Parasitic Bipolar Leakage in III-V FETs: Impact of Substrate Architecture," *arXiv*, vol. eprint arXiv:1705.06731, pp. 1–8, 2017.

[88] M. Radosavljevic, T. Ashley, A. Andreev, S. D. Coomber, G. Dewey, M. T. Emeny, M. Fearn, D. G. Hayes, K. P. Hilton, M. K. Hudait, R. Jefferies, T. Martin, R. Pillarisetty, W. Rachmady, T. Rakshit, S. J. Smith, M. J. Uren, D. J. Wallis, P. J. Wilding, and R. Chau, "High-performance 40nm gate length InSb p-channel compressively strained quantum well field effect transistors for low-power (VCC=0.5V) logic applications," in *IEEE International Electron Devices Meeting (IEDM)*, 2008, pp. 1–4.

[89] A. Nainani, T. Irisawa, Z. Yuan, Y. Sun, T. Krishnamohan, M. Reason, B. R. Bennett, J. B. Boos, M. G. Ancona, Y. Nishi, and K. C. Saraswat, "Development of high-k dielectric for

antimonides and a sub 350°c iiiâĂŞv pmosfet outperforming germanium," in *International Electron Devices Meeting (IEDM)*, 2010, pp. 6.4.1–6.4.4.

[90] K. Takei, M. Madsen, H. Fang, R. Kapadia, S. Chuang, H. S. Kim, C.-H. Liu, E. Plis, J. Nah, S. Krishna, Y.-L. Chueh, J. Guo, and A. Javey, "Nanoscale InGaSb Heterostructure Membranes on Si Substrates for High Hole Mobility Transistors," *Nano Letters*, vol. 12, no. 4, pp. 2060–2066, 2012.

[91] W. Lu, J. K. Kim, J. F. Klem, S. D. Hawkins, and J. A. del Alamo, "An ingasb p-channel finfet," in *IEEE International Electron Devices Meeting (IEDM)*, 2015, pp. 31.6.1–31.6.4.

[92] W. Lu, I. P. Roh, D. . Geum, S. . Kim, J. D. Song, L. Kong, and J. A. del Alamo, "10-nm Fin-width InGaSb p-channel self-aligned FinFETs using antimonide-compatible digital etch," in *IEEE International Electron Devices Meeting (IEDM)*, 2017, pp. 17.7.1–17.7.4.

[93] 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.

# Chapter 3

# Fabrication

## 3.1 Overview

This chapter provides an overview of the nanofabrication techniques relevant to the work presented in this thesis, namely: wafer growth of III-V materials, the patterning of samples via lithographic techniques, pattern transfer via wet/dry etching, and metal and thin film deposition. The description of each of the techniques is not exhaustive, but rather it is intended to provide the reader with sufficient detail such that the fabrication processes presented in later chapters can be fully understood. Particular focus is given to electron-beam lithography (EBL) and dry etch, which proved the most challenging due to the nature of the devices fabricated.

## 3.2 Molecular Beam Epitaxy

Molecular beam epitaxy (MBE) is a technique for growing thin films of crystalline material with ultra high purity and sub-nm precision [1]. The working principal of MBE is as follows: individual elements are heated in separate effusion cells until they start to sublime. The gaseous elements, or, 'molecular beam', emitted from each effusion cell condenses onto a heated substrate, where they react with each other, and, under the right conditions, form crystalline epitaxial monolayers of the desired material. In the growth of GaAs for example, solid sources of gallium and arsenic reside in their own individual effusion cells, and the emitted flux of gaseous Ga and As elements react on the sample surface to form single crystal GaAs. The thickness of the grown layer can be precisely controlled, down to individual monolayers of atoms, by computer controlled shutters which either block or permit the molecular beam emitted from each effusion cell. Compounds can be grown with as many elements as there are effusion cells, and the composition of the compound precisely controlled by controlling the flux of each molecular beam via the electronic shutters. Dopant elements can also be introduced in this manner. A simplified schematic of an MBE tool is shown in Fig. 3.1, illustrating the above. MBE achieves ultra-high purity growth due to the combination of not using carrier gasses to transport the desired elements to the substrate, and the ultra high-vacuum growth environment; which, aside from being desirable to minimise contamination, is critical to ensure that the elements emitted from the effusion cells have sufficiently long mean free paths such that they do not interact with each other until they reach the substrate.

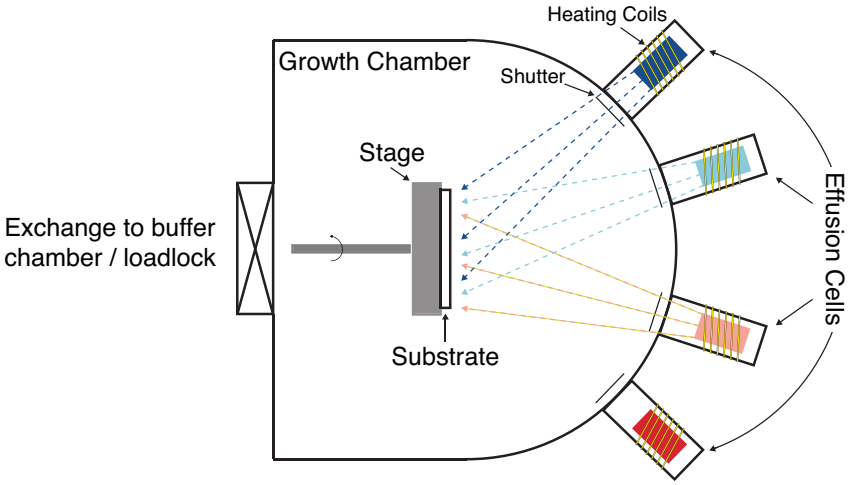


Figure 3.1: Schematic diagram of a molecular beam epitaxy (MBE) tool.

The main drawback of MBE is throughput: the advantage of precision afforded by a layer by layer growth mechanism is counter balanced by slow growth rates which limit the productivity. The other most commonly used growth technique for III-V semiconductors is MOCVD, which offers faster growth rates at the expense of precission. All material used for the work presented

in this thesis was grown by MBE.

## 3.3 Lithography Overview

In the context of nanofabrication, lithography refers to the process of defining a pattern in a thin film, named a resist, which has been deposited on to the surface of a sample. The defined pattern serves to either mask or expose desired regions of the sample to subsequent processing steps, such as etching to isolate a device or metal deposition to form contacts for example. There are a variety of lithography techniques, each of which have advantages and disadvantages in terms of resolution, throughput and cost, however in this work the discussion is limited to photolithography and EBL.

### 3.3.1 Resists

Resists are radiation sensitive polymers whose solubility in certain solvents changes subsequent to irradiation. As such, arbitrary patterns can be defined by irradiating selected regions of a resist coated sample and subsequently developing in the corresponding solvent. Resists can be categorised as either positive tone or negative tone, where their solubility increases or decreases upon irradiation respectively. Figure. 3.3(a-c) depicts the pattern yielded in a positive and negative tone resist from a given identical exposure. Also shown are subsequent metal deposition (d) and etching (e) processes which illustrate common ways in which lithographic patterning, and specifically both positive and negative tone resists, are utilised as part of nanofabrication processes.

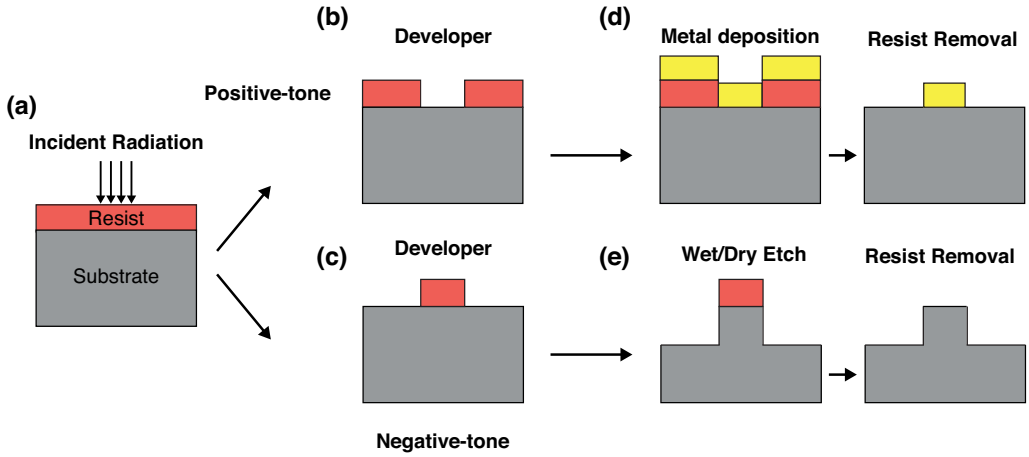


Figure 3.2: A schematic diagram illustrating the difference between positive and negative tone resits and the subsequent processing steps that they are commonly used for: lift-off and etching.

Depositing metal onto a sample using the process depicted in Fig. 3.2 is known as ‘lift off’. A successful lift-off process is critically dependent on there being a discontinuity in the deposited metal film surrounding the patterned feature, through which solvent can pass to dissolve the

resist and lift-off the unwanted metal. This can be achieved by purposefully altering the sidewall profile of the resit in order to minimise any unwanted metal sidewall coverage. This is commonly achieved by either employing a process which yields an undercut resist profile such as that shown in Fig. 3.3(a), or by utilising a bi-layer of resits which have different sensitivities, such as the case depicted in Fig. 3.3(b) for a positive tone resist. For the former, this can be achieved by a variety of processes including the use of negative tone resists which naturally undercut [2] and image reversal baking [2] for example. For the latter, the use of a more sensitive resist for the bottom layer results in a developed pattern which has a feature size fractionally larger than in the top layer, yielding a profile with an overhang.

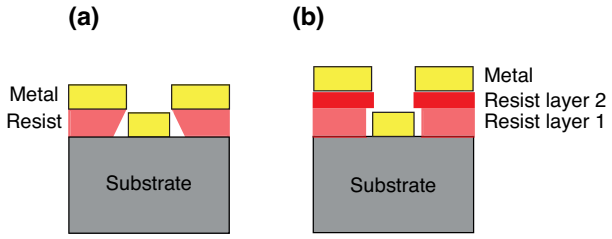


Figure 3.3: Suitable resist profiles for a successful lift-off process: (a) an undercut profile and (b) the use of a resist bi-layer to yield an overhanging profile.

In addition to utilising a process which yields favourable resist sidewalls profiles, it is also critical that the method of metal deposition is non-conformal, such is the case for metal deposition by ebeam evaporation, which is discussed in Section 3.9.

Given the lengthy discussion of the requirements for CMOS compatible process modules in the previous chapter, it is worth noting that while lift-off is commonly used in research environments and for III-V fabrication, it is not a CMOS compatible process. It was used however to form the S/D and gate, source and drain contacts of the junctionless FinFET devices discussed in Chapter 4 as a means to simplify the process. The results presented however are not critically dependant on this, and the process could easily be altered to be fully CMOS compatible.

### 3.3.2 Resist Application

Resist application is dependant on the specific resist used. Typically the sample is cleaned and de-greased prior to baking in order to remove residual moisture from the surface. A solution of resist is then dispensed onto the sample and spun at high speed to yield a conformal coating, before a final bake to evaporate the solvent and yield a solid film.

## 3.4 Photolithography

Photolithography utilises photosensitive resists which are patterned with ultraviolet light projected through a mask, as depicted in Fig. 3.4. This is a high throughput process as each layer of a pattern is written with a single exposure and is thus the lithographic technology utilised in industry for mass production. The smallest possible feature size that can be written however, referred to as the critical dimension (CD), is ultimately limited by the wavelength of the light source. Advancements have been made in reducing the wavelength of the illumination source as well as improving mask technology such that 50 nm features can be patterned, however this is an extremely costly process which is only economically viable in a mass production environment. In research laboratories, as is the case the in the James Watt Nanofabrication centre (JWNC), photolithography is commonly used for the patterning of large features (CD > 500 nm), while an alternative type of lithography such as EBL is used where smaller features are required.

The photolithographic tool used in this work was a Karl Suss Microtech MA6 Mask Aligner (MA6) which uses ultraviolet light as the illumination source and includes an i-line filter (365 nm wavelength) which facilitates the patterning of features with CD$\geq$ 500 nm.

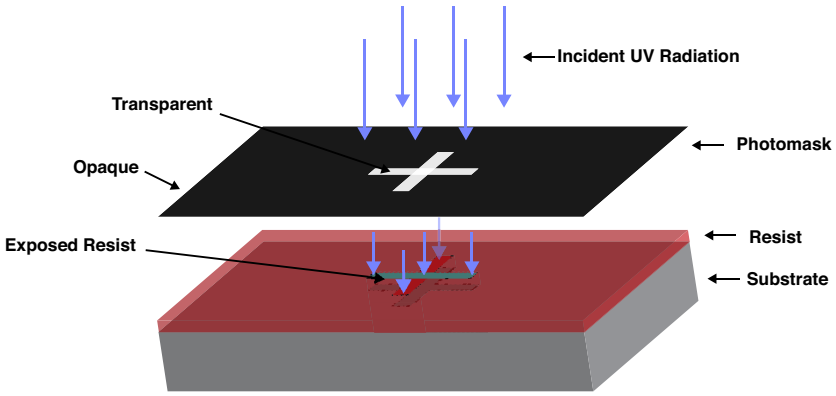


Figure 3.4: Illustration of the process of photolithography, where ultra-violet radiation is incident on a sample through a partially transparent mask.

## 3.5 Electron beam Lithography

Unlike photolithogrpahy, EBL is a maskless technology which selectively patterns desired regions of an electron sensitive resist by controllably deflecting a beam of electrons. Using this method, patterns can be written with CDs as small $\sim$10 nm. EBL tools exist in a variety of forms, and the following provides a basic overview of the tool at the JWNC: it's classification, how patterns are written, layer to layer alignment, and how pattern files are prepared.

### 3.5.1 Tool Classification and Basic Operation

The EBL tool at the JWNC is is a Vistec VB6 UHR EWF which utilises a Gaussian-beam and operates at 100 kV. Gaussian-beam tools focus the electron beam to as small as possible given the beam current, and patterns samples using a succession of point exposures, pixel by pixel. The intensity of the beam across the diameter of the focussed spot is Gaussian in shape, hence the name. The alternative option is a shaped-beam tool, which utilises a wide beam of uniform intensity that can be manipulated into a variety of trapezium shapes using a series of interchangeable apertures, and therefore each shape can patterned with a single exposure. As such Guassian beam tools suffer in terms of writing speed in comparison to shaped-beam, however they offer the ultimate resolution, and the VB6 at Glasgow can readily pattern features with CDs of $\sim$10 nm. A simplified schematic for for the VB6 is shown in Fig 3.5 (a). The key components are the electron gun which generates the electrons, the column which focusses the electron beam and controllably deflects it to the desired regions of the sample, and the stage where the sample is mounted under high vacuum. The column comprises a variety of electrostatic and electromagnetic lenses and apertures which serve to focus the beam and block electrons which have strayed significantly from the optical axis. The beam blankers are a pair of electrostatic plates that can quickly divert the beam from the optical axis and prevent electrons reaching the sample surface, effectively turning the beam on and off. The deflection coils are electromagnetic coils that controllably deflect the beam in order to expose the desired region of the sample. Aberrations in the electron beam increase with radial distance from the optical axis, and as such there is a limit in the extent to which the beam can be deflected to write a pattern, known as the main field size. The maximum main field size of the VB6 is 1.301072 mm. In order to write patterns which are larger than this, the stage can be moved in either x or y directions by linear stepper motors. To account for the inaccuracy of the movement of the motors, the stage is interferometrically monitored and the beam is deflected to negate the error in position; a process known as beam error feedback (BEF). The VB6 utilises vector scanning to write patterns, which is where the beam is deflected to the start of each shape individually. This is opposed to raster scanning tools which scan the beam across the entire field while blanking the beam selectively. For clarity, Fig. 3.5 (b) illustrates the patterning of 3 shapes across two main fields by means of vector scanning, beam blanking, beam deflection and stage movement with BEF. The process starts with the deflection of the blanked beam to the bottom left corner of shape 1. The beam blanker is subsequently turned off and the pattern written by deflecting the beam using the deflection coils as indicated by the solid blue arrow. Once the end of the shape is reached, the beam is deflected to the start of shape 2, during which time the beam is again blanked so as not to expose the resist between patterns. This is indicated by the blue dashed line. Once the beam has been deflected to the correct position, the beam blankers are turned off and Shape 2 is patterned as per Shape 1. The position of Shape 3 is out with Main Field 1 - beyond the point

which the beam can be maximally deflected - and therefore the position of the stage is moved by the linear stepper motors while the beam is blanked. This is indicated by the red dashed arrow. As shown, there is error in this process and the beam is deflected to the exacts start position of Shape 3 using BEF.

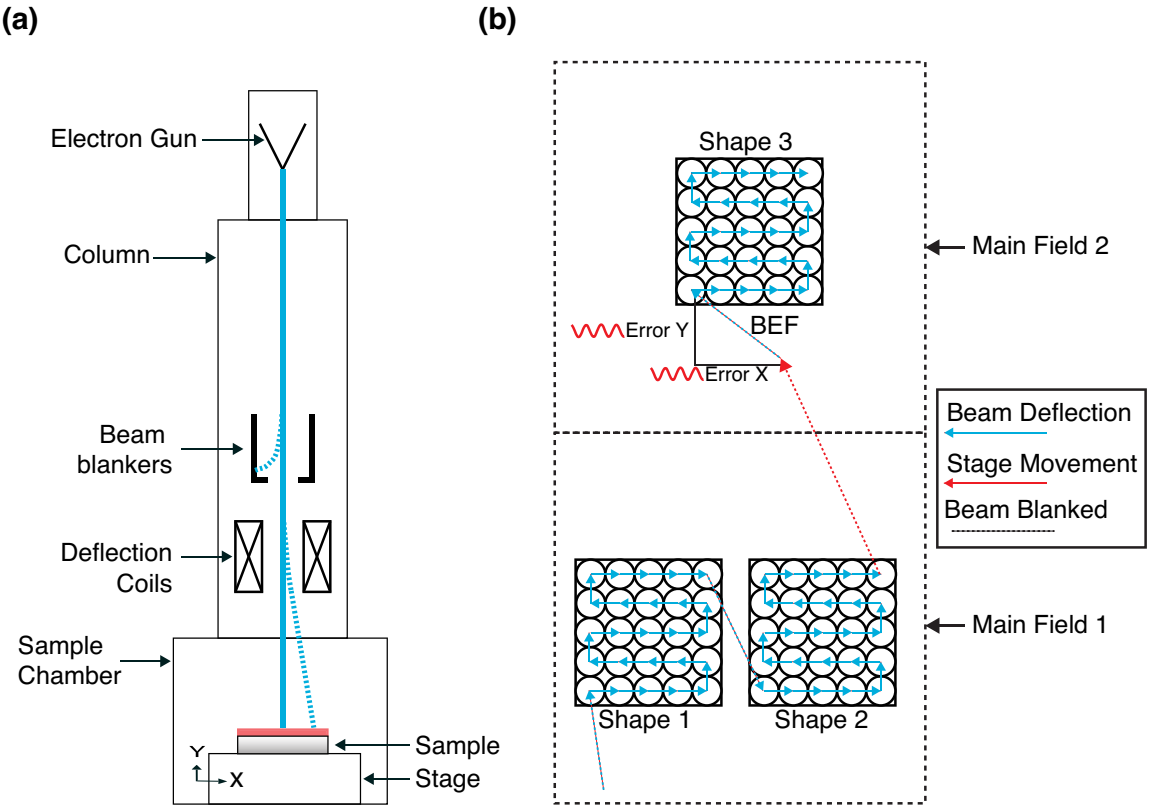


Figure 3.5: (a) Simplified schematic diagram of the Vistec VB6 UHR EWF Electron Beam Lithography tool. (b) Illustration of how a pattern is written via beam deflection, beam blanking, stage movement and beam error feedback.

### 3.5.2 Resist Exposure: Electron-Solid Interactions

Ideally, only the regions of resist that interacted with the incident beam of electrons would be exposed, and therefore arbitrary patterns could be written with minimum features corresponding to the size of the smallest electron beam diameter, which in the case of the VB6 is $\sim 4$ *nm*. In reality this not the case however due to scattering events that occur when electrons interact with matter. The two important scattering mechanisms with regards to EBL are Rutherford scattering and electron-electron scattering. Rutherford scattering is an elastic process which occurs when electrons interact with atomic nuclei, causing them to undergo large changes in direction while retaining their energy. The mean free path for Rutherford scattering increases proportionally with the square of electron energy [3] (and therefore beam voltage) and decreases proportionally

with atomic number [3]. Conversely, electron-electron scattering is an inelastic process which causes the electrons to lose energy and slow down [3]. The rate with which electrons lose energy as a function of distance due to this process is described by the continuous slowing down approximation of Bethe [4], which increases approximately linearly with material density and decreases inversely proportionally with electron energy [4]. It is this loss in energy that exposes the resist and therefore lower energy electrons contribute more to resist exposure than do higher energy. The consequence of these scattering processes is depicted in Fig. 3.6. The thickness of resist for typical high resolution EBL exposures is smaller than the mean free path for Rutherford scattering in the resist (the elastic scattering mean free path in PMMA at 100 kV is 250 nm for example) and therefore the majority of electrons pass through the resist undeflected. The inelastic scattering mean free path is much shorter however and electron-electron scattering events occur which slow the electrons down and generates secondary electrons. These secondary electrons typically have low energies (<200 eV) and it is therefore these electrons which are predominantly responsible for resist exposure. Due to their low energies they have a short range (< 5 nm) however this can serve to effectively increase the spot size. This increase in beam diameter in nanometers, $d_{\mathrm{f}}$, is given empirically by [5]

$$d_f = 0.9(\frac{R_t}{V_b})^{\frac{3}{2}} \tag{3.1}$$

where $R_{\mathrm{f}}$ is thickness of resist in nm and $V_{\mathrm{b}}$ is the beam voltage in kV. The 100 kV beam voltage of the VB6 makes this negligible in most circumstances however.

All electrons that penetrate the resist undergo multiple large angle Rutherford scattering events, which results in some electrons undergoing deflection angles of $> 180^{\circ}$ and re-entering the resist some distance away from the initial point of incidence. These are known as backscattered electrons and they have lower energy than incident electrons due to electron-electron scattering events along the way. This results in what is known as the 'proximity effect', where unwanted exposure of resist occurs within the backscattering radius. The region labelled as the 'interaction volume' in Fig. 3.6 illustrates the depth into the sample where Rutherford scattering starts to dominate and the majority of electrons undergo large deflections. The closer this interaction volume is to the surface of the sample, the smaller the backscattering radius and the higher the intensity of exposure immediately surrounding the patterned feature. The interaction volume is pushed further into the substrate with increasing beam voltage and decreasing atomic weight. III-V compounds have a higher atomic weight than Si, and as a consequence they suffer more from the proximity effect. This can be seen from Fig. 3.7 which compares the calculated deposited energy as a function of radial distance from a given point exposure, known as a point spread function (PSF), for. both a Si substrate and a III-V wafer[1].

[1]The layer structure of the III-V wafer in this example is that of the Junctionles FinFET wafer discussed in

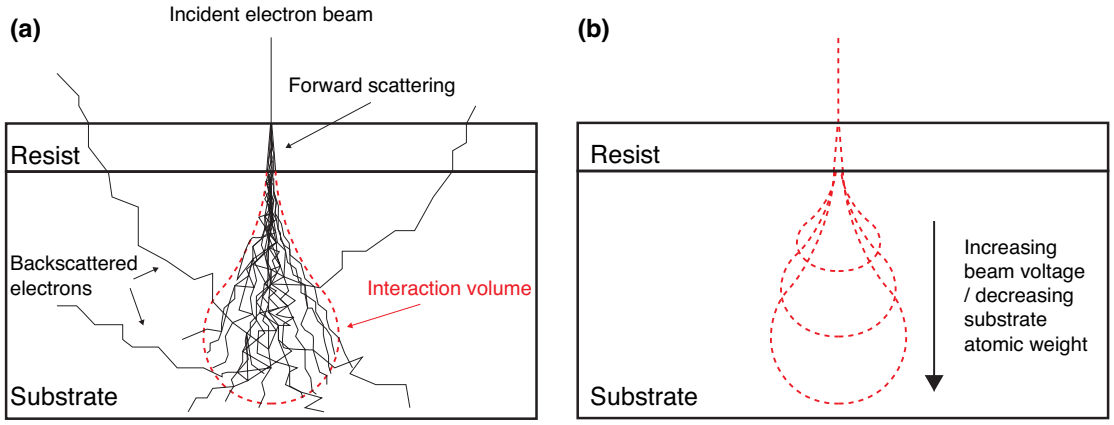


Figure 3.6: (a) Illustration of the effects of electron-solid interactions. (b) The impact of beam voltage and substrate atomic weight on the interaction volume for electron-solid interactions.

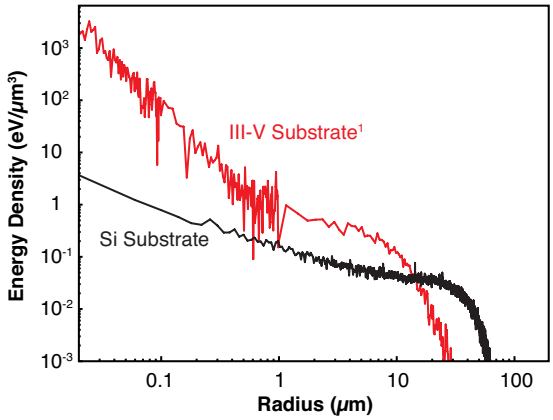


Figure 3.7: Calculated point spread functions (PSF) for Si (black) and III-V (red) substrates.

The effects of the above scattering processes place limitations on what can be patterned by EBL, with the most limiting mechanism dependant on the specific pattern. For example, if one were to write a narrow isolated line, the dominant mechanism is forward scattering, which serves to broaden the beam diameter and limit the minimum CD.

For patterns which are non-isolated or larger than the backscattered radius, the proximity effect can result in the unwanted exposure of regions between shapes or the distortion of individual shapes.

### 3.5.3 Proximity Effect Correction

The proximity effect can be accounted for by patterning shapes with a non uniform dose such that the combination of the total forward and backscattered doses is equal at all points. The

Chapter 4. The PSFs were calculated using Monte Carlo simulations

total dose received at a given a distance from the point of exposure can be approximated by the following double Gaussian function [5]

$$PSF(r) = \frac{1}{\pi(1+\eta)}[\frac{1}{\alpha^2}exp(\frac{-r^2}{\alpha^2}) + \frac{\eta}{\beta^2}exp(\frac{-r^2}{\beta^2})] \tag{3.2}$$

where $\alpha$ is the radius of the forward scattered beam, $\beta$ is the radius of the backscattered beam and $\eta = \alpha/\beta$. $\alpha$ and $\beta$ can be calculated for a given resist and substrate layer structure by fitting Eqn. 3.2 to the PSFs determined from Monte Carlo simulation.

### 3.5.4 Writing Parameters

The beam deflection is controlled by analogue signals output from the pattern generator which are generated by a digital to analogue converter (DAC). The minimum amount by which the beam can be deflected is dependant on the resolution of the DAC, which in the case of the VB6 is 20 bit. This corresponds to $2^{20}$ distinct positions which can be defined along each of the x and y axes over the size of one main field. In the case of a field size of 1.301072 mm this gives a minimum beam deflection, known as the beam step size (BSS) of 1.25 nm. The field size of the VB6 can be reduced in order to decrease the BSS, however this is a timely process which requires lengthy calibration. As such, the the VB6 is set up with the three default field sizes which correspond to BSS of 0.5 nm, 1 nm and 1.5 nm.

The variable resolution unit (VRU) allows the user to change increase the minimum amount by which the beam is deflected, while retaining placement accuracy of the smallest BSS. This is often to desirable in order to achieve the correct dose while staying within the limits of the hardware, to increase the speed of writing a pattern which does not require a small BSS. VRU is an integer multiple of BSS and is set by the user along with area dose and beam current in order to define their pattern.

### 3.5.5 Layer to Layer Alignment

Often processes require more than one layer of lithography, and subsequent layers have to be accurately aligned relative to each other. The Vb6 aligns samples by identifying markers on the sample, which form a series of reference points that allow the tool to define a patterning grid. These markers can either be defined globally, or locally. Global markers are defined on the chip scale and can provide alignment on the order of 100 nm. Local markers are defined within each cell of the pattern (i.e. each 1.3 mm $\times$ 1.3 mm square), which means that the pattern is

effectively re-aligned following a stage movement. Local markers can allow for layer to layer alignment as low as 20 nm.

Local alignment markers can also be used to counter the effect of drift, which is where the lithography tool effectively loses track of the beams position with respect to the sample. This can be caused by thermal expansion of the materials amongst other effects, and can be reduced by local alignment markers as they force the tool to re-align for every local cell.

### 3.5.6 File preparation

Patterns for electron beam lithography are designed in a CAD program; L-edit used in this work. The final design is output as a GDSII file, and subsequently fractured into polygons and paths that can be written by the e-beam tool. The only shapes available in GDSII are polygons and paths. A path, or track, is a series of straight-line segments with a given width. Both polygons and paths have a maximum of 200 vertices, which seemed generous when the format was designed but is appears restrictive now. If you need more points you must subdivide your shape; for instance a circle could be coded as a series of pie wedges. As well as being fractured, proximity error corrections are applied as described earlier in the chapter. Finally, the VRU, dose, and beam-size are set before the pattern file is transferred to the Vb6, and the job is run.

## 3.6 Etching

In general terms, etching is the selective removal of material by either chemical or physical processes. Wet etching is where the sample is immersed in a solution that removes material purely through chemical reactions.

### 3.6.1 Wet Etching

In this work wet etching was only used for cleaning the surface of samples by selectively etching their native oxide. Wet etching can be used for pattern transfer however its isotropic nature makes it unsuitable for patterns such as highly scaled Fins; it is often used for much larger patterns where the sidewall profile is unimportant, such as in the definition of etched mesas.

## 3.7 Dry Etching

### 3.7.1 Reactive Ion Etching and Inductively Coupled Reactive Ion Etching

Unlike wet etching, anisotropic etching can be achieved via reactive ion etching (RIE), which is one such variant of a "Dry Etch" process, so-called as it is a "dry" plasma that etches the sample, as opposed to a wet liquid. A plasma is a partially ionised gas which contains equal numbers of positive ions and negative electrons in addition to non-ionised gas particles and neutral radicals. In the case of an RIE tool, the plasma is generated by applying radio frequency power to parallel electrodes immersed a gas or mixture of gases, which rips electrons from atoms due to the oscillating electric field, which in turn are accelerated into other atoms resulting in further ionisation.

A simplified schematic of an RIE etch tool is shown in Fig. 3.8. The selected gas(es) are pumped into the chamber and subsequently ionised due to the RF field applied between the parallel electrodes. Electrons in the plasma are able to follow the applied RF power however the more massive ions are not. This results in electrons being accelerated across the chamber and accumulating on the platter, which unlike the chamber walls and top electrode, is electrically isolated due a DC blocking capacitor. This results in the formation of an ion sheath, also referred to as dark space, between the platter and plasma, across which there is a self induced DC bias. This bias results in an electric field which accelerates ions directionally towards the sample. RIE is a combination of both physical (sputtering) and chemical etching mechanisms. Chemically inert

ions which have been accelerated by the DC self bias can eject atoms from the sample surface due to their energetic collision. This is a purely physical process and is anisotropic provided that the mean free path of the ion is longer than the thickness of the ion sheath. If this were not the case the ion would likely scatter within the sheath and collide with the sample surface at non-normal angle of incidence. Etching by sputtering is an extremely slow process and can damage the semiconductor by breaking bonds beneath the surface. Conversely if the gas(es) used are not inert, material can be removed from the sample by a chemical process where ions which have been accelerated by the DC self bias, or neutral radicals which have diffused across the ion sheath, adsorb onto the sample surface, react and form by products, and lastly the by-products desorb from the sample.

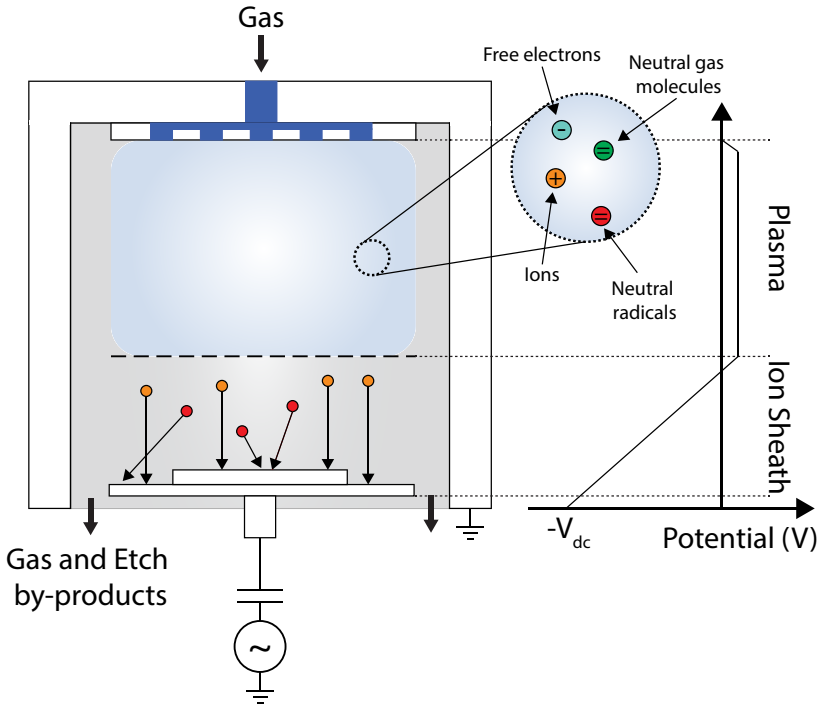


Figure 3.8: Schematic diagram of an RIE tool with the DC self bias shown.

This can also be an anisotropic process provided that the etch rate due to directionally accelerated ions is higher than the isotropic etching of the sample due to neutral radicals. A phenomena known as ion-assisted etching can be exploited to yield an anisotropic etch process, whereby the combination of sputtering due to ion bombardment and etching due to neutral radicals yields a significantly increased etch rate compared to either process individually. It is thought that this occurs due to ions locally increasing the temperature of regions of the sample, which in turn increases the reactivity with radicals.

Achieving an anisotropic etch profile can readily be achieved by increasing the RF power to in turn increase the DC self bias and thus increase the sputtering and ion assisted sputtering etch components. In the context of etching an InGaAs fin however, as discussed in Section 4.5, this

approach is not viable as the semiconductor is easily damaged due to sputtering which manifests as degradation in device performance. Where an etch process is constrained by requiring a low DC self bias, an anisotropic etch can be achieved by balancing isotropic chemical etching due to radicals with passivation either due to polymer deposition as a result of the gases used or non-volatile etch byproducts. This is approach is limited in an RIE system, where the plasma density (and therefore chemical reactivity) and DC self bias cannot be controlled independently, but are both proportional to the RF power. An inductively coupled plasma reactive ion etching (ICP-RIE) tool, the simplified schematic for which is shown in Fig. 3.9, overcomes this limitation: the plasma is generated by applying RF power to an inductive coil, whereas the platen is powered by a separate RF source.

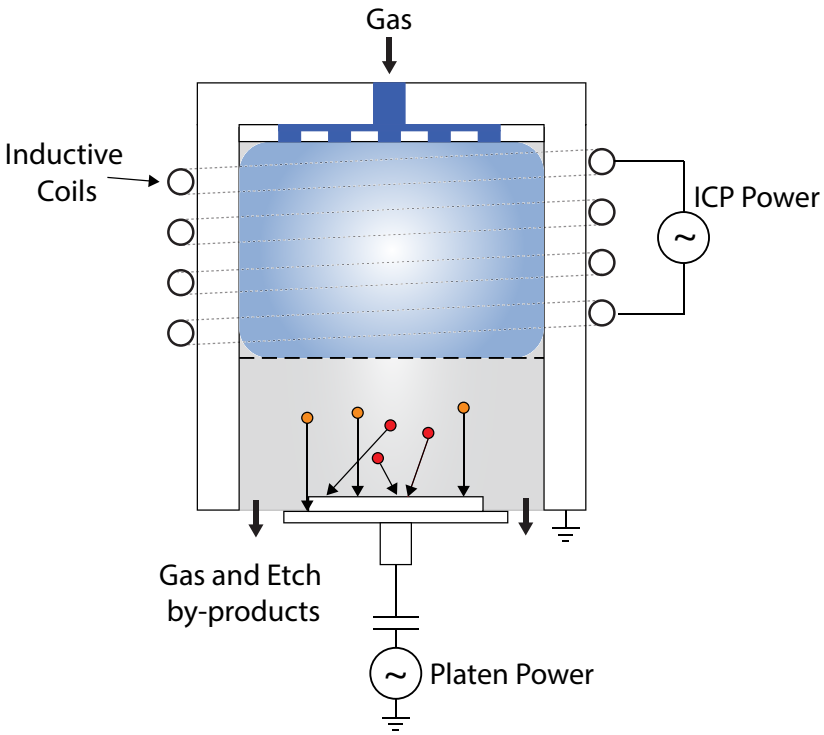


Figure 3.9: Schematic diagram of an ICP-RIE tool.

## 3.8 Atomic Layer Deposition

chemical vapour deposition (CVD) is a technique which deposits thin films on substrates by passing gaseous precursors over a heated sample which they react with or decompose on to form a solid polymer film. Atomic layer deposition is regarded as a subclass of CVD and it deposits thin films in a self-limiting cyclic manner which yields excellent control in the thickness of film deposited and also produces films which are extremely conformal - a requirement that is pivotal when forming gate dielectrics on non-planar transistors. To illustrate the process, Fig. 3.10 shows the 4 steps of depositing a mono layer of thermal $Al_2O_3$ an -OH terminated substrate. In step 1, the precursor TMA is pumped into the chamber and dissociatively chemisorbs onto the surface of the sample. When the sample is saturated the chamber is purged and any remaining TMA and byproducts are removed. Subsequently water vapour is pumped into the chamber which reacts with $CH_3$ which forms $CH_4$ as a by products and leaves a hydroxylated $Al_2O_3$, followed by another purge to remove any precursor and by-products. This process can then be repeated for an arbitrary number of cycles in order to deposit the desired film thickness.

ALD has proved an indispensable technology in producing scaled EOT gate stacks with low $D_{it}$ on III-V's.

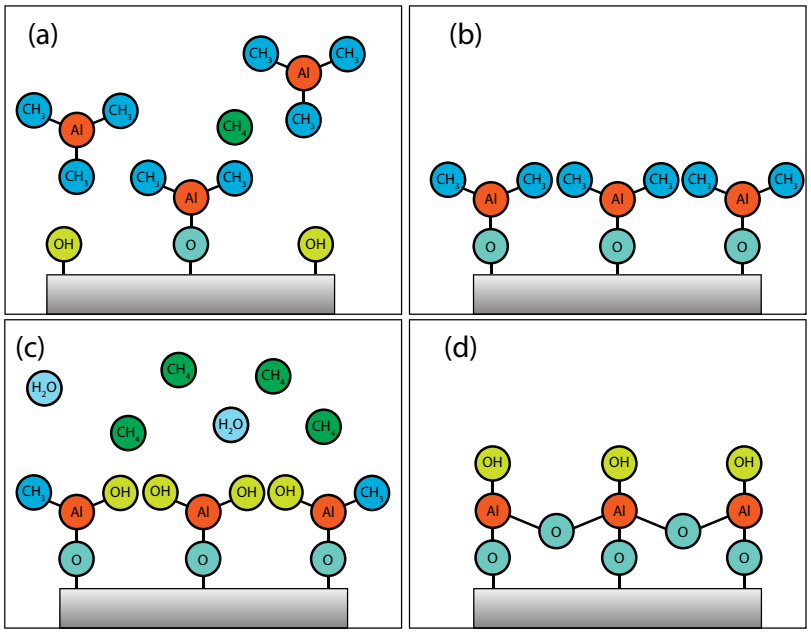


Figure 3.10: An ALD growth cycle schematically depicting $Al_2O_3$ deposition from TMA ($Al(CH_3)_3$) and $H_2O$ precursors on an OH-terminated substrate surface: (a) the pulsed TMA precursor is chemisorbed until the surface is saturated. Al ions bond with O, and in the process produce $CH_4$ as reaction by-products, (b) un-reacted TMA and $CH_4$ by-products are purged, leaving behind a $CH_3$-terminated surface, (c) pulsed $H_2O$ precursor is chemisorbed until the surface is saturated. Al-O bridges are formed and OH groups replace $CH_3$ ligands, with $CH_4$ generated as by-products and (d) un-reacted $H_2O$ and $CH_4$ by-products are purged, leaving behind an OH-terminated surface.

## 3.9 Metal deposition

Metal deposition was required to form alignment markers for EBL and source, drain and gate contacts for the FinFET process detailed in Chapter 4.3, as well as to deposit the gate metal and ohmic contacts to the MOS capacitors presented throughout Chapter 5. All of these were deposited by lift-off and therefore required a directional, non-conformal metal deposition technique. Electron beam (ebeam) evaporation achieves this by magnetically focussing an electron beam onto the surface of the desired source metal, which causes its temperature to increase until surface atoms vaporise, at which point they traverse the chamber and coat the surface of the sample. A non-conformal coating is achieved by depositing at low pressure ($1 \times 10^{-7} mbar$) such that the mean free path of evaporated particles is longer than the distance between source and sample, which prevents collisions from occurring and randomising the direction. The thickness of the deposited film is monitored in real time using a quartz crystal whose frequency of oscillation changes as metal is deposited onto it and increases its mass. A metal deposition run involves heating the metal source and monitoring the deposition rate via the crystal with the sample shutter closed until the desired rate is achieved. Subsequently the sample shutter is opened and the sample holder rotates to improve the uniformity of the deposition across the sample. Metal can be thermally evaporated using the same principle by directly heating the crucible, however ebeam evaporation is preferable as the molten source does not come into contact with the crucible, which minimises contamination. The ebeam evaporation tool used includes the capability to generate an Ar plasma in the chamber, which can be used for the in-situ etching of native oxides on the sample surface prior to metal deposition. This is utilised in Chapter 4 to etch the native oxide of the InGaAs surface prior to ohmic contact deposition, which serves to reduce contact resistance.

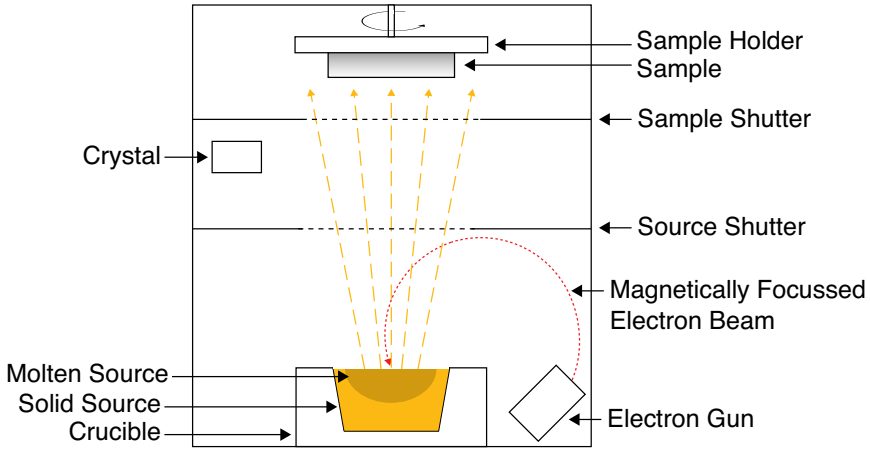


Figure 3.11: Schematic of an metal evaporation tool.

## 3.10 Cluster Tool

The work presented in Chapter 4 was entirely reliant on the ability to expose samples to a plasma in-situ prior high-k deposition. This was facilitated by the 'Cluster Tool' in the JWNC which comprises ICP-RIE, ICP-Deposition and ALD chambers connected via a central vacuum exchange chamber, which allows samples to be processed and transferred with exposure to air. This is critical for fabricating antimonide devices, which are highly reactive in air and rapidly oxidise. For this work, only the ICP-RIE and ALD chambers were used.

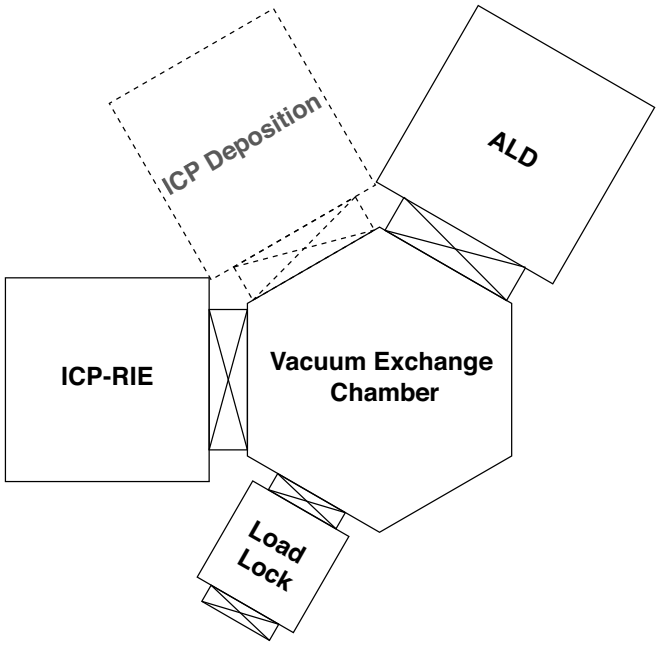


Figure 3.12: (a) Annotated image of the cluster tool and (b) the corresponding schematic showing the ICP-RIE, ICP deposition and ALD chambers which are interconnected via a central vacuum exchange chamber.

## 3.11 Thermal Processing

### 3.11.1 Rapid Thermal Annealing

Rapid Thermal Annealing (RTA) is the process of rapidly heating a sample via halogen lamps which can heat samples in excess of 1000 $^{\circ}C$ on a times scale of several seconds. Such high temperatures are not compatible with or required for III-V processing, however it's asset is the rate with which it can heat samples to a desired temperature, which is utilised in this work to anneal ohmic contacts for a sepcific, optimal time, which limits the diffusion of metal atoms and prevents degradation to the resistance of the contacts formed.

### 3.11.2 Forming Gas Annealing

A strip heater was utilised to anneal the samples discussed in Chapter 5 in forming gas. This simply comprises a resistive heating element which the sample sits atop in a forming gas (95% $N_2$: 5%$H_2$) purged environment.

# Bibliography

[1] M. A. Herman and H. Sitter, "Molecular beam epitaxy: Fundamentals and current status," 1996.

[2] H. Moritz, "Optical single layer lift-off process," *IEEE Transactions on Electron Devices*, vol. 32, pp. 672–676, Mar 1985.

[3] e. b. J. M. Stephen Thoms, ch. 2. first ed.

[4] R. Shimizu and D. Ze-Jun, "Monte Carlo modelling of electron-solid interactions," *Reports Prog. Phys.*, vol. 55, no. 4, pp. 487–531, 1992.

[5] M. A. McCord and e. b. P. R.-C. Michael J. Rooks, ch. 2. first ed.

# Chapter 4

# Results I - Junctionless InGaAs FinFETs

## 4.1 Background

### 4.1.1 The Junctionless Architecture

Conventional MOSFETs rely on p-n junctions between the channel and S/D regions to either block or permit current flow depending on the applied gate bias [1]. As these devices are scaled, extremely abrupt doping concentration gradients are required between junctions in order to define short channel lengths while retaining electrostatic integrity [1]. This places limitations on the thermal processing for device fabrication and necessitates costly millisecond annealing techniques in order to minimise the diffusion of dopant atoms [2]. Colinge *et al*. demonstrated a silicon nanowire device in 2010 which circumvents the above issues by employing the junctionless gated resistor architecture [2]: an architecture which was in fact first proposed by Lilienfeld in 1925 [3], however, its realisation has only been made possible by modern nano fabrication techniques. A schematic diagram of a junctionless device is shown in Fig. 4.1. It comprises a fin or nanowire of uniformly doped semiconductor (Si for the Colinge device [2]), without junctions, which is isolated by a buried oxide layer (BOX) or wide band gap semiconductor. Current flow can by modulated via a capacitively coupled gate, and the device can be turned off provided that the geometry of the fin or wire is sufficiently thin and narrow such that it can be fully depleted [2]. In the on-state, this device operates at, or slightly more accumulated than flat band, and is essentially a resistor with conductivity $\sigma$=$q\mu N_{\mathrm{D}}$. This has the consequence that the transverse electric field is much smaller (zero at flat band) than an inversion mode (IM) device at a given gate overdrive, and therefore carriers are not confined to the surface but are free to flow through the body of the device. This reduces the effects of surface roughness scattering

and minimises the mobility degradation from that of the bulk value [2]. Furthermore, it has been shown that due to their bulk conduction, the relative amplitude of random telegraph noise (RTN) of junctionless transistors is considerably less than their IM counterparts, as there are fewer carriers at the dielectric interface to be captured by trap states [4].

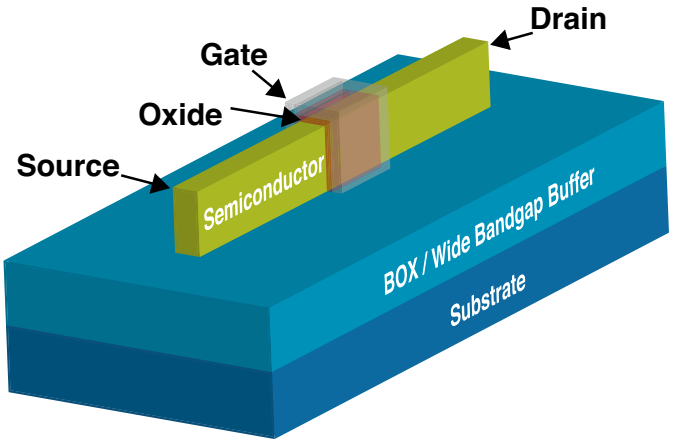


Figure 4.1: Schematic diagram of the junctionless metal oxide semiconductor field effect transistor (MOSFET) architecture.

Colinge proposes that junctionless devices could offer a reduced CV/I gate delay,$\tau$, compared to conventional MOSFETs and thus could operate at higher speeds [2]. In a IM MOSFET, the on-current can be increased by increasing $C_{ox}$, however, the increase in speed that would result due to an increased $I_{on}$ is exactly counterbalanced by the increased gate capacitance of the next device. In a junctionless device, $I_{on}$ can be increased independently of $C_{ox}$ by increasing the doping concentration, which Colinge proposes relieves the requirement to scale $C_{ox}$ to the same extent as for IM devices [2]. Consequently, it is argued that at a given $I_{on}$, the reduced $C_{ox}$ of a junctionless device offers a reduced $\tau$. This is somewhat contentious, however, as devices with scaled gate lengths would still require scaled EOTs to negate SCEs [1]. Further studies are required to determine if this indeed could offer a speed advantage while retaining good subthreshold performance, however, it is interesting to note the possible ramifications of this architecture due to its fundamentally different operation.

A criticism of junctionless devices is that their variability is worse than conventional MOSFETs due to their high channel doping concentrations, which makes them more susceptible to threshold variations due to random dopant fluctuations [5] and body thickness variations [2]. This is of course problematic for VLSI and presents its own fabrication challenges, however, the architecture merits further research given its potential benefits. Conversely, it has been suggested that the simplified fabrication process afforded by this architecture would reduce device to device variations [6].

### 4.1.2 III-V Junctionless Devices

#### The Advantage of III-Vs

The dominant scattering mechanism in a junctionless device is impurity scattering due to its necessarily highly doped channel [2, 7]. III-V compounds suffer less from mobility degradation due to impurities than Si, however, meaning that the advantage offered by III-Vs in terms of electron transport is in fact even greater in junctionless devices than it is for IM: the bulk electron mobility of $In_{0.53}Ga_{0.47}As$ is approximately 8.5 times that of Si at room temperature with $N_D \sim 10^{14}$ cm$^{-3}$ [8, 9], this increases to a factor of $\sim$13 at a doping concentration of $N_D \sim 10^{19}$ cm$^{-3}$ [8, 10]. Not only do III-Vs offer a potential advantage in performance for junctionless devices, but in fact the junctionless architecture offers an even greater simplification to the fabrication process of scaled III-V devices than it does Si: circumventing the need to form highly doped, self-aligned S/D regions by complex re-growth processes [11].

#### Device Variability

Leung *et al*. published a simulation study which demonstrated InGaAs junctionless devices to have lower variability than Si due to random dopant fluctuations (RDF) [12]. It was argued that due to the low DOS of InGaAs, local variations in potential along the channel result in smaller variations in electron density and thus have less of an impact on $V_t$ and $I_{on}$ [12]. It was further proposed that the higher permittivity of InGaAs yields lower fluctuations in $V_t$, which is inversely proportional to $\varepsilon_r$ [5, 12]. Conversely, Zagni *et al*. argued that the variability is worse in scaled InGaAs devices than in Si due to its increased quantum confinement effect, which results in a lower inversion capacitance, $C_{inv}$, and therefore a greater sensitivity to body thickness variations and RDF [13]. It is precisely such a disagreement which further merits the experimental investigation of these devices, for which there is currently little experimental data to validate these simulations, which may well not encompass all aspects of the devices to sufficiently reflect the reality.

#### Literature Review and Novelty

Yokoyama *et al*. [14] demonstrated the first III-V junctionless device in 2010, which employed an ultra thin body (UTB)-XOI architecture. The device comprised an $In_{0.53}Ga_{0.47}As$ channel on top of an $Al_2O_3$ BOX which was formed by DWB. Long channel MOSFETs were fabricated ($L_g$=50 $\mu$m) and the impact of channel thickness, $t_{ch}$, ($t_{ch}$ = 3.5 nm and 7 nm) and doping concentration ($N_D = 1 \times 10^{17}, 1 \times 10^{18}, 1 \times 10^{19}$ cm$^{-3}$) investigated. Promising results were

demonstrated with a minimum SS of 150 mV/Dec. and an $I_{on}/I_{off}$ ratio $\sim 10^7$ at a gate overdrive of 1 V. Further planar devices have since been demonstrated [15–18] with gate lengths as short as 6 nm [16]; exhibiting a maximum on-current and transconductance of 420 $\mu A/\mu m$ and 1.48 $mS/\mu m$ respectively at $V_{DD}$= 0.7 $V$, albeit with significant SCEs (DIBL$\sim$ 500 mV/V). The more moderately scale variant of the same device ($L_g$=36 nm) demonstrated a greatly improved electrostatic integrity with an SS = 115 mV/Dec and DIBL = 165 mV/V [16].

Non-planar, III-V junctionless devices have also been demonstrated, which include $In_xGa_{1-x}As$ FinFETs with x = 0.53 [19, 20] and x = 0.7 [19, 21, 22], as well as a GaAs, GAA lateral nanowire [23]. As expected, these devices offer superior immunity to SCEs compared to their planar counterparts, and Zota *et al*. demonstrated a 25 nm gate length device in 2017 with a minimum SS of 76 mV/Dec. and DIBL = 33 mV/V [22]. A complete summary of metrics relevant to low power digital logic applications for all III-V junctionless devices published to date is given in the following sections, against which the performance of the devices fabricated are benchmarked.

One advantage of the FinFET architecture which has yet been exploited for III-V junctionless devices, is the ability to increase current density per chip area by employing high aspect ratio fins. Such devices would require unpinned, low $D_{it}$ dielectric interfaces to the [110] fin sidewalls to allow for the effective modulation of channel current. To date, all non planar III-V junctionless devices that have been demonstrated have employed fin heights which are smaller than the maximum depletion width of the channel material given its doping concentration, and as such, can be well modulated by the top gate only, and offer little insight into the effectiveness of the gated sidewalls. Accordingly, this chapter presents the process development, fabrication, measurement and analysis of InGaAs junctionless FinFET devices which utilise high aspect ratio fins, predominantly gated by the dielectric interface to their [110] orientated sidewalls, allowing for an increased drive current per chip surface width compared to anything published thus far.

## 4.2 Epitaxial Layer Structure

### 4.2.1 Wafer Parameters

The layer structure of the junctionless FinFET wafer, which was grown by Dr Matthew Steer, is shown in Fig. 4.2(a) and comprised a 1.5 $\mu$m thick, lattice matched, p-doped ($N_A = 8 \times 10^{15}$ cm$^{-3}$) $In_{0.52}Al_{0.48}As$ buffer layer and a 50 nm thick, n-doped ($N_D = 6 \times 10^{18}$ cm$^{-3}$) $In_{0.53}Ga_{0.47}As$ device layer, both of which were grown by MBE on a semi-insulating (SI) InP (100) substrate. The p and n-type dopant elements were Be and Te respectively. $In_xGa_{1-x}As$ with an InAs mole fraction of 53% was chosen as the channel material for this initial investigation given the maturity of its gate stack [24] and contact technology [25, 26], as well as to minimise the formation of defects due to lattice mismatch. The wide bandgap InAlAs buffer layer serves to confine the InGaAs channel, effectively mimicking the structure of the Colinge SOI device. The band diagram and electron concentration for this layer structure at flat band, as calculated by solving the Poisson and Schrödinger equations self consistently [27], are shown Fig. 4.2(b), illustrating the confinement in the channel. The doping concentration of the InGaAs device layer of $6 \times 10^{18}$ cm$^{-3}$ was the highest possible with the particular MBE system used, and was chosen to minimise the contact resistance [25] and to maximise the on-state performance [2]. The InAlAs buffer was lightly p-doped to further increase the barrier height to the channel.

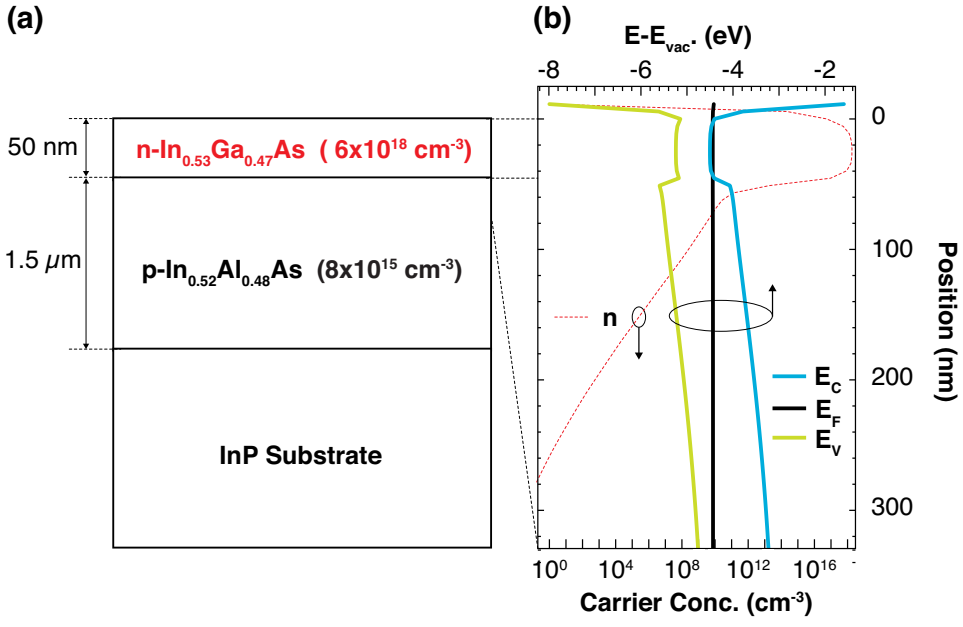


Figure 4.2: (a) Schematic epitaxial layer structure of the material used to fabricate junctionless InGaAs FinFET devices. (b) The corresponding band alignments and carrier concentrations as function of depth into the wafer as calculated by solving the Poisson and Schrödinger equations self consistently [27].

### 4.2.2 Characterisation

Minimal bowing of the wafer (∼6 nm/ mm) was measured by a Bruker Contour GT optical profiler, as shown in Fig. 4.3(a), indicating that the buffer and channel are closely lattice matched. This measurement was performed by Dr Kevin Gallagher. A $40 \times 40$ $\mu$m AFM scan of the surface measured by a Bruker Icon AFM in tapping mode (Fig. 4.3(b)) shows a number of bow-tie shaped defects which are approximately 1 $\mu$m wide, 2 $\mu$m long and protrude from the sample surface by ∼10 nm. The shape of these defects is unfamiliar to anything reported in the literature, however, it is suspected that they originate in thick InAlAs buffer layer. Fig. 4.3(c) and (d) show $500 \times 500$ nm AFM scans of the surface at the centre and edge of the wafer respectively. The degradation in root mean square (RMS) surface roughness, $r_q$, from 0.5 nm at the centre to 2.0 nm at the edge is indicative of a temperature gradient across the wafer during growth. The low measured $r_q$ at the centre validates the quality of the wafer for use in device fabrication.

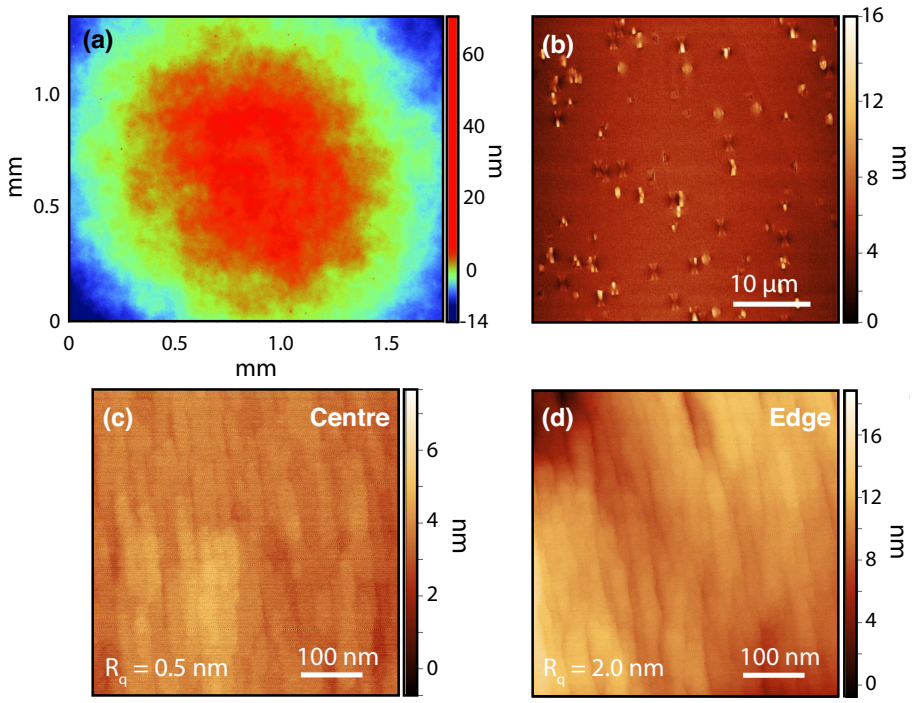


Figure 4.3: (a) Surface profile of the as grown InGaAs JFinFET wafer measured by an optical profiler. Minmial bowing of the wafer illustrates close lattice matching between the epitaxial layers. (b) $40 \times 40$ $\mu$m AFM scan of the surface at the centre of the wafer showing bow-tie shaped defects which protrude from the surface by approximately 10 nm. (c,d) $500 \times 500$ nm AFM scans of the centre and edge of the wafer respectively illustrating excellent surface roughness between the bow-tie like defects. The increased $r_q$ at the edge of the wafer is indicative of a temperature gradient from centre to edge during growth.

## 4.3 Device Parameters and Process Flow

For the first iteration of the device, it was chosen to investigate the impact of fin width on the performance of single-fin FinFETs, at a fixed $L_g$ of 200 nm. Nominal fin widths of 10 nm, 15 nm, and 20 nm were investigated. The top-down, gate-first fabrication process employed is summarised in Fig. 4.4. The following provides an overview of each process module, describing the critical requirements and consequences on device performance, and details the process development required. Full details of the corresponding process optimisation are given in the following sections.

1. **Etch Mask Patterning by EBL.** The etch mask was based on a design by Dr Muhammad Mirza for fabricating Si junctionless nanowires [7]. It comprised fin, S/D and contact regions and was patterned by EBL using HSQ resist. Differences in the magnitude of proximity exposure between Si and InGaAs required optimisation of the HSQ development process, which is discussed in Section 4.4. Also discussed, is the strategy employed to minimise the fin line edge roughness (LER), which is critical to avoid degradation to the on-state performance [28].

2. **Mesa/Fin Dry Etch.** Samples were etched using an inductively coupled plasma (ICP)-RIE process which formed both the contact mesas for device isolation, as well as the fin. Simulation studies have shown that non-vertical fin sidewalls serve to degrade off-state performance [29], and therefore, a highly anisotropic etch process is desirable, with the constraint of inducing minimal damage to the etched sidewalls: etch damage would degrade the dielectric interface to the fin and limit the devices performance, or in the worst case, pin the Fermi level at the interface such that the current cannot be modulated. A low damage etch process had been previously established within the group by Dr Olesya Ignatova and Dr Xu Li. Section 4.5 details the further development of this process, with the aim of improving sidewall verticality while retaining its low damage properties.

3. **HSQ Etch /Digital Etch.** Following the mesa/fin etch process, the remaining HSQ was selectively etched using ICP-RIE with $CF_4$ etch chemistry. Two samples were processed for comparison, one with, and one without the inclusion of a digital etch process, which is employed to recover any damage inflicted by the dry etch process [30]. It has been previously demonstrated that 3 cycles of digital etching is sufficient to yield the maximum improvement in damage recovery when an InGaAs surface is etched with this baseline fin etch process [31]. A such, 3 cycles were implemented in this device fabrication process, where the sample was oxidised using an $O_2$ plasma and the oxide selectively etched in sulphuric acid.

4. **High-k Deposition.** A bi-layer of $Al_2O_3/HfO_2$ (nominally 1/3 nm thick) was deposited

at Tyndall National Institute (TNI) using their optimal ammonium sulphide surface passivation ($(NH_4)_2S$) process [24], which has demonstrated fully unpinned InGaAs-$Al_2O_3$ interfaces with a peak $D_{it}$ of just $9 \times 10^{11}$ cm$^{-2}$eV$^{-1}$ situated $\sim E_v$ + 0.4 eV, which decreases towards the conduction band edge [24, 32]. The $Al_2O_3$ and $HfO_2$ layers were deposited in a Cambridge NanoTech Fiji F200LLC ALD reactor using a thermal process at 250°$C$. The $Al_2O_3$ was deposited via alternating cycles of TMA and water ($H_2O$), while tetrakis-ethylmethylaminohafnium (TEMAHf) and $H_2O$ were employed as precursors for $HfO_2$ deposition. This process has demonstrated accumulation capacitances of 1.6 $\mu$Fcm$^{-2}$ [33].

5. **Gate Metal Deposition.** The gate contact was formed by patterning the sample using EBL with PMMA resist, and subsequently lifting-off Pt/Au (20/200 nm), which was deposited by ebeam evaporation. Pt was chosen for its high work function in order to yield a more positive threshold voltage than lower work function metals. The deposited gate metal length was 2 $\mu$m as shown in Fig. 4.5. This was chosen to allow its formation by lift-off in order to simplify the process and remove the need to employ a more complex subtractive process, which would be required to form smaller gate metal lengths. Given the geometry of the S/D regions and the high doping concentration, the gate is only able to fully deplete the fin, the length of which defines the gate length of the device.

6. **Ohmic Contact Formation.** The ohmic contact regions were also patterned using EBL with PMMA resist. Following resist development, the exposed high-k was selectively dry etched using $SiCl_4$ chemistry with in-situ interferometric monitoring. The contact system employed was annealed Ni-InGaAs. Prior to metal deposition, the native oxides present on the sample surface were removed by in-situ Ar sputtering, and subsequently, the following contact layers were deposited by ebeam evaporation: Ni/Ti/Pd/Au (4/15/15/100 nm). Following lift-off, the sample was annealed in a rapid thermal anneal(er) (RTA) at 350°$C$ for 60 s in an $N_2$ purged environment. This process had been previously developed within the group by Dr Uthayasankaran Peralagu and Dr Olesya Ignatova and exhibited a specific contact resistance of $1.92 \times 10^{-6}$ $\Omega$cm$^2$ on n-$In_{0.53}Ga_{0.47}As$ doped at $5 \times 10^{19}$ cm$^{-3}$. The distance between the contact and gate edge was chosen to be 100 nm in an attempt to minimise the access resistance should there be charge in the oxide that depletes the surface of the un-gated, or 'underlapped', regions, while not placing too stringent demands on the alignment process. This is depicted in Fig. 4.5 for clarity.

Full details of the above processing steps are provided in Appendix A.

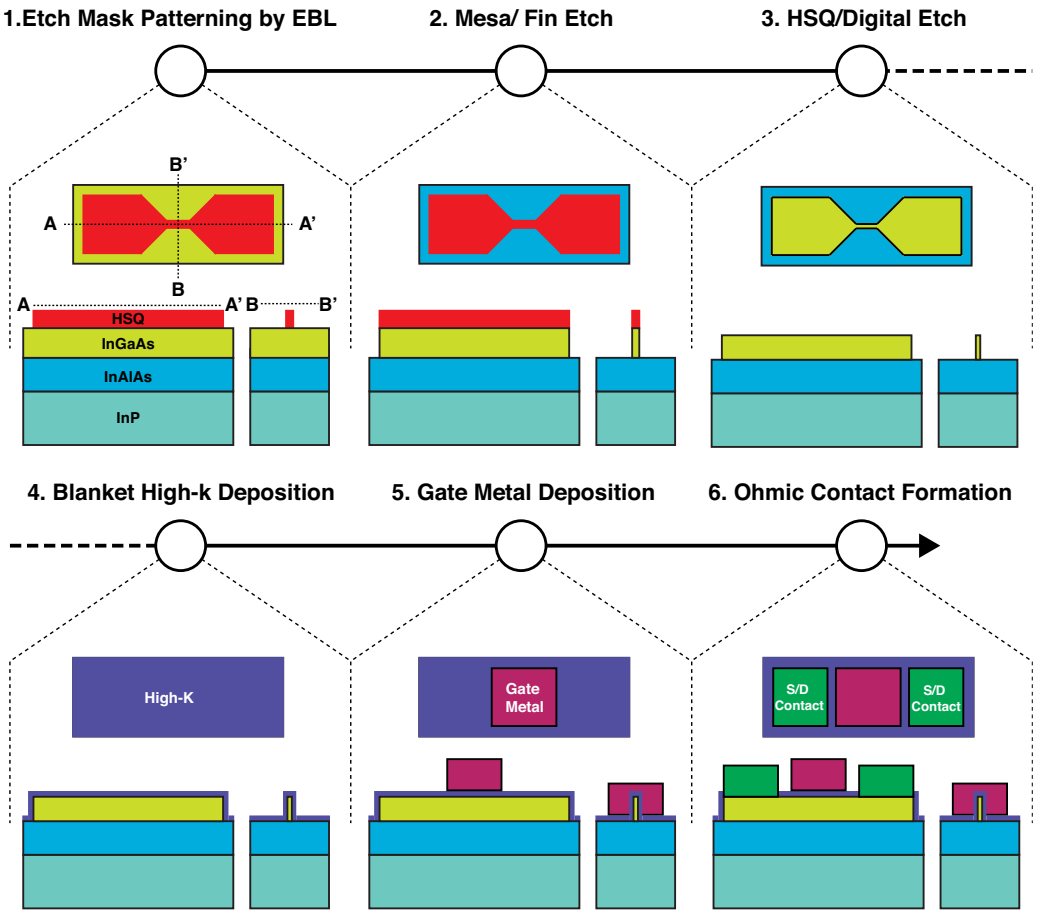


Figure 4.4: Simplified schematic process flow for the top-down, gate-first, InGaAs junctionless FinFET comprising: etch mask patterning by electron beam lithography (EBL) in HSQ resist; mesa and fin definition using an inductively coupled plasma reactive ion etching (ICP-RIE) process; HSQ removal and digital etch to recover dry etch damage; blanket deposition of $Al_2O_3/HfO_2$ (1 nm/3 nm) high-k bi-layer; gate metal deposition via EBL, ebeam evaporation of Pt/Au (20/200 nm) and lift-off; and S/D contact formation by EBL, high-k dry etch selective to InGaAs, in-situ native oxide etching with Ar sputtering, ebeam evaporation of Ni/Ti/Pd/Au (4/15/15/100 nm) and lift-off.

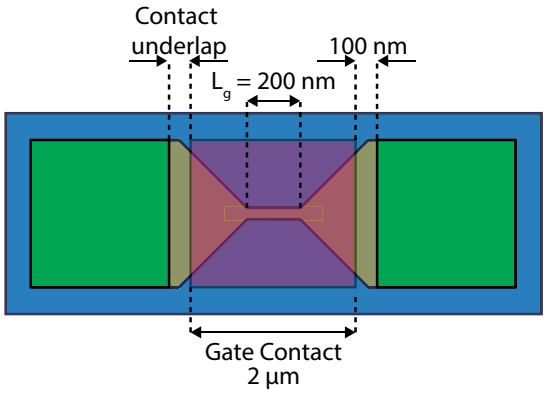


Figure 4.5: Schematic of the top view of the finished device, illustrating the gate (fin) length, $L_g$, the gate contact size, and the distance between gate metal edge and contact regions.

## 4.4 Mask Defintion

### 4.4.1 Post Exposure Baking of HSQ Resist

The HSQ etch mask, which comprised fin, S/D, contact and bond pad components, is shown in Fig. 4.6. In order to minimise unwanted proximity exposure, the mask was patterned in two lithographic layers which were aligned to each other, with the first layer (L1) patterning fin, S/D and contact features, and the second (L2) the bond pad regions. An initial dose test was conducted for each component individually to narrow the process window, and a further subsequent dose test conducted with the corresponding components for L1 combined, and their doses varied about their previously determined isolated dose. Proximity error correction (PEC) was included, which utilised a PSF that was specific to this epitaxial layer structure as determined by Monte Carlo simulation [34]. Nonetheless, it was found that there was no combination of doses which yielded the correct pattern dimensions without overexposing the HSQ surrounding the desired features. An example of this is shown in Fig. 4.7(a) where the mask components were written with the parameters given in Table 4.1 and the resist was processed as per Table 4.2, excluding the step highlighted in red.

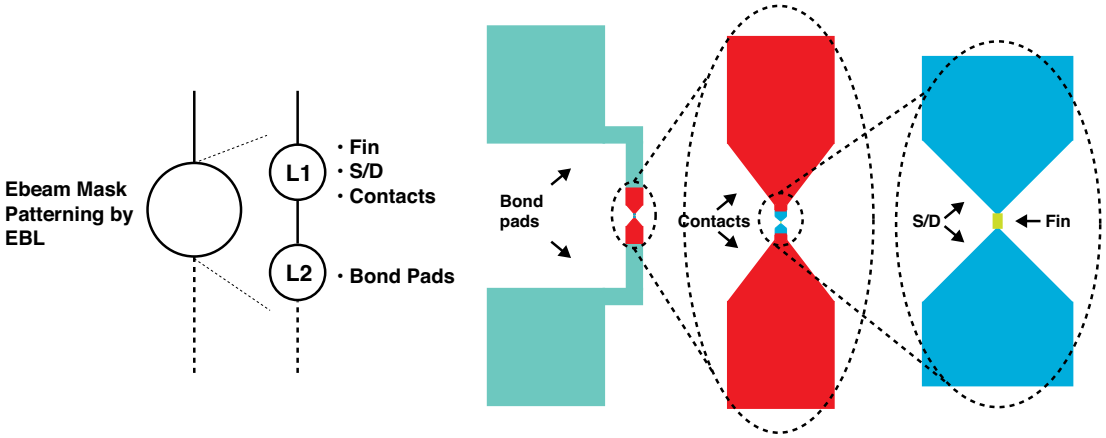


Figure 4.6: Process flow and schematic diagram of the HSQ mask defined by electron beam lithography. The mask is written in two layers to minmise proximity exposure. Layer 1 (L1) patterns the fin, S/D and contact regions and layer (L2) patterns the bond pad mesa.

Patterning the combined features included in L1 was not problematic on Si, which this mask was originally developed for, however, proximity exposure is worse in III-Vs due to their increased backscaterring coefficient [35]. While it would have been possible to circumvent this issue by patterning the mask in more lithographic layers, it was preferable to investigate a process which reduced the impact of proximity exposure: this would result in a higher throughput of samples and could also prove advantageous for fabricating future iterations with more scaled gate lengths. Incorporating a post exposure bake (PEB) into the development process, as highlighted in red

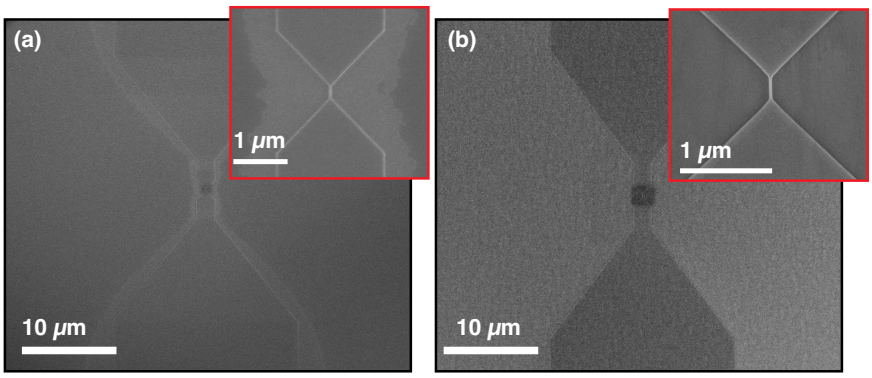


Figure 4.7: SEM images of the first lithographic layer (L1) written in HSQ with the ebeam parameters shown in Table 4.1 with (a) and without (b) the inclusion of a post exposure bake (PEB) during HSQ development. The dark region around the fin visible in (b) is due to hydrocarbon contamination from previous inspected at high magnification on an SEM.

| Mask Component | Beam Current (nA) | Spot Size (nm) | BSS (nm) | Area Dose ($\mu$Ccm$^{-2}$) |
|---|---|---|---|---|
| Fin | 1 | 4 | 1.25 | 30000$^{\dagger}$ |
| Source/Drain | 1 | 4 | 2.5 | 1950 |
| Contacts | 32 | 24 | 20 | 970 |

$^{\dagger}$ The fin was written using a technique called undersize-overdose, as discussed below, which results in what appears to be an anomalously high area dose.

Table 4.1: HSQ application and development process with (highlighted in red) and without incorporating a post exposure bake (PEB).

| | |
|---|---|
| Sample Clean | 5 min. acetone at 50°$C$<br>5 min. IPA<br>60 s. SurPass3000<br>60 s. DI water rinse<br>$N_2$ blow dry |
| Resist Parameters | 3:1 HSQ:MIBK spun at 5k rpm |
| Soft Bake | 90°$C$ for 2 min. on hotplate |
| Development | PEB at 195°$C$ for 2 min. on hotplate<br>30 s TMAH (25%)<br>60 s DI water rinse<br>15 s IPA rinse<br>$N_2$ blow dry |

Table 4.2: Electron beam lithography exposure parameters for lithographic layer L1.

in Table 4.2, where the sample is baked at high temperature subsequent to ebeam exposure and prior to development, has been shown improve the contrast and aid the removal of proximity exposed regions between features [36]. This is somewhat counter intuitive, as increasing the

baking temperature of HSQ prior to ebeam exposure results in thermally activated cross linking which reduces its contrast [37]. The improvement gained by employing a PEB, however, is attributed to the formation of $SiO_2$, which passivates the HSQ surface and prevents it bonding with contaminants which are insoluble in the developer [36]. Fig. **??**(b) shows the resulting HSQ mask, with no residual HSQ present, written with identical parameters as in (a) but for the inclusion of the PEB step highlighted in red in Table. 4.2. It should be noted that the dark region around the fin visible in (b) is hydrocarbon contamination from having previously inspected this area at high magnification, which is shown inset.

### 4.4.2 Undersize-Overdose

The fin was written using a technique known as undersize-overdose, which is where the width of a feature is defined to be smaller than the desired result, and subsequently patterned with a high dose which results in overexposure and yields the correct dimension [38, 39]. This process has been shown to reduce the LER of patterned features and also increase the process latitude. This can be understood by considering the fin width vs dose plot shown in Fig. 4.8. The data for this was collected by conducting a dose test with a fixed design fin width of 5 nm, (shown as the black vertical lines in (a)), and measuring the resulting patterned widths using an SEM. By employing the undersize-overdose method, the dose required to pattern the desired feature dimensions is situated in a region of the dose-fin width curve which is less sensitive to changes is dose. Consequently, any variations in beam current when writing a pattern result in a smaller change in feature size and thus smaller LER [39].

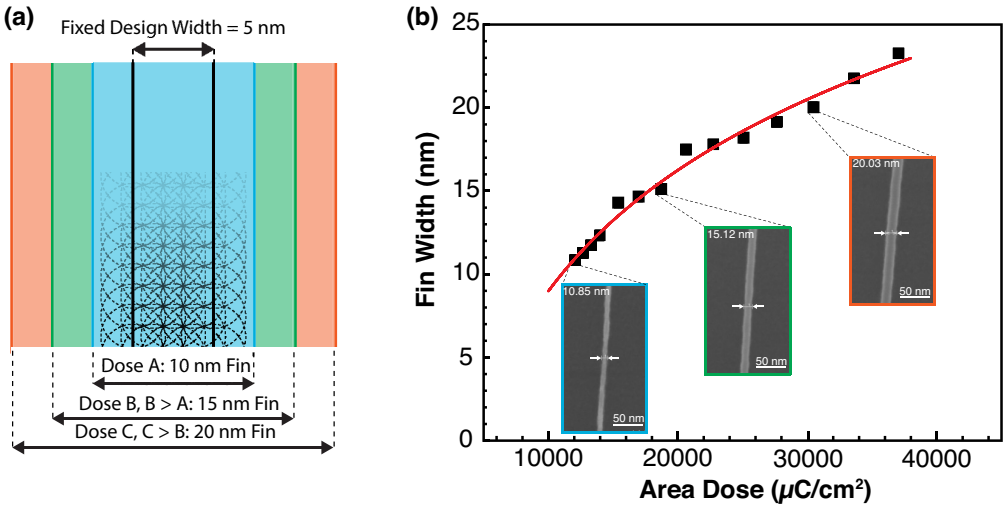


Figure 4.8: (a) Illustration of the undersize-overdose electron beam lithography (EBL) writing strategy. The solid vertical black lines show the design width of the fin in L-Edit, which was 5 nm for all patterned fin widths. Varying fin widths can be written by changing the dose with which the black rectangle is written. (b) Experimental dose vs fin width data from patterning fins as illustrated in (a), with the EBL and resist parameters given in Tables 4.1 and 4.2 respectively, including the PEB step. The circles in (a) illustrate the placement of the EBL beam shots at the given parameters.

## 4.5 Anisotropic Etching of In-Containing III-Vs

III-V compounds are highly susceptible to damage due to physical sputtering [40, 41], which constrains a low damage etch process to minimise the induced direct current (DC) self bias to minimise the energy with which ions bombard the sample. Given that etch anisotropy generally increases with increasing DC bias [42], designing a highly anisotropic, low damage etch process for III-V compounds presents a difficult fabrication challenge. This is particularly the case for In-containing compounds, as their etch byproduct with $Cl_2$, which historically is the commonly used etchant gas for III-Vs, is non-volatile at temperatures below 200°C [43, 44]. This results in the accumulation of $InCl_x$ on the sample, leading to rough surface morphologies, poor etch anisotropy and low etch rates [44]. Etching In-containing III-Vs in $CH_4/H_2$ mixtures has demonstrated highly vertical etched sidewalls with smooth surfaces for large feature sizes such as mesas [45]. Interactions between atoms, ions and radicals within the $CH_4/H_2$ plasma can result in the deposition of an organic polymer film [45]. The mechanism of etch anisotropy with this etch chemistry is thought to be due to balancing the anisotropic etching due to physical sputtering and ion assisted etching, the isotropic chemical etching due to radicals generated in the plasma, and the simultaneous passivation of the surface due to polymer film deposition [45]. The phenomena of positively sloped sidewalls, where the base of the etched feature is larger than the etch mask, is commonly observed when etching In-containing compounds [46, 47]. For $CH_4/H_2$ etch chemistry, this is attributed to polymer deposition rates which exceed the effective anisotropic etch rate [46, 47]. Haneji *et al*. demonstrated a process of cyclic etching using $CH_4/H_2$ and by $O_2$, where the to $O_2$ step was intended to ash any unwanted polymer that had been deposited. This was shown to improve the sidewall verticality for various In-containing III-Vs compared to continuously etching in $CH_4/H_2$ [46]. While $CH_4/H_2$ based etch processes circumvent the issue of the low volatility of $InCl_x$, their etch rates are prohibitively slow and etching with these gasses alone is not feasible at low bias [48]. Constantine *et al*. demonstrated that etching InP in a $Cl_2/CH_4/H_2$ mixture retained the excellent surface morphology and etch anisotropy of $CH_4/H_2$ plasmas, while significantly increasing the etch rate [48]. The low damage InGaAs recipe previously developed within the group by Dr Ignatova and Dr Li combines the promising results of the $Cl_2/CH_4/H_2$ chemistry, with the effectiveness of using $O_2$ to ash any deposited organic polymer, and employs continuous etching with $Cl_2/CH_4/H_2/O_2$ etch chemistry. This process yielded visually smooth surfaces with a sidewall angle of 76.8° at a DC self bias of 113 V. The complete recipe and corresponding etched fin are shown in Fig. 4.9. Dry etch processes for etching InGAs features of this size are scarce in the literature, with the first demonstration of sub 20 nm vertical features demonstrated in 2014 by Zhao *et al*. [49].Their process used $BCl_3/SiCl_4$/Ar chemistry, however, it utilised a high DC bias of 280 V [49], which illustrates the promise of the $Cl_2/CH_4/H_2/O_2$ process. The following details an array of etch experiments which endeavoured to improve the sidewall verticality of this base recipe, while

retaining the smooth surface roughness and maintaining the DC self bias at $\sim$ 100 V . While this target bias is somewhat arbitrary, previous investigations examining the degradation of qualitative MOS characteristics for samples which had been exposed to the baseline etch process, show the damage induced at this bias to be minimal and readily recoverable via subsequent plasma processing [31].

| | Fin Etch |
|---|---|
| Etch Chemistry | $Cl_2/CH_4/H_2/O_2$ |
| Flow Rate (sccm) | 6/10/15/0.5 |
| ICP Power (W) | 250 |
| Platen Power (W) | 25 |
| Pressure (mT) | 2 |
| Temperature ($^{\circ}C$)) | 120 |
| DC Bias (V) | 113 |
| Time (mm:ss) | 02:30 |

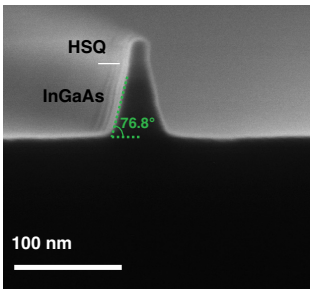


Figure 4.9: Inductively coupled plasma - reactive ion etching (ICP-RIE) parameters for the baseline, low damage InGaAs etch process and a scanning electron microscope (SEM) image of the corresponding etched fin profile.

### 4.5.1 InGaAs Dry Etch Process Development

Initial tests were undertaken on an Oxford instruments ICP180 ICP-RIE etch tool with circular HSQ etch masks which pertained to the development of a vertical nanowire process. Ultimately, however, it was the recipe yielded from this set of experiments that was implemented as part of the junctionless FinFET process.

#### Etch Recipe 1: Increased $H_2$ Flow

Hayes *et al*. demonstrated that when etching InP in a $CH_4/H_2$ discharge, the polymer deposition rate increased as the ratio of $CH_4$ to $H_2$ was increased [45]. In attempt to decrease the polymer deposition rate and reduce surface passivation, the $H_2$ flow rate was increased by $\sim$ 50% from 15 to 22 standard cubic centimetres per minute (SCCM). This improved the sidewall angle from 76.8° to 81.3°. A marginal decrease in InGaAs etch rate was observed from 20 nm/min to $\sim$18 nm/min, and the InGaAs:HSQ selectivity decreased from 3.9:1 to 1.9:1, indicating that there was indeed less polymer being deposited.

| | Fin Etch |
|---|---|
| Etch Chemistry | $Cl_2/CH_4/H_2/O_2$ |
| Flow Rate (sccm) | 6/10/22/0.5 |
| ICP Power (W) | 250 |
| Platen Power (W) | 25 |
| Temperature ($^{\circ}C$ | 120 |
| DC Bias (V) | 113 |
| Time (mm:ss) | 02:30 |

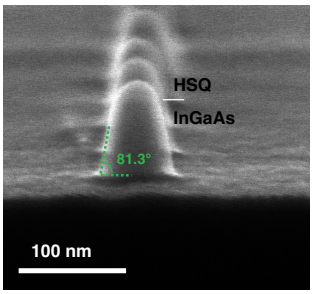


Figure 4.10: Inductively coupled plasma - reactive ion etching (ICP-RIE) parameters for Etch Recipe 1 and a scanning electron microscope (SEM) image of the corresponding etched fin profile. The changes from the baseline process are highlighted in red.

### Etch Recipe 2: Increased $H_2$ and $O_2$ Flow

To further decrease passivation due to polymer deposition, the $O_2$ flow was increased by 100 % from 0.5 to 1.0 SCCM, while retaining the increased $H_2$ flow from the previous experiment. This further improved the sidewall angle to 83.7°, and also improved the InGaAs:HSQ selectivity to 3.5:1. The etch rate remained at 18 nm/min and there appeared to be no degradation to the surface morphology.

| | Fin Etch |
|---|---|
| Etch Chemistry | $Cl_2/CH_4/H_2/O_2$ |
| Flow Rate (sccm) | 6/10/22/1 |
| ICP Power (W) | 250 |
| Platen Power (W) | 25 |
| Pressure (mT) | 2 |
| Temperature ($^{\circ}C$ | 120) |
| DC Bias (V) | 113 |
| Time (mm:ss) | 02:30 |

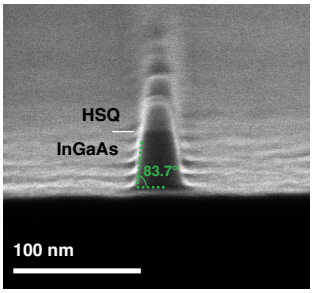


Figure 4.11: Inductively coupled plasma - reactive ion etching (ICP-RIE) parameters for Etch Recipe 2 and a scanning electron microscope (SEM) image of the corresponding etched fin profile. The changes from the baseline process are highlighted in red.

### Etch Recipe 3: Increased ICP Power

Figure 4.12 shows the impact of increasing the ICP power: with the flow rates fixed at the values for Etch Recipe 2, the ICP power was increased from 250 W to 500 W. This significantly improved the resulting profile, with the upper third of the nanowire being extremely vertical, and lower two thirds tapering at angle of 86.2°. Increasing the ICP power increases the plasma density, and therefore, the reactive chemical component of the etch. This vast improvement with

increased chemical etching indicates that the previous, more positively sloped processes, were dominated by low energy physical sputtering.

| | Fin Etch |
|---|---|
| Etch Chemistry | $Cl_2/CH_4/H_2/O_2$ |
| Flow Rate (sccm) | 6/10/22/1 |
| ICP Power (W) | 500 |
| Platen Power (W) | 25 |
| Pressure (mT) | 2 |
| Temperature ($°C$ | 120) |
| DC Bias (V) | 88 |
| Time (mm:ss) | 02:30 |

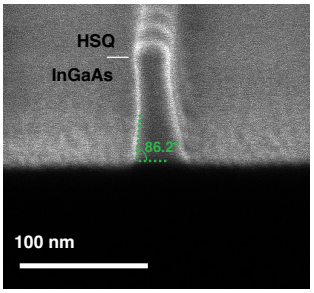


Figure 4.12: Inductively coupled plasma - reactive ion etching (ICP-RIE) parameters for Etch Recipe 3 and a scanning electron microscope (SEM) image of the corresponding etched fin profile. The changes from the baseline process are highlighted in red.

### Etch Recipe 4: Overetching

Over-etching, where the sample is etched beyond the point that the desired etch depth is reached, with any vertical etch component suppressed by the InAlAs etch stop layer, has been shown to improve sidewall verticality by scattering incoming ions towards the foot of the etched feature and thus aiding its removal [50]. Fig 4.13 shows the resulting etched wires when etched with Etch Recipe 3, with the addition of a 30 s over etch. It can be see that this process effectively reduces the foot at the base of the wire, however, the surface morphology has degraded severely due to the complete consumption of the HSQ mask, which has resulted in micromasking [51].

| | Fin Etch | Over Etch |
|---|---|---|
| Etch Chemistry | $Cl_2/CH_4/H_2/O_2$ | |
| Flow Rate (sccm) | 6/10/221 | |
| ICP Power (W) | 500 | |
| Platen Power (W) | 25 | 10 |
| Pressure (mT) | 2 | |
| Temperature ($°C$) | 120 | |
| DC Bias (V) | 88 | 34 |
| Time (mm:ss) | 02:30 | 00:30 |

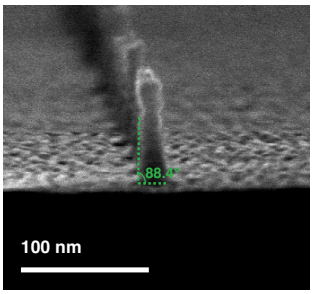


Figure 4.13: Inductively coupled plasma - reactive ion etching (ICP-RIE) parameters for Etch Recipe 4 and a scanning electron microscope (SEM) of the corresponding etched fin profile. The changes from the baseline process are highlighted in red.

**Etch Recipe 5: Final Fin Etch Process**

Table 4.3 shows the final ICP-RIE parameters used for device fabrication and Fig 4.14 shows the corresponding cross sectional SEM images of etched fins of 10 nm, 15 nm and 20 nm design widths. This recipe differs from that of Etch Recipe 4 as the process was transferred from the ICP180 to an an Oxford instruments Cobra ICP-RIE etch tool, which is part of a clustered ICP-RIE, ICP deposition, ALD and scanning Auger system. Transferring the process required increasing the flow rates by $\sim 20\%$ to account for the the larger chamber size, as well as marginally increasing the $CH_4$:$H_2$ ratio. Given the smaller etch depth of the fins compared to the nanowires in Fig. 4.13, the HSQ mask is left intact and the effectiveness of overetching is evident with no degradation observable in sidewall roughness compared to Etch Recipe 3.

| | Fin Etch | Over Etch |
|---|---|---|
| Etch Chemistry | $Cl_2/CH_4/H_2/O_2$ | |
| Flow Rate (sccm) | 7.2/12/18/0.5 | |
| ICP Power (W) | 500 | |
| Platen Power (W) | 25 | 10 |
| Pressure (mT) | 2 | |
| Temperature ($^\circ C$) | 120 | |
| DC Bias (V) | 111 | 66 |
| Time (mm:ss) | 01:25 | 00:30 |

Table 4.3: Inductively coupled plasma - reactive ion etching (ICP-RIE) parameters for the Cobra fin etch recipe used in device fabrication.

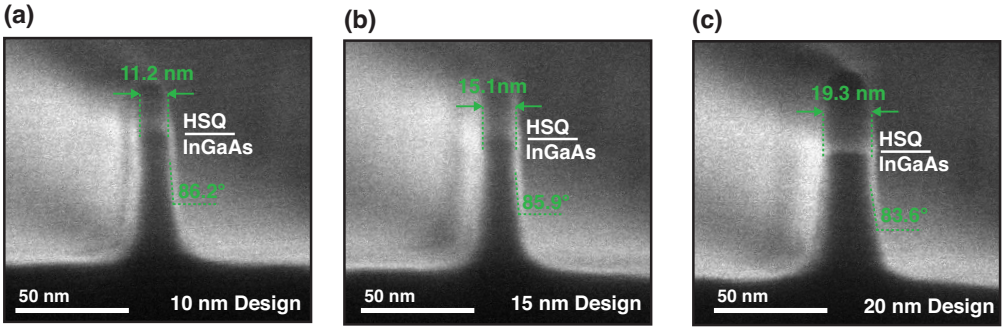


Figure 4.14: Scanning electron microscope (SEM) images of etched InGaAs fins using the recipe given in Table 4.3 for (a) 10 nm, (b) 15 nm and (c) 20 nm design widths.

### 4.5.2 Digital Etch

Fig. 4.15 shows bright field (BF) TEM images of the three fin widths following complete device fabrication, including digital etch. These images were taken at TNI using a JEOL 20000FX microscope operated at 200 kV. The samples were prepared for TEM using focussed ion beam

(FIB) thinning on a FEI 200 workstation. The etched sidewall angle of each fin width appears to be significantly less vertical than that measured post fin etch, and the base of the fins appear wider than observed in the SEM images shown in Fig. 4.14. It is suspected that there is variability in the etch process, which could arise, for example, as a result of different etch chemistries used by other users and result in contamination in the etch chamber. Furthermore, it is possible that the digital etch process, and/or the $(NH_4)_2S$ passivation process, have preferentially etched the top of the fins compared to the bottom, resulting in the more tapered profile shown. Further investigation is required. What appears to be an inhomogeneous dielectric layer surrounding the fins is thought to have resulted from damage during the FIB process, as low gate leakage currents were measured (shown in the following section). The varying contrast across the fins is also thought to be due to damage from the FIB.

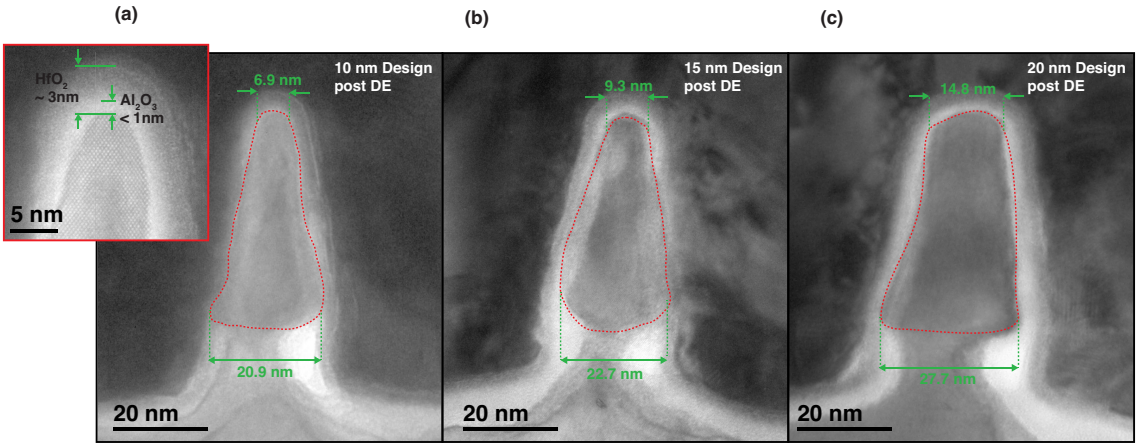


Figure 4.15: Transmission electron microscope (TEM) images of nominally designed 10 nm, 15 nm and 20 nm fins, following complete device fabrication. The red line illustrates the perimeter of the InGaAs fin. The bi-layer dielectric is visible in the higher magnification image shown inset to (a), showing $Al_2O_3$ and $HfO_2$ thicknesses of < 1nm and $\sim$ 3 nm respectively.

## 4.6 Electrical Results

For clarity, when comparing the impact of digital etch and fin width on electrical performance, the following naming convention is adopted. Samples that have been processed with and without digital etching are referred to as 'DE' and 'NoDE' respectively. Individual devices are referred to by their nominal fin width prior to digital etching, irrespective of whether they have been digital etched or not. Digital etching of course reduces the actual width and this accounted for during analysis. The '10 nm-DE' device, therefore, corresponds to the narrowest fin on the digitally etched sample. Both samples DE and NoDE comprised 4 cells each containing $33 \times 10$ nm fin width devices, $11 \times 15$ nm fin width devices, and $11 \times 20$ nm fin width devices, totalling 220 transistors per sample. DC electrical measurements were performed at room temperature using a Keysight B1500 Semiconductor Analyser and Cascade Microtech automated probe station with four-point Kelvin probes. This was used in conjunction with a purpose written MATLAB script, which allowed for the automated measurement and analysis of all devices on each sample. The total yield for both samples was found to be $\sim 40$ %, however, error in the alignment process resulted in the failure of all devices in 2 cells of sample DE, halving the total number of devices available for comparison. The total number of working devices for each sample and fin width is summarised in Fig. 4.16, where a device was deemed to be 'working' if the gate leakage current was less than 1 nA. Unless otherwise stated, all values for devices on sample DE were normalised to the gated perimeter as measured from the TEM images shown in Fig. 4.15. Given the unexpected fin geometries observed in these images, it is not clear what the correct length of the gated perimeter is for devices on sample NoDE, for which there are no TEM images. For simplicity, the perimeter of the fins shown in Fig. 4.14 was used. This may well induce some error, however, as expected the devices on sample DE exhibited superior performance than those on sample NoDE, and consequently, this error will only manifest as an under or overestimation of the benefit of digital etching, and does not effect the quoted metrics of the best performing devices.

Fig. 4.17 shows typical measured transfer and output data for each fin width from sample DE, illustrating well behaved transistor characteristics with low gate leakage. The following sub-sections provide a detailed analysis and discussion of the measured devices metrics, concluding on the impact of fin width and DE, and benchmarking the best performing device against those published in the literature.

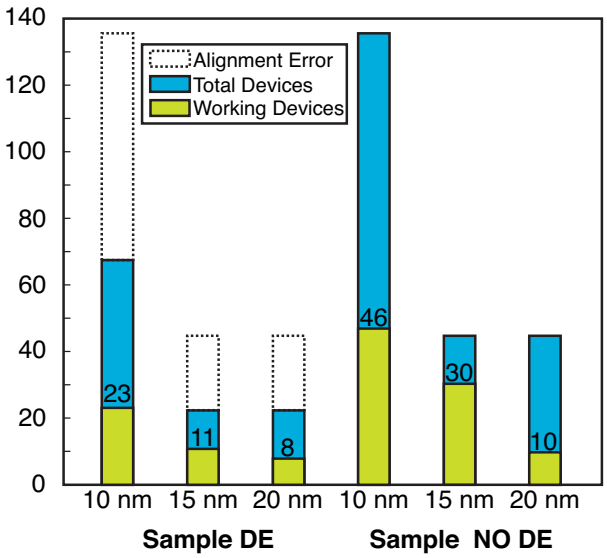


Figure 4.16: Comparison of the yield for samples processed without and without a digital etch (DE). Devices were deemed to be 'working' if their gate leakage was less than 1 nA.

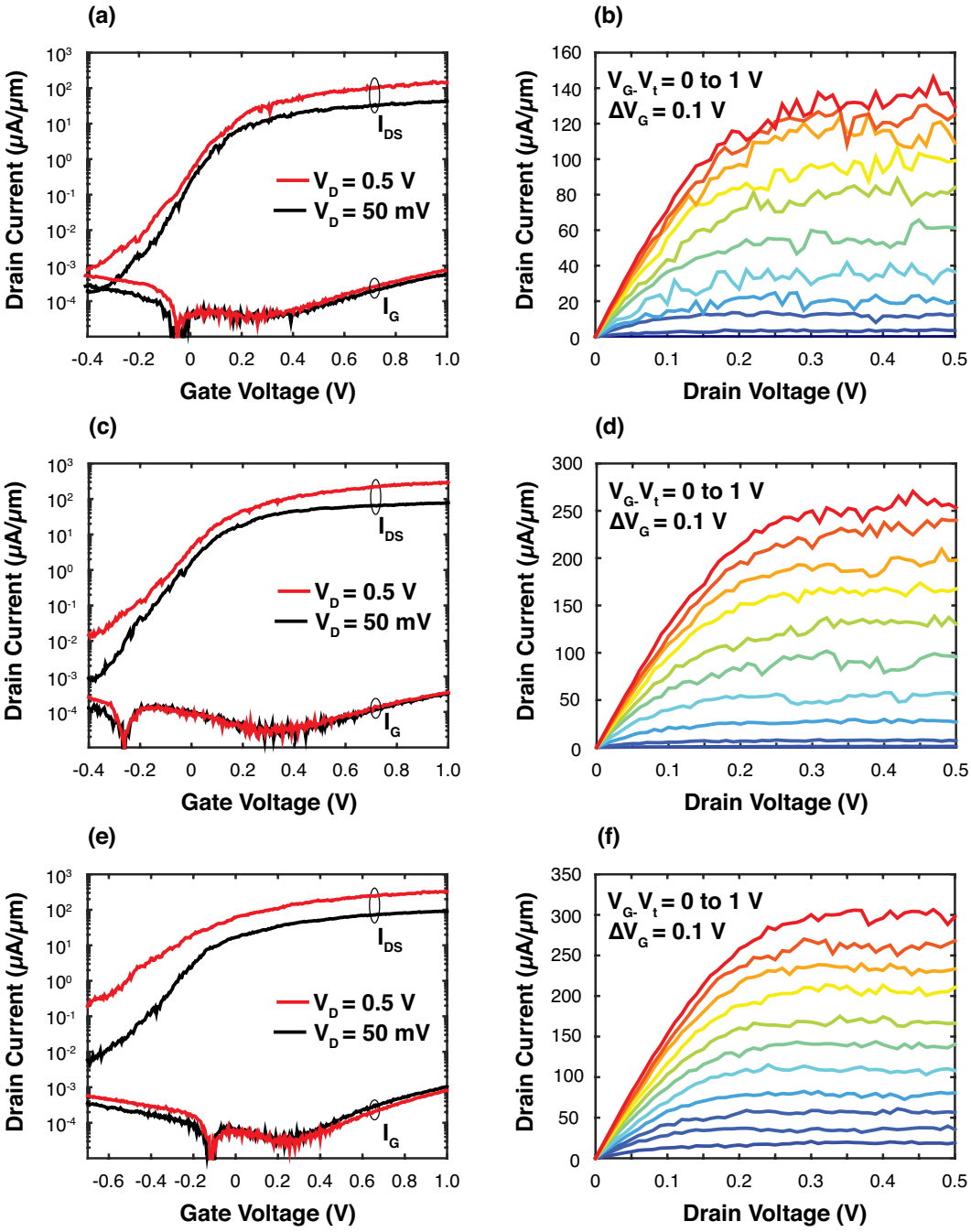


Figure 4.17: Measured room temperature transfer ($I_d$-$V_g$) and output ($I_d$-$V_d$) characteristics for (a,b) 10 nm, (c,d) 15 nm and (e,f) 20 nm devices which included the digital etch (DE) process.

### 4.6.1 Noise

Significant noise was found to be present in the $I_d$-$V_g$ and $I_d$-$V_d$ responses for both samples DE and NoDE, the magnitude of which appeared qualitatively to increase with decreasing fin width, as evident in Fig. 4.17. It is suspected that this is RTN due to the trapping and de-trapping of carriers in the channel with border traps in the oxide [52, 53]. This mechanism would have a greater effect on narrow fins as a larger percentage of the channel would be depleted by a given charge in the oxide than that of a wider fin. Quantum confinement effects will cause the energy of first sub-band of the conduction band to increase as the fin width is reduced [54], meaning that the Fermi level of the narrowest fin in the on-state will be situated higher in energy than that of wider fins. It has been shown that the density of border traps in $Al_2O_3$ increases with energy above the conduction band edge of InGaAs [55], and therefore, it could also be the case that the Fermi level of the narrowest device is aligned with a higher density of border traps than the wider devices. The magnitude of the noise will be somewhat exacerbated due to the fact that these are single fin devices, however: commonly data is reported for devices with multiple fins in parallel, for which a random fluctuation in current in any individual fin would yield a much smaller variation in total device current at any point in time. Further investigation is required to ascertain the source of this noise and to quantify its dependancy of fin width.

### 4.6.2 Quantifying the Impact of Digital Etching and Fin Width

The following compares the electrical performance of devices processed with and without digital etching, in terms of $V_t$, SS, DIBL, $R_{on}$, $g_m$ and $I_{on}$, all as a function of fin width. Significant variability was observed in the electrical characteristics of nominally identical devices for both samples, and therefore, in order to illustrate the overall trends in these metrics, comparison was made using the ten devices of each fin width from each sample that exhibited the highest $I_{on}$ values, with $I_{on}$ measured at $V_g$ = 0.5 V from $I_{off}$ = 100 nA/$\mu$m with $V_d$ = 0.5 V. The following plots the mean value from each group of ten devices for each metric. The error bars shown correspond to one standard deviation.

#### Threshold Voltage ($V_t$)

Figure 4.18 illustrates the trend in $V_t$, which was extracted at a fixed current of $I_d$= 500 nA/$\mu$m with $V_d$ = 50 mV. For both samples, the overall trend is an increasing $V_t$ with decreasing fin width. This is expected due to the quantum confinement effects discussed in the previous section requiring an increasingly positive gate voltage to accumulate the same charge for decreasing fin widths [56]. The negative shift in $V_t$ for sample DE has been observed previously with etched

vertical InGaAs nanowires and is attributed to the removal of positive ions incorporated in the InGaAs sidewalls during the dry etch process [30].

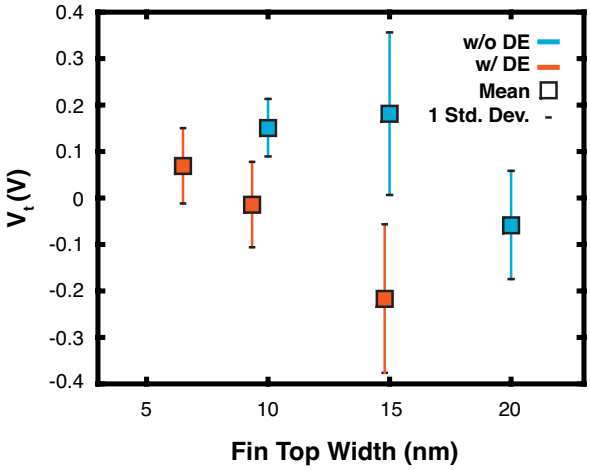


Figure 4.18: Comparison of the threshold voltage ($V_t$) as a function of fin width, as extracted at fixed $I_d$ = 500 nA/$\mu$m and $V_d$ = 50 mV, with and without digital etching (DE).

### Subthreshold Swing (SS) and Drain Induced Barrier Lowering (DIBL)

Figure 4.19 shows the minimum SS, extracted at $V_d$=50 mV, as a function of fin width for both samples. Given the noise in the transfer characteristics, SS was calculated over a large $\Delta V_g$ increment of 0.1 V, which corresponds to 20 data points. The quoted minimum value for each device corresponds to the single steepest point over any $\Delta V_g$, which is commonly how SS is reported in the literature. As expected, the minimum SS reduces with the reducing fin width as a consequence of 2-dimensional (2D) electrostatics [57].

Fig. 4.20 compares the impact of fin geometry on DIBL. For both samples, the minimum measured DIBL decreases with reducing fin width, which is expected due the increased capacitive coupling between the gate and channel for narrower fins [54]. The mean value of DIBL for all fin widths however is larger than expected given the scaled channel geometries and comparatively long gate length. Furthermore, the effects of significant DIBL were not evident in the output characteristics of any device, which would cause a high output conductance. It is suspected that the measured shift in $V_t$ is not due to the encroachment of the electric field from drain to source - DIBL - but rather another mechanism, such as charging of the oxide between voltage sweeps. Further investigation is required, however.

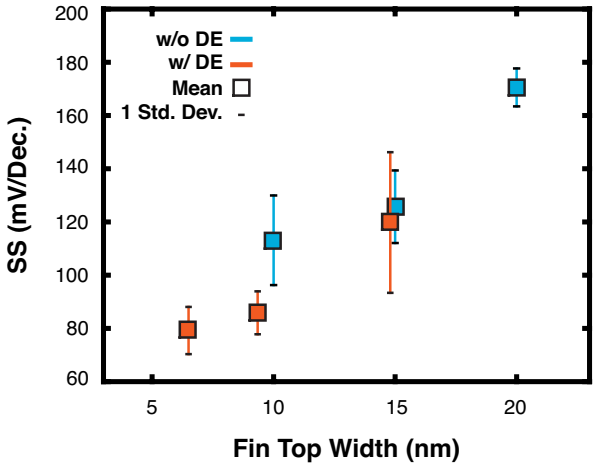


Figure 4.19: Comparison of the extracted minimum subthreshold swing (SS) as a function of fin width, with and without digital etching (DE). SS was calculated of a large gate voltage increment of 0.1 V in order to minimise the error due to noise.

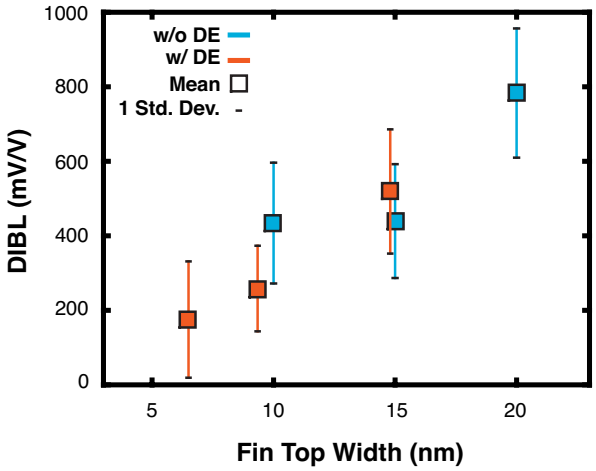


Figure 4.20: Comparison of the extracted minimum drain induced barrier lowering (DIBL) as a function of fin width, with and without digital etching (DE).

### On-resistance ($R_{on}$)

The on-resistance of the devices, $R_{on}$, plotted in Fig. 4.21, was comparable to other devices of similar geometries reported in the literature [58]. Commonly, the contributions to $R_{on}$ from the channel and contacts, $R_{ch}$ and $R_{sd}$ respectively, where $R_{on}$=$R_{ch}$+$R_{sd}$, are calculated by measuring $R_{on}$ of devices with different gate lengths (nominally identical otherwise) and extrapolating the data to zero $L_g$ to yield $R_{ch}$ [59]. This approach was not possible with this dataset, however, given the fixed gate length of all devices. Alternatively, $R_{ch}$ and $R_{sd}$ can be extracted through the analysis of $I_d$-$V_d$ measurements of a single device [60]. This approach requires the comparisons

of drain currents at closely separated drain voltages ($\Delta V_d \sim$ 10 mV) in the linear regime, however, and produced nonsensical results given the noise present in these devices.

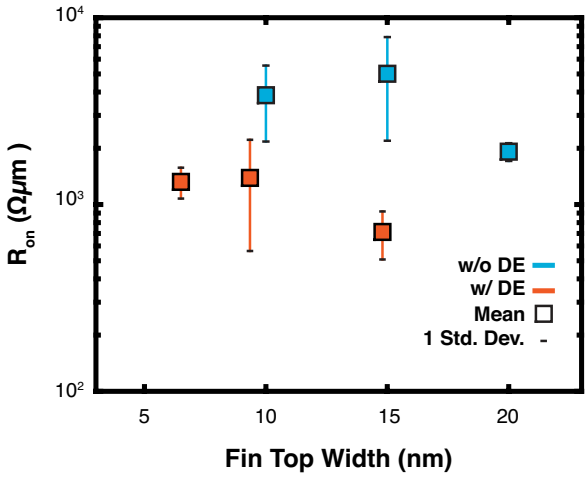


Figure 4.21: Comparison of the extracted on-resistance ($R_{on}$) as a function of fin width, with and without digital etching (DE).

### Transconductance ($g_m$) and On-current ($I_{on}$)

The transconductance was found to degrade with fin width, as shown in Fig. 4.22, which plots the maximum transconductance at $V_d$=0.5 V. This is expected due to mobility degradation as carries are pushed towards the surface of the fin and the surface roughness scattering component increases [54]. Digital etching improves the mean value of $g_m$ by a factor of $\sim$ 3.6 and $\sim$8.3 for fins which are approximately 10 and 15 nm wide respectively.

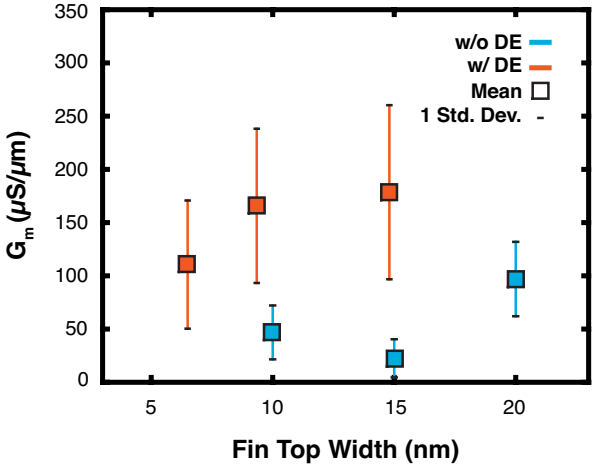


Figure 4.22: Comparison of the extracted maximum transconductance ($g_m$), extracted at $V_d$=0.5 V, as a function of fin width, with and without digital etching (DE).

Fig. 4.23 shows $I_{on}$ as a function of fin width, taken at a gate overdrive of 0.5 V from $I_{off}$=100 nA/$\mu$m. The on-current was found to increase with reducing fin width, despite the degradation in $g_m$, which is a consequence of the superior off-state performance of the narrower devices. Digital etching improves the mean value of $I_{on}$ for devices which are approximately 10 nm wide by a factor of $\sim$4.6.

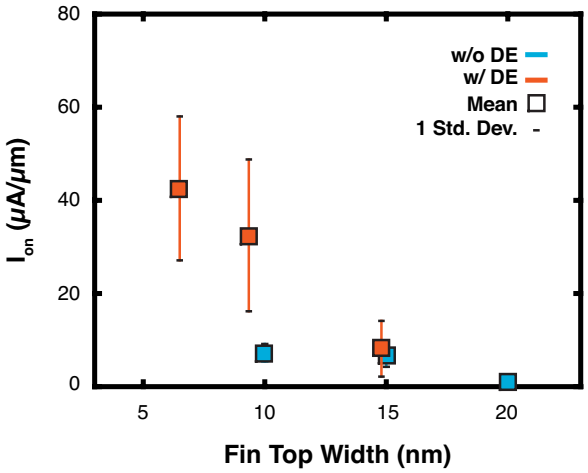


Figure 4.23: Comparison of on-current ($I_{on}$) as a function of fin width, with and without digital etching (DE). On-current was extracted at a gate overdrive of 0.5 V from $I_{off}$=100 nA/$\mu$m at $V_d$=0.5 V.

### 4.6.3 Benchmarking Performance

Figure 4.24 shows the $I_d$-$V_g$ data, in addition to $g_{m,sat}(V_g)$, for the device which exhibited the single highest $I_{on}$ value; a 10 nm device from sample DE. This devices exhibits excellent electrostatic integrity, with a DIBL of 54.4 mV/V and minimum SS values at $V_d$ = 50 mV and 0.5 V of 73 mV/Dec. and 91.1 mV/Dec. respectively. A maximum $g_{m,sat}$ of 387.7 $\mu$S/$\mu$m was measured which, combined with the excellent off-state performance, yields an $I_{on}$ of 80.51 $\mu$A/$\mu$m. It should be noted that there will be a non-negligible error in $I_{on}$ due to the difficulty in discerning the correct length of the gated perimeter, $L_{perim}$, from the TEM image shown in Fig. 4.15(a). Error in $L_{perim}$ results in error in the gate voltage corresponding to $I_{off}$, the magnitude of which is dependant on the local SS at that point. This error propagates to the gate voltage corresponding to $I_{on}$, and the magnitude of the total error in $I_{on}$ is dependant on the local $g_m$ at this point. For benchmarking $I_{on}$, an error of $\pm$25% was assumed in $L_{perim}$, and values for upper and lower bounds of $I_{on}$ calculated accordingly. Fig. 4.25(a) shows $I_{on}$ plotted against the state of the art III-V junctionless devices as a function of gate length, with the channel doping concentration of each device also shown. The on-current of the device fabricated in this work is shown to be competitive with the state of the art at this gate length, when normalised in this manner. The advantage of employing high aspect ratio fins is explicit, however, when $I_{on}$ is normalised to fin width, $W_{fin}$. Commonly in the literature, $I_{on}$ per $W_{fin}$ is calculated subsequent to $I_{on}$ per $L_{perim}$, by multiplying by $L_{perim}/W_{fin}$ [22]. Fig. 4.25(b) re-plots all non-planar devices published with their $I_{on}$ per $L_{perim}$ values re-normalised to $W_{fin}$ in this manner. It should be noted that this overestimates $I_{on}$, as $I_{off}$ is not correctly normalised to $W_{fin}$, however this was done for simplicity given that the raw electrical data for published devices was not available. The data point plotted for the device fabricated in this work was correctly normalised to $W_{fin}$ by also re-normalising its $I_{off}$ value. Despite this overestimation of the performance of devices in the literature, the device fabricated in this work exhibits the highest $I_{on}$ per $W_{fin}$ of any III-V junctionless device demonstrated thus far, with a value of 371.8 $\mu$A/$\mu$m.

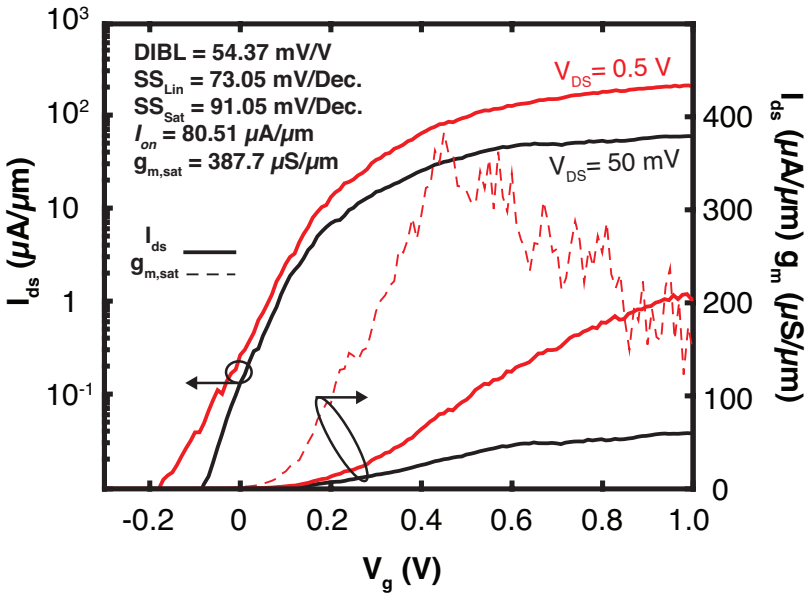


Figure 4.24: Measured transfer characteristics ($I_d$-$V_g$) and transconductance ($g_{m,sat}$) in saturation ($V_d$=0.5 V) as a function of gate voltage for the 10 nm, digitally etched (DE), device with the highest $I_{on}$ value.

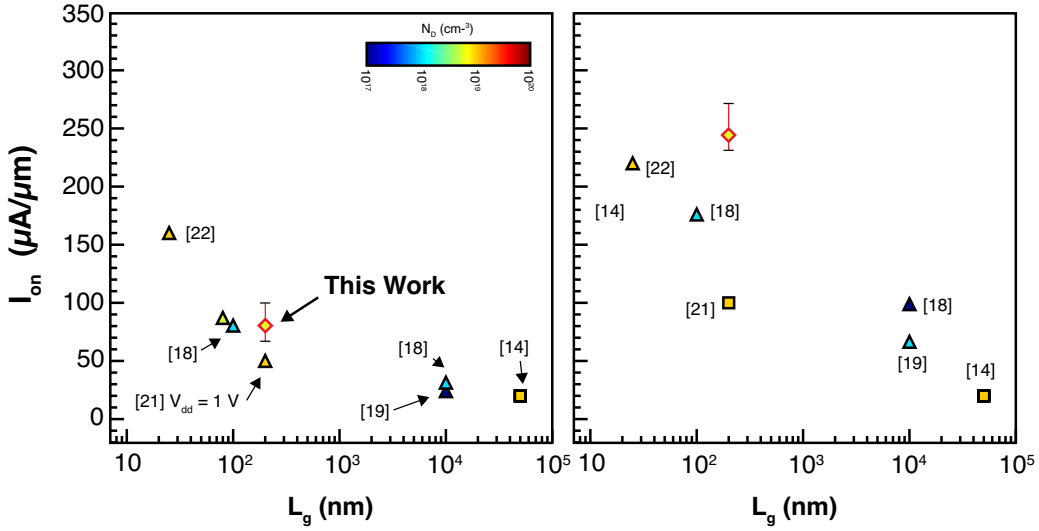


Figure 4.25: Comparison of the on-current ($I_{on}$) achieved in this work to the state of the art III-V junctionless devices reported in the literature, with current normalised by (a) gated perimeter and (b) surface width.

## 4.7 Device Variability

Given the relatively small number of devices for comparison, it is not possible to discern meaningful statistical data from the distribution of metrics and compare them to variability studies which discern the expected distribution due to line edge roughness [12] and RDF [13], for example. The following, however, suggests a variety of sources of variability and discusses how they would fundamentally be expected to impact upon device performance.

### Contamination from the Digital Etch Process

Throughout the development of the digital etch process, a large degree of contamination was observed on the sample surface subsequent to wet etching, an example of which is shown in Fig. 4.26. Efforts were made to identify the source of this contamination, however, it was found that its extent was largely variable form sample to sample, and on occasion was not present at all, making its origin difficult to identify. During the etch process, appropriate glassware (glass beakers) and tweezers (PTFE) were used given the etchant and this is not thought a possible source. It is suspected that it may originate with the exposed InAlAs surface, which is highly reactive and oxides readily in air. Further investigation is required to remove this contamination from the process, which would likely improve both the device yield and variability: any residual contaminant in the gated or contact region of a device would significantly effect the electrical properties of the gate stack contact resistance respectively.

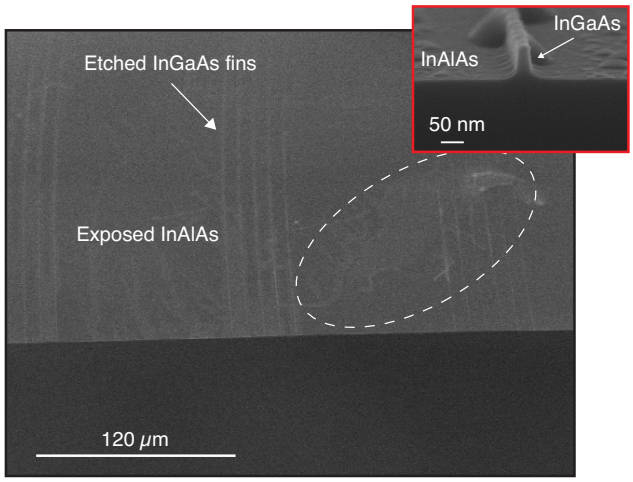


Figure 4.26: Contamination present on the sample surface post digital etch.

### Variability in the Etch Mask Definition and Dry Etch Process

Given the observed differences between the fin geometry of a completed device (shown in Fig. 4.15) and that of the etched fins inspected during the process development (shown in Fig. 4.14), there is clearly a degree of variability in the etch process. Without further TEM images of other completed devices, it is not clear whether it is to be expected that there is significant variability across individual samples, or whether the process is variable from sample to sample. Even with a more robust etch process, there would be some degree of variability in the etch masks themselves in terms of feature size and LER. The extracted device metrics are extremely sensitive to any geometric variability as this effects both on and off-state performance due to the impact on electrostatics and mobility. Furthermore, any variability in the true $L_{\text{perim}}$ of a device impacts the extracted $I_{\text{on}}$ value, as discussed in the previous section.

### Material Composition

A further source of variability could stem from the use of a ternary compound as the channel material as opposed to a binary. When forming narrow fins of a ternary channel material by dry etching, the precise composition of the remaining channel material will likely differ from device to device, and also along the length of the channel. As a consequence, barriers will be introduced into the channel as a result of the locally varying electronic band structures which will degrade the transport properties.

### Random Dopant Fluctuations

Variations in the local concentration of dopant elements - RDF - induce variability in device performance, predominantly in terms of $V_{\text{t}}$ [61]. The impact of RDF is exacerbated for extremely scaled devices as there a fewer dopant elements in the channel, and therefore, individual dopant elements induce a large percentage change [61].

## 4.8 Chapter Summary and Future Work

### 4.8.1 Summary

This chapter presented the design, fabrication, measurement and analysis of high aspect ratio, top-down, InGaAs junctionless FinFETs with fin top widths scaled to less than 10 nm. To enable the fabrication of these devices, a highly anisotropic, low damage, ICP-RIE etch process was developed. The etch process utilised $Cl_2/CH_4/H_2/O_2$ etch chemistry and yielded highly vertical sidewalls ($\sim$86° to the sample surface) despite inducing a low DC bias (111 V). It is proposed that the mechanism of anisotropy is predominantly due to isotropic chemical etching due to the dense $Cl_2$ and $CH_4/H_2$ plasma, which is counterbalanced by surface passivation due to organic polymer deposited by the $CH_4/H_2$ plasma. The physical sputtering component is not thought to play a dominant role.

Devices were fabricated and the impact of fin width and digital etching was assessed on single fin FinFETs, which had a fixed $L_g$ of 200 nm. The fins fabricated were nominally 10 nm, 15 nm and 20 nm wide. As expected, the incorporation of digital etching was found to effectively reduce dry etch induced damage and yielded devices with superior performance compared to those which were not digital etched, in all quantified metrics. The maximum transconductance degraded with decreasing fin width, however, the improvement gained in SS for narrower devices outweighed this degradation, and the 10 nm, digitally etched, FinFETs demonstrated the highest $I_{on}$ of all devices. The single best performing 10 nm device (in terms of $I_{on}$) exhibited an extremely low minimum SS of 73 mV/Dec., a competitive $g_{m,sat}$ of 387.7 $\mu$S/$\mu$, and value of $I_{on}$ normalised to the gated perimeter of 80.51 $\mu$A/$\mu$m.

With $I_{on}$ normalised in the manner above, the on-state performance of the optimal device was competitive with the state of the art at equivalent gate lengths. When $I_{on}$ of the same device is re-normalised by fin base width, which pertains to the current per chip area, it exhibits an on-current value of 371.8 $\mu$A/$\mu$m, which is the highest $I_{on}$ per fin width of any III-V junctionless device published to date.

### 4.8.2 Future Work

Given the target application of low power digital logic, the primary focus of future work should be the scaling of $L_g$ to lengths comparable to current state of the art Si CMOS in order to assess the performance and viability of this technology at the device geometries required for implementation. Secondary to this, should be to target a fully Si CMOS compatible process flow for the reasons discussed at length throughout Chapters 1 and 2. The following elaborates on both

points, discussing a proposed fabrication process flow that could facilitate these investigations and detailing any work done thus far pertaining to these issues.

### Scaled Future Device Generations

Scaling $L_g$ with the existing process flow is problematic: as the S/D regions are patterned using a negative tone resist, the minimum gap length that can be patterned between S/D regions in one lithographic layer is limited by proximity exposure. It would be possible to pattern source and drain regions in separate lithographic layers, however, this limits the gap spacing to 2 $\times$ the alignment accuracy of the EBL tool, which in the case of the Vb6, equates to a minimum gap of $\sim$40 nm. Further, this would likely induce more variability into the process. A self aligned solution would be far preferable. A proposed solution to this is to utilise a combination of positive tone and negative tone resists to pattern the etch mask for S/D and fin regions respectively, as shown in Fig. 4.27. With the fin(s) patterned initially in HSQ as per the existing process, the S/D regions could be patterned subsequently in PMMA which would only require exposure of the resist around the border of the device, and in a region overlapping the fins to determine $L_g$. Using this process, the gate length could be precisely controlled by varying the line width of a single pixel PMMA line, which is capable of a minimum feature size of $\sim$ 10 nm. In practice, 10 nm would likely not be achievable, as the etch resistance of PMMA is poor in comparison to HSQ and thus a significantly thicker layer would be required, which limits the resolution of the CD. It is suggested, however, that this process would enable the realisation of substantially shorter gate lengths than the existing process, with little added complexity.

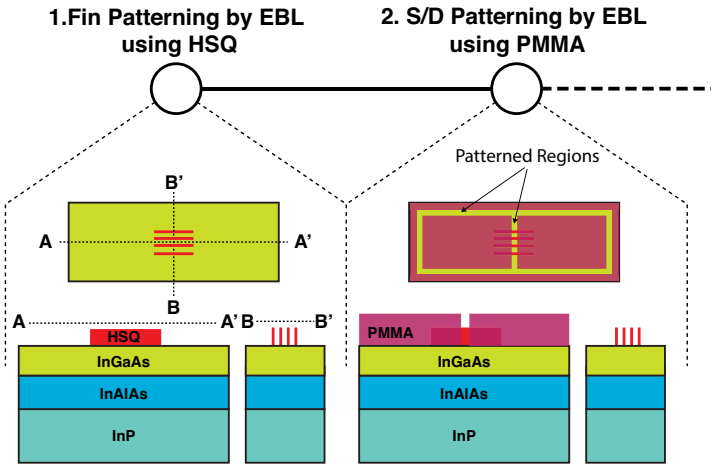


Figure 4.27: Proposed fabrication process flow to from the etch mask for source/drain and fin regions of the device with scaled gate lengths. This could be achieved by using a positive tone resist (PMMA) to from the S/D regions as opposed to a negative tone (HSQ) and therefore removing the limiting factor of proximity exposure.

Additional recommendations for future device generations include:

- **Multiple Fins.** Multiple fins in parallel will be required in order to meet the drive current demands of logic circuits. To this end, future devices should include multiple fins with a scaled fin to fin pitch in order to maximise the current per chip surface width.

- **Self-Aligned S/D Contacts.** Lithographically aligned S/D contacts were used in the first generation of these devices in the interest of process simplicity. Future generations should incorporate self-aligned contacts in order to minimise the access resistance to the channel. There a variety of possible contact technologies and spacer processes that could be implemented to achieve this, such as: annealed Ni-InGaAs contacts with the removal of unreacted Ni in HCl [25]; non-annealed Mo contacts [62]; and self-aligned SiN spacers [63].

- **Scaled Metal Gate Length.** As with the use of lithographically aligned contacts, lift-off was used to deposit the gate metal in the interest of processing simplicity. Future devices should implement a subtractive process to from the gate; where a blanket layer of metal is deposited on the sample and selected regions are subsequently etched. This would allow the overlap between the gate contact and S/D regions to be minimised, even in the case of an extremely scaled $L_g$. It is not thought that the large overlapped regions in the existing devices significantly impacts device performance, as the gate cannot fully depleted regions of the device other than the fin. It is possible, however, that the overlapped regions aid the low off-current exhibited by existing devices, and it is important, therefore, that devices which more closely resemble those used in VLSI, with minimal overlap regions in order to minimise the device footprint, be investigated. An etch process previously developed within the group by Dr Menglin Cao for dry etching Mo selectively to $Al_2O_3$ could be used for this purpose [64]. Furthermore, the use of a such process as opposed to lift-off pertains to Si CMOS compatibility.

- **Lateral Gate All Around Nanowires.** Lateral nanowires could be formed by selectively wet etching the InAlAs buffer prior to the formation of the stack. This would improve the electrostatic integrity of the device, which may be required in order to minimise SCEs at more scaled gate lengths.

- **Doping Concentration.**. Increasing the doping concentration can increase the current of a junctionless device and reduce contact resistance, however, it also degrades the electron transport properties due to increased scattering, and requires more confined channel geometries such that the devices can be full depleted in order to turn off, which in turn may further degrade the transport properties due to mobility degradation. Various doping concentrations should be investigated to find the optimal trade off between the above parameters.

### A Fully Si CMOS Compatible Process

With regards to to yielding a fully Si CMOS compatible junctionless process, InGaAs XOI wafers have been grown by Dr Lukas Czornomaz in IBM Zurich as part of the Compose3 collaboration between, among others, IBM and the University of Glasgow. The wafers were formed by DWB and have the same layer structure as per Ref. [65]. Two wafers were grown with doping concentrations of $6\times10$ cm$^{-3}$ and $1\times10$ cm$^{-3}$. The existing process is compatible with these new wafers and the addition of a substractive metal gate process, as discussed above, would yield a fully Si CMOS compatible process.

# Bibliography


[1] S. M. Sze and K. K. NG, *Physics of Semiconductor Devices*, 3rd ed. Wiley, 2006.

[2] J.-P. Colinge, C.-W. Lee, A. Afzalian, N. D. Akhavan, R. Yan, I. Ferain, P. Razavi, B. O'Neill, A. Blake, M. White, A.-M. Kelleher, B. McCarthy, and R. Murphy, "Nanowire transistors without junctions," *Nature Nanotechnology*, vol. 5, no. 3, pp. 225–229, 2010.

[3] L. Edgar, "Device for controlling electric current," US Patent, US1 900 018A, 03 28, 1933.

[4] A. N. Nazarov, I. Ferain, N. D. Akhavan, P. Razavi, R. Yu, and J. P. Colinge, "Random telegraph-signal noise in junctionless transistors," *Applied Physics Letters*, vol. 98, no. 9, pp. 1–4, 2011.

[5] A. Gnudi, S. Reggiani, E. Gnani, and G. Baccarani, "Analysis of threshold voltage variability due to random dopant fluctuations in junctionless FETs," *IEEE Electron Device Letters*, vol. 33, no. 3, pp. 336–338, 2012.

[6] C.-W. Lee, A. Afzalian, N. D. Akhavan, R. Yan, I. Ferain, and J. P. Colinge, "Junctionless multigate field-effect transistor," *Applied Physics Letters*, vol. 94, no. 5, pp. 4–6, 2009.

[7] M. M. Mirza, D. A. Maclaren, A. Samarelli, B. M. Holmes, H. Zhou, S. Thoms, D. Macintyre, and D. J. Paul, "Determining the electronic performance limitations in top-down-fabricated Si nanowires with mean widths down to 4 nm," *Nano Letters*, vol. 14, no. 11, pp. 6056–6060, 2014.

[8] "Si Electrical Properties", IOFFE, 2018. [Online]. Available: http://www.ioffe.ru/SVA/NSM/Semicond/Si/electric.html

[9] "InGaAs Electrical Properties", IOFFE, 2018. [Online]. Available: http://www.ioffe.ru/SVA/NSM/Semicond/GaInAs/electric.html

[10] T. Orzali, A. Vert, R. T. P. Lee, A. Norvilas, G. Huang, J. L. Herman, R. J. W. Hill, and S. S. P. Rao, "Heavily tellurium doped n-type InGaAs grown by MOCVD on 300 mm Si wafers," *Journal of Crystal Growth*, vol. 426, pp. 234–247, 2015.

[11] U. Singisetti, M. A. Wistey, G. J. Burek, A. K. Baraskar, B. J. Thibeault, A. C. Gossard, M. J. W. Rodwell, B. Shin, E. J. Kim, P. C. McIntyre, B. Yu, Y. Yuan, D. Wang, Y. Taur, P. Asbeck, and Y. Lee, "$In_{0.53}Ga_{0.47}As$ Channel MOSFETs With Self-Aligned InAs Source/Drain Formed by MEE Regrowth," *IEEE Electron Device Letters*, vol. 30, no. 11, pp. 1128–1130, 2009.

[12] G. Leung, A. Pan, and C. O. Chui, "Junctionless Silicon and $In_{0.53}Ga_{0.47}As$ Transistors – Part II: Device Variability From Random Dopant Fluctuation," *IEEE Tranactions on Electron Devices*, vol. 62, no. 10, pp. 3208–3214, 2015.

[13] N. Zagni, F. M. Puglisi, G. Verzellesi, and P. Pavan, "Combined variability/sensitivity analysis in III-V and silicon FETs for future technological nodes," in *IEEE International Reliability Physics Symposium (IRPS)*, 2017, pp. XT–4.1–XT–4.5.

[14] M. Yokoyama, R. Iida, S. H. Kim, N. Taoka, Y. Urabe, T. Yasuda, H. Takagi, H. Yamada, N. Fukuhara, M. Hata, M. Sugiyama, Y. Nakano, M. Takenaka, and S. Takagi, "Extremely-thin-body ingaas-on-insulator mosfets on si fabricated by direct wafer bonding," in *IEEE International Electron Devices Meeting (IEDM)*, 2010, pp. 3.1.1–3.1.4.

[15] K. H. Goh, Y. Guo, X. Gong, G. Liang, and Y. Yeo, "Near ballistic sub-7 nm junctionless fet featuring 1 nm extremely-thin channel and raised s/d structure," in *IEEE International Electron Devices Meeting (IEDM)*, 2013, pp. 16.5.1–16.5.4.

[16] K. H. Goh and Y. Yeo, "Novel short-channel $In_{0.53}Ga_{0.47}As$ junctionless nanowire nFET with raised s/d structure: An ultimately scaled 1-D transistor architecture," in *Silicon Nanoelectronics Workshop (SNW)*, June 2014, pp. 1–2.

[17] K. Goh, S. Yadav, K. L. Low, G. Liang, X. Gong, and Y. Yeo, "Gate-All-Around $In_{0.53}Ga_{0.47}As$ Junctionless Nanowire FET With Tapered Source/Drain Structure," *IEEE Transactions on Electron Devices*, vol. 63, no. 3, pp. 1027–1033, 2016.

[18] V. Djara, K. Cherkaoui, T. Lopez, É. O'Connor, I. M. Povey, K. K. Thomas, and P. K. Hurley, "Junctionless InGaAs MOSFETs with InAlAs barrier isolation and channel thinning by digital wet etching," in *71st Device Research Conference (DRC)*, 2013, pp. 131–132.

[19] T. Irisawa, M. Oda, K. Ikeda, Y. Moriyama, E. Mieda, W. Jevasuwan, T. Maeda, O. Ichikawa, T. Ishihara, M. Hata, and T. Tezuka, "High mobility p-n junction-less InGaAs-OI tri-gate nMOSFETs with metal source/drain for ultra-low-power CMOS applications," in *IEEE International SOI Conference (SOI)*, 2012, pp. 1–2.

[20] V. Djara, L. Czornomaz, N. Daix, D. Caimi, V. Deshpande, E. Uccelli, M. Sousa, C. Marchiori, and J. Fompeyrine, "Tri-gate $In_{0.53}Ga_{0.47}As$-on-insulator junctionless field effect

transistors," in *EUROSOI-ULIS 2015: Joint International EUROSOI Workshop and International Conference on Ultimate Integration on Silicon*, 2015, pp. 97–100.

[21] H. X. Guo, X. Zhang, Z. Zhu, E. Y. J. Kong, and Y. C. Yeo, "Junctionless Pi-gate transistor with indium gallium arsenide channel," *Electronics Letters*, vol. 49, no. 6, pp. 402–403, 2013.

[22] C. B. Zota, M. Borg, L. Wernersson, and E. Lind, "Record performance for junctionless transistors in ingaas mosfets," in *Symposium on VLSI Technology*, 2017, pp. T34–T35.

[23] Y. Song, C. Zhang, R. Dowdy, K. Chabak, P. K. Mohseni, W. Choi, and X. Li, "III-V junctionless gate-all-around nanowire MOSFETS for high linearity low power applications," *IEEE Electron Device Letters*, vol. 35, no. 3, pp. 324–326, 2014.

[24] É. O'Connor, B. Brennan, V. Djara, K. Cherkaoui, S. Monaghan, S. B. Newcomb, R. Contreras, M. Milojevic, G. Hughes, M. E. Pemble, R. M. Wallace, and P. K. Hurley, "A systematic study of $(NH_4)_2S$ passivation (22%, 10%, 5%, or 1%) on the interface properties of the $Al_2O_3/In_{0.53}Ga_{0.47}As$/InP system for n-type and p-type $In_{0.53}Ga_{0.47}As$ epitaxial layers," *Journal of Applied Physics*, vol. 109, no. 2, pp. 0–10, 2011.

[25] S. Kim, M. Yokoyama, N. Taoka, R. Iida, S. Lee, R. Nakane, Y. Urabe, N. Miyata, T. Yasuda, H. Yamada, N. Fukuhara, M. Hata, M. Takenaka, and S. Takagi, "Self-Aligned Metal Source/Drain $In_xGa_{1-x}As$ n-Metal-Oxide-Semiconductor Field-Effect Transistors Using Ni-InGaAs Alloy," *Applied Physics Express*, vol. 4, no. 2, p. 024201, 2011.

[26] A. Vardi, W. Lu, X. Zhao, and J. A. Del Alamo, "Nanoscale Mo Ohmic contacts to III-V Fins," *IEEE Electron Device Letters*, vol. 36, no. 2, pp. 126–128, 2015.

[27] G. Snider. (2017) 1D Poisson Solver. [Online]. Available: https://www3.nd.edu/ gsnider/

[28] X. Wang, A. R. Brown, B. Cheng, and A. Asenov, "Statistical variability and reliability in nanoscale finfets," in *IEEE International Electron Devices Meeting (IEDM)*, 2011, pp. 5.4.1–5.4.4.

[29] N. Agrawal, Y. Kimura, R. Arghavani, and S. Datta, "Impact of Transistor Architecture (Bulk Planar, Trigate on Bulk, Ultrathin-Body Planar SOI) and Material (Silicon or III-V Semiconductor) on Variation for Logic and SRAM Applications," *IEEE Transactions on Electron Devices*, vol. 60, no. 10, pp. 3298–3304, 2013.

[30] X. Zhao, J. Lin, C. Heidelberger, E. A. Fitzgerald, and J. A. del Alamo, "Vertical nanowire InGaAs MOSFETs fabricated by a top-down approach," in *IEEE International Electron Devices Meeting (IEDM)*, 2013, pp. 28.4.1–28.4.4.

[31] U. Peralagu, X. Li, O. Ignatova, M. Steer, I. Povey, P. Hurley, and I. Thayne, "Demonstration of III-V fins with vertical sidewalls using $Cl_2/CH_4/H_2/O_2$ dry etch chemistry in conjunction with digital etching for recovery of etch damage." in *45th IEEE Semiconductor Interface Spcialists Conference (SISC)*, 2014.

[32] 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, pp. 2009–2012, 2011.

[33] S. Monaghan, A. O'Mahony, K. Cherkaoui, E. OConnor, I. M. Povey, M. G. Nolan, D. O'Connell, M. E. Pemble, P. K. Hurley, G. Provenzano, F. Crupi, and S. B. Newcomb, "Electrical analysis of three-stage passivated $In_{0.53}Ga_{0.47}As$ capacitors with varying $HfO_2$ thicknesses and incorporating an $Al_2O_3$ interface control layer," *Journal of Vacuum Science & Technology B*, vol. 29, no. 1, pp. 01A807–8, 2011.

[34] "PENELOPE (Penetration and ENErgy LOss of Positrons and Electrons)", SourceForge, 2018. [Online]. Available: http://pypenelope.sourceforge.net/

[35] S. Thoms and J. McGeough, *Micromachining of Engineering Materials*, 1st ed. Marcel Dekker, Inc., 2002.

[36] D.-H. Kim, S.-K. Kang, G.-Y. Yeom, and J.-H. Jang, "Nanometer-scale fabrication of hydrogen silsesquioxane (hsq) films with post exposure baking," *Journal of Nanoscience and Nanotechnology*, vol. 13, no. 3, pp. 1918–1922, 2013.

[37] W. Henschel, Y. M. Georgiev, and H. Kurz, "Study of a high contrast process for hydrogen silsesquioxane as a negative tone electron beam resist," *Journal of Vacuum Science & Technology B*, vol. 21, no. 5, p. 2018, 2003.

[38] V. A. Guzenko, B. Pedrini, A. Menzel, and C. David, "Fabrication of nanoparticles with 3D shape control for X-ray scattering experiments," *Microelectron. Eng.*, vol. 121, pp. 127–130, 2014.

[39] X. Zhao, S.-Y. Lee, J. Choi, S.-H. Lee, I.-K. Shin, and C.-U. Jeon, "Minimization of line edge roughness and critical dimension error in electron-beam lithography," *J. Vac. Sci. Technol. B, Nanotechnol. Microelectron. Mater. Process. Meas. Phenom.*, vol. 32, no. 6, p. 06F505, 2014.

[40] I. Thayne, R. Hill, A. Asenov, R. Droopad, and M. Passlack, "Review of Current Status of III-V MOSFETs," *Meet. Electrochem. Soc.*, vol. 215, no. 5, p. 1329, 2009.

[41] S. Murad, M. Rahman, N. Johnson, and S. Thoms, "Dry etching damage in III-V semiconductors," *Journal of Vacuum Science & Technology B*, vol. 14, no. 6, p. 3658, 1996.

[42] K. Nojiri, *Dry Etching Technology for Semiconductors*, 1st ed. Springer, 2015.

[43] K. Asakawa, T. Yoshikawa, S. Kohmoto, Y. Nambu, and Y. Sugimoto, "Chlorine-based dry etching of III-V compound semiconductors for optoelectronic application," *Japanese Journal of Applied Physics*, vol. 37, no. 2, pp. 373–387, 1998.

[44] T. Maeda, J. Lee, R. Shul, J. Han, J. Hong, E. Lambers, S. Pearton, C. Abernathy, and W. Hobson, "Inductively coupled plasma etching of III-V semiconductors in $BCl_3$-based chemistries: I. GaAs, GaN, GaP, GaSb and AlGaAs," *Applied Surface Science*, vol. 143, pp. 174–182, 1999.

[45] T. R. Hayes, "Reactive ion etching of InP using $CH_4/H_2$ mixtures: Mechanisms of etching and anisotropy," *Journal of Vacuum Science & Technology B*, vol. 7, no. 1989, p. 1130, 1989.

[46] N. Haneji, G. Segami, T. Ide, T. Suzuki, T. Arakawa, K. Tada, Y. Shimogaki, and Y. Nakano, "Electron Cyclotron Resonance-Reactive Ion Etching of III-V Semiconductors by Cyclic Injection of $CH_4/H_2$/Ar and $O_2$ with Constant Ar Flow," *Japanese Journal of Applied Physics*, vol. 42, no. Part 1, No. 6B, pp. 3958–3961, 2003.

[47] M. Gross, "Modeling of sloped sidewalls formed by simultaneous etching and deposition," *Journal of Vacuum Science & Technology B*, vol. 7, no. 1989, p. 534, 1989.

[48] C. Constantine, C. Barratt, S. J. Pearton, F. Ren, and J. R. Lothian, "Smooth, low-bias plasma etching of InP in microwave Cl2/CH 4/H2 mixtures," *Applied Physics Letters*, vol. 61, no. 24, pp. 2899–2901, 1992.

[49] X. Zhao and J. A. Del Alamo, "Nanometer-scale vertical-sidewall reactive ion etching of ingaas for 3-D III-V MOSFETs," *IEEE Electron Device Letters*, vol. 35, no. 5, pp. 521–523, 2014.

[50] U. Peralagu, X. Li, O. Ignatova, Y. Fu, D. Millar, M. Steer, I. Povey, K. Hossain, M. Jain, T. Golding, R. Droopadd, P. Hurley, and I. Thayne, "Towards a Vertical and Damage Free Post-Etch InGaAs Fin Profile: Dry Etch Processing, Sidewall Damage Assessment and Mitigation Options," *ECS Transactions*, vol. 69, no. 5, pp. 15–36, 2015.

[51] J. R. Lothian, F. Ren, and S. J. Pearton, "Mask erosion during dry etching of deep features in III-V semiconductor structures," *Semiconductor Science and Technology*, vol. 7, pp. 1199–1209, 1999.

[52] Z. Ji, X. Zhang, J. Franco, R. Gao, M. Duan, J. F. Zhang, W. D. Zhang, B. Kaczer, A. Alian, D. Linten, D. Zhou, N. Collaert, S. D. Gendt, and G. Groeseneken, "An Investigation on Border Traps in III-V MOSFETs," *IEEE Transactions on Electron Devices*, vol. 62, no. 11, pp. 3633–3639, 2015.

[53] T. Grasser, H. Reisinger, P. J. Wagner, and B. Kaczer, "Time-dependent defect spectroscopy for characterization of border traps in metal-oxide-semiconductor transistors," *Physical Review B*, vol. 82, no. 24, pp. 1–10, 2010.

[54] J. P. Colinge and A. Chandrakasan, *FinFETs and other multi-gate transistors*. Springer, 2008.

[55] A. Vais, J. Franco, D. Lin, V. Putcha, S. Sioncke, A. Mocuta, N. Collaert, A. Thean, and K. De Meyer, "On the distribution of oxide defect levels in Al2O3 and HfO2 high-k dielectrics deposited on InGaAs metal-oxide-semiconductor devices studied by capacitance-voltage hysteresis," *J. Appl. Phys.*, vol. 121, no. 14, pp. 3–10, 2017.

[56] R. D. Trevisoli, R. T. Doria, M. de Souza, and M. A. Pavanello, "Threshold voltage in junctionless nanowire transistors," *Semiconductor Science and Technology*, vol. 26, no. 10, p. 105009, 2011.

[57] H. J. Cho, J. D. Choe, M. Li, J. Y. Kim, S. H. Chung, C. W. Oh, E.-J. Yoon, D.-W. Kim, D. Park, and K. Kim, "Fin width scaling criteria of body-tied finfet in sub-50 nm regime," in *Device Research Conference, 2004. 62nd DRC. Conference Digest [Includes 'Late News Papers' volume]*, 2004, pp. 209–210 vol.1.

[58] A. Vardi and J. A. Del Alamo, "Sub-10-nm Fin-Width Self-Aligned InGaAs FinFETs," *IEEE Electron Device Letters*, vol. 37, no. 9, pp. 1104–1107, 2016.

[59] D. K. Schroder, *Semiconductor Material and Device Characterization: Third Edition*. Wiley, 2005.

[60] A. Ortiz-Conde, F. J. García Sánchez, J. J. Liou, A. Cerdeira, M. Estrada, and Y. Yue, "A review of recent MOSFET threshold voltage extraction methods," *Microelectronics Reliability*, vol. 42, no. 4-5, pp. 583–596, 2002.

[61] A. Asenov, "Random dopant induced threshold voltage lowering and fluctuations in sub-0.1$\mu$m MOSFET's: A 3-D "atomistic" simulation study," *IEEE Transactions on Electron Devices*, vol. 45, no. 12, pp. 2505–2513, 1998.

[62] A. Vardi, J. Lin, W. Lu, X. Zhao, A. Fernando-Saavedra, and J. A. del Alamo, "A si-compatible fabrication process for scaled self-aligned ingaas finfets," *IEEE Transactions on Semiconductor Manufacturing*, vol. 30, no. 4, pp. 468–474, Nov 2017.

[63] X. Li., R. Hill, H. Zhou, C. Wilkinson, and I. Thayne, "A low damage SiN sidewall spacer for self-aligned sub 100 nm III-V MOSFETs," *Microelectronic Engineering*, vol. 85, no. 5-6, pp. 996–999, 2008.

[64] M. Cao, X. Li, and I. Thayne, "An Anisotropic Low Power, Low DC Bias, $SF_6/C_4F_8$ Inductively Coupled Plasma Etch Process of Molybdenum with Critical Dimension of 30 nm Suitable for Compound Semiconductor Devices," in *8th International Conference on Electron, Ion, and Photon Beam Technology and Nanofabrication (EIPBN)*, 2014.

[65] L. Czornomaz, N. Daix, D. Caimi, M. Sousa, R. Erni, M. D. Rossell, M. El-Kazzi, C. Rossel, C. Marchiori, E. Uccelli, M. Richter, H. Siegwart, and J. Fompeyrine, "An integration path for gate-first utb iii-v-on-insulator mosfets with silicon, using direct wafer bonding and donor wafer recycling," in *IEEE International Electron Devices Meeting (IEDM)*, 2012, pp. 23.4.1–23.4.4.

# Chapter 5

# Results II - In-situ $H_2$ Plasma Surface Cleaning of InGaSb

## 5.1 Background

The excellent performance demonstrated by scaled, III-V, n-type MOSFETs has thus far remained elusive for equivalent complementary devices. To date, Ge remains the most technologically mature, high mobility candidate to replace Si PMOS, however, conversely, Ge NMOS exhibits poor drive currents and low mobility. This has prompted research into hybrid CMOS technologies which combine III-V NMOS devices with Si/Ge PMOS. Such devices require complex fabrication processes, however, in order to co-integrate the different material systems onto a common substrate, and to concurrently fabricate CMOS devices where each device polarity requires the formation of different gate stacks, utilising different surface cleaning processes with different thermal budgets.

Unlike III-V aresenide compounds, antimonides exhibit excellent transport properties for both electrons and holes, with InSb and GaSb having bulk electron/hole mobilities of 77,000 [1]/ 1,000 $cm^2V^{-1}s^{-1}$ [2] respectively, as summarised in Fig. 5.1. $In_xGa_{1-x}Sb$ ternary compounds have been shown to offer the combined optimal performance for electrons and holes in the same material: the incorporation of In maintains excellent electron transport [3] while room temperature hole mobilities as high as 1,500 $cm^2\ V^{-1}\ s^{-1}$ have been demonstrated in strained p-$In_{0.4}Ga_{0.6}Sb$ quantum wells [4]. As such, CMOS devices which have a common channel material of $In_xGa_{1-x}Sb$ have the potential to not only outperform Si CMOS, but to do so while having greatly reduced fabrication complexity in comparison to hybrid technologies, specifically if a common gate stack was applicable to both device polarities. Of course, identifying a material with high bulk mobilities for both electrons and holes does not necessarily translate to

the fabrication of high performance complementary devices, as evidenced by Ge, whose poor NMOS performance is predominantly attributed to gate stack non-idealities, such as high $D_{it}$ near the conduction band edge, among others. A critical enabling technology therefore in order to full exploit the bulk transport properties of the $In_xGa_{1-x}Sb$ material system, is the ability to form a low $D_{it}$, high-k, unpinned dielectric interface.

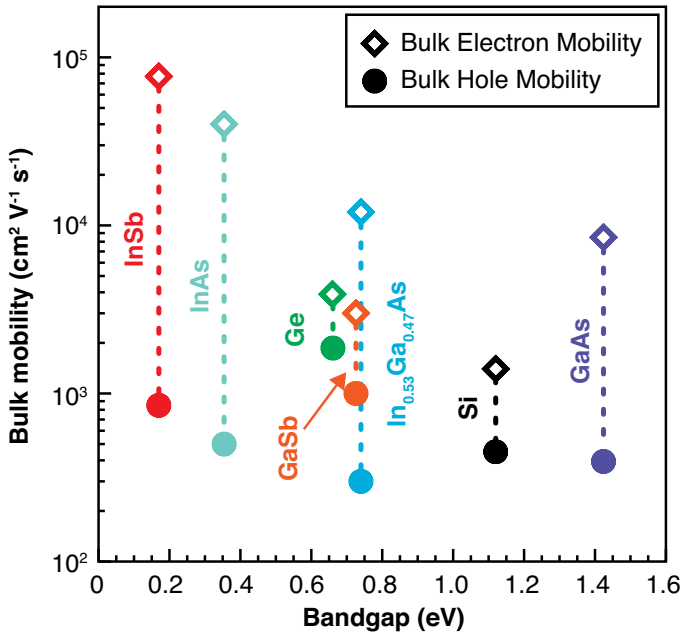


Figure 5.1: Bulk electron/hole mobility plotted against bandgap energy for Si, Ge, and a variety of III-V compounds.

### 5.1.1 Antimonide Surface pre-treatments on GaSb and InSb

To date, all InGaSb devices demonstrated utilise buried channels [3, 5–13] which fundamentally limits the scaling potential due to gating through a capping layer; being able gate the InGaSb channel directly would circumvent this limitation. Systematic studies on improving the electrical properties of the dielectric interface to antimonides have been limited to GaSb [8–10, 12, 14–22] and InSb only [23–30], however. For the former, ex-situ hydrochloric acid (HCl) [8–10, 13, 16] and ammonium sulfide $(NH_4)_2S$ [14, 17, 18] surface treatments, as well as in-situ $H_2$ plasma exposure [19–22] have yielded promising results.

## 5.2 Chapter Outline

This chapter presents the results of the first investigation into the impact of in-situ $H_2$ plasma exposure on the electrical properties of the $In_{0.3}Ga_{0.7}Sb$-$Al_2O_3$ interface. The epitaxial layer structures of the samples used throughout these experiments are detailed in Section 5.3. The parameter space was initially explored in terms of varying $H_2$ plasma power; a set of experiments deemed the 'Power Series'. This is detailed in Section 5.4, which further includes an assessment of the impact of both in-situ TMA pre-cleaning and a post gate metal FGA, as well as chemical analysis of the dielectric interfaces of selected samples via energy dispersive X-ray spectroscopy (EDX) and X-ray photoelectron spectroscopy (XPS) measurements. Given the well documented difficulties in correctly extracting quantitative metrics for interfaces to narrow band gap materials [31, 32], the impact of the various surface treatments on the electrical properties of the interfaces were assessed qualitatively using multifrequency CV measurements following the fabrication of MOSCAPs. These measurements were taken both at room temperature and at 30 K, and were characterised in terms of the following metrics: capacitance modulation, $C_{\text{mod}}$ [19], stretch out [33], and frequency dispersion in accumulation [34], which pertain to freedom of Fermi level ($E_{\text{F}}$) movement [35], magnitude of $D_{\text{it}}$ [33], and magnitude of $N_{\text{BT}}$ [34] respectively. The optimal process in terms of these metrics was subsequently employed for the investigation into $H_2$ plasma exposure time, deemed the 'Time Series', which is discussed in Section 5.5. Samples which were subjected to the optimal process overall in terms of the above metrics latterly underwent a more comprehensive analysis, including the use of a novel impedance spectroscopy method which has been demonstrated to discern genuine surface inversion from false inversion due to $D_{\text{it}}$ [36], as well as the extraction of the activation energy of minority carriers at gate biases corresponding nominally to inversion, both of which are also presented in Section 5.5. Sections 5.6-5.9 present the simulation of the InGaSb-$Al_2O_3$ interface via an equivalent circuit model which has been previously utilised to quantitatively assess dielectric interfaces to InGaAs and InAs [37, 38]. This the most robust method known for extracting $D_{\text{it}}$, $N_{\text{BT}}$, and the gate voltage-surface potential relationship ($V_{\text{g}}$-$\psi_{\text{s}}$) for narrowband gap MOS systems.

## 5.3 Material Growth and Characterisation

$In_{0.3}Ga_{0.7}Sb$ epitaxial layers were grown by MBE on GaAs (100) substrates. An InSb mole fraction of 30 % was chosen as simulations have shown mole fractions between 20-40 % to offer the maximum drive current for n-type devices [3]: increasing the In concentration increases $v_{inj}$ and decreases the DOS, and mole fractions between 20-40 % yield the optimal trade off between these two parameters [3]. The complete layer structure of the wafer used for the Power Series is shown in Fig. 5.2 (a) and comprised: a SI GaAs (100) substrate; 250 nm GaAs re-growth and a 200 nm GaSb relaxed buffer layer, both of which were unintentionally doped; a 3 $\mu$m $In_{0.3}Ga_{0.7}Sb$ buffer layer, uniformly doped at a nominal value of $1 \times 10^{18}$ cm$^{-3}$; and a 500 nm $In_{0.3}Ga_{0.7}Sb$ capacitor layer, uniformly doped at a nominal value of $2 \times 10^{17}$ cm$^{-3}$. Both p and n-type variants were grown, for which beryllium and tellurium were the dopant elements respectively. A SI substrate was used due to availability at the time of growth, and consequently required the formation of a top ohmic contact for MOSCAP fabrication, as discussed in Section 5.4.2. The wafer used for the Time Series was grown subsequently and employed a doped substrate and doped buffer layers to facilitate simpler MOSCAP fabrication via a back ohmic contact. For these samples, the $In_{0.3}Ga_{0.7}Sb$ capacitor layer had a nominal doping concentration of $1 \times 10^{17}$ cm$^{-3}$.

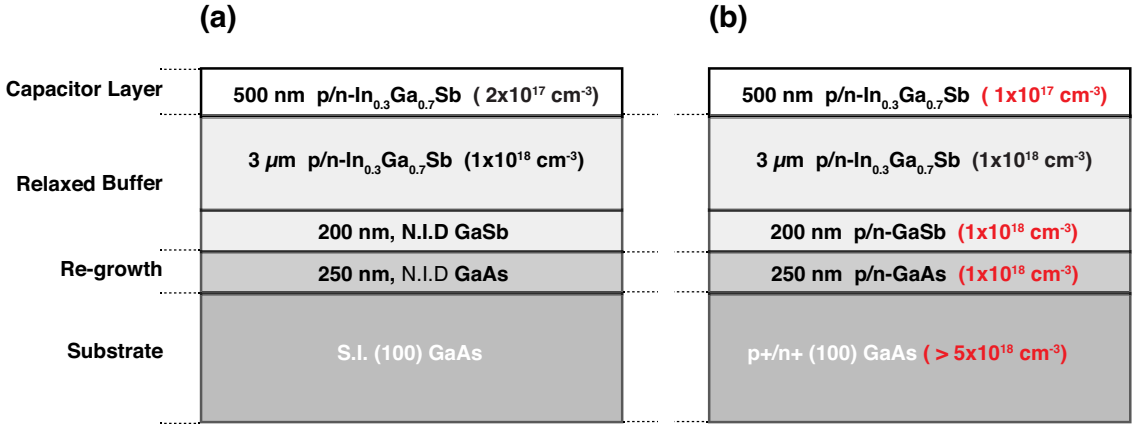


Figure 5.2: Schematic diagram of the epitaxial layer structures grown by molecular beam epitaxy (MBE) for the investigation into the effect of (a) plasma power and (b) exposure time. The wafer shown in (a) employed a semi-insulating substrate due to availability at the time of growth. The wafer in (b), which was grown subsequently to (a), utilised a doped substrate with doped buffer layers to facilitate a simpler metal oxide semiconductor capacitor fabrication process, as discussed in Section 5.5.

AFM measurements of the surface of the various grown wafers, taken using a Bruker Icon AFM in tapping mode and shown in Fig. 5.3(a-c) for the n-type Time Series wafer, reveal the surface to comprise mounds which are on the order of $\sim 6$ $\mu$m in length and width, 15-25 nm in height, and at an approximate area density of $12.5 \times 10^6$ cm$^{-2}$. These are thought to originate

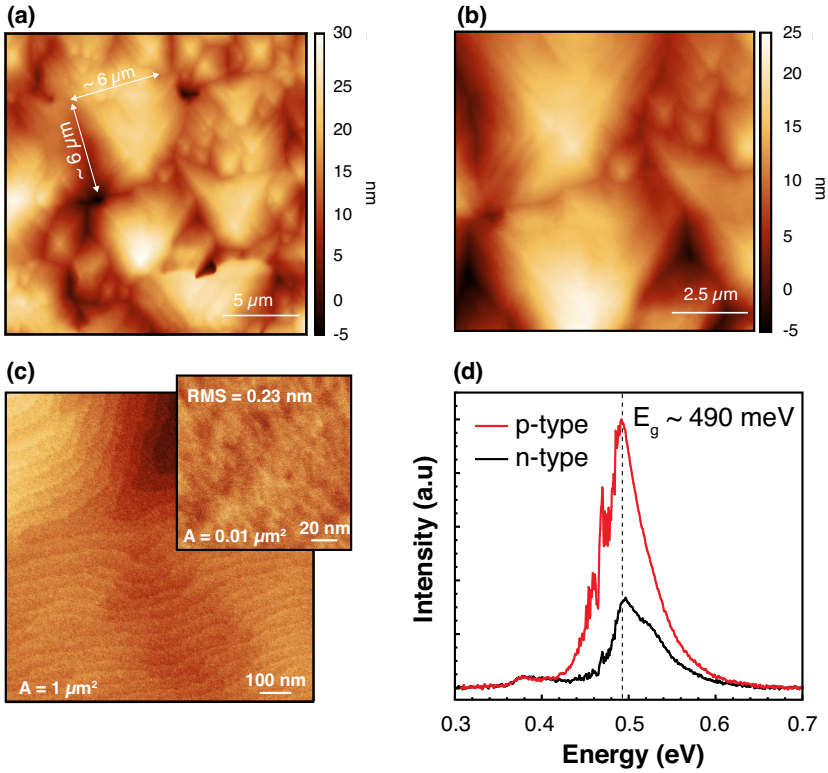


Figure 5.3: (a-c) Atomic force microscopy (AFM) scans of the surface of the n-type Time Series wafer illustrating large mounds (a,b) thought to originate from dislocations between epitaxial layers. The small area scan shown in (c) illustrates the exceptionally smooth surface of these mounds, with $r_q$ measured as 0.23 nm. The surfaces of all other wafers grown were equivalent. (d) Photoluminescence measurements of p and n-type samples from the Power Series wafers showing their band gap to be ∼490 meV.

from dislocations formed at the interfaces between the various heterolayers and are commonly observed on MBE grown antimonide surfaces. Smaller area scans show the surface of these mounds to be incredibly smooth however, with mono layer step edges visible in (c) and an $r_q$ of 0.23 nm measured over an area of 100 nm×100 nm (shown inset).

The band gap of the samples was measured by Dr Ross Millar using photoluminescence (PL) measurements taken using a diode pumped solid state laser emitting at 532 nm wavelength. Sample emission was collected by gold parabolic mirrors and coupled into a Bruker Vertex 70 Fourier-transform infrared spectroscopy (FTIR) system with a liquid nitrogen cooled InSb detector. The pump is rejected by a 1.65 $\mu$m long-pass filter. In order to remove the contribution of ambient blackbody radiation from the room, the tool was operated in 'step-scan' mode. For this measurement, the pump is modulated by an optical chopper, and the detector output is passed through a lock-in amplifier, which discerns signals at the chopper frequency, therefore removing the ambient radiation from the measurement. For both p and n-type samples, the band gap was measured as ∼490 meV, which corresponds an InSb mole fraction of ∼0.275.

## 5.4 Power Series

### 5.4.1 Parameter Space

The experimental matrix for the Power Series is summarised in Fig. 5.4. Prior to plasma exposure, all samples were subjected to an ex-situ HCl surface clean and subsequently loaded into a central vacuum load lock, which is part of a clustered ICP-RIE and ALD tool. The samples were transferred to the load lock in less than 1 minute. Samples were exposed to the $H_2$ plasma in the ICP-RIE chamber with the following fixed parameters: $H_2$:Ar (1:7) plasma chemistry, 90 mT chamber pressure, 30 minute exposure time, 150 °C platen temperature, and 2 W platen power. The exposure time and temperature were based on promising results on GaSb [19]. The platen power was fixed at 2 W to minimise the DC self bias and thus minimise any damage to the semiconductor surface due to sputtering. Six ICP powers were investigated: 150, 250, 500, 1000, 1500 and 2000 W. Additionally a control sample, subsequently referred to as 0 W, was included which had no plasma exposure. Following $H_2$ plasma treatment, samples were transferred, under vacuum, to the ALD chamber where $Al_2O_3$ was deposited via a thermal process at 200°C, comprising 80 alternating cycles of TMA and water ($H_2O$) exposure. Further samples were processed with identical parameters but for the inclusion of in-situ TMA exposure in the ALD chamber immediately prior to $Al_2O_3$ deposition, which has demonstrated the reduction of native oxides or 'self-cleaning' on GaSb [20] and other III-Vs [39–41].

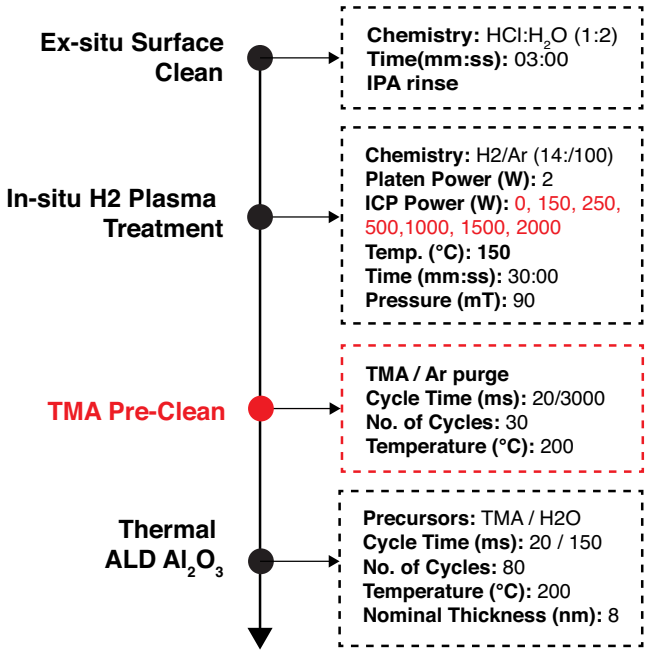


Figure 5.4: Experimental matrix for samples processed as part of the Power Series. The parameters which were varied are highlighted in red. A total of 14 samples were processed corresponding to the 7 ICP powers highlighted, processed with and without TMA pre-cleaning.

### 5.4.2 MOSCAP Fabrication

MOSCAPs for the Power Series were fabricated as per the process flow shown in Fig. 5.5. The gate metal (20 nm Pt/ 200 nm Au) was deposited by ebeam evaporation through a shadow mask. Top ohmic contacts were patterned using photolithography and the exposed $Al_2O_3$ in these regions was selectively dry etched using silicon tetrachloride $SiCl_4$ chemistry. The contact metal was subsequently deposited using ebeam evaporation and lift-off. The contact stack comprised Ti/Pt/Au (30 nm/50 nm/100 nm). Full details of this fabrication process are given in Appendix B.

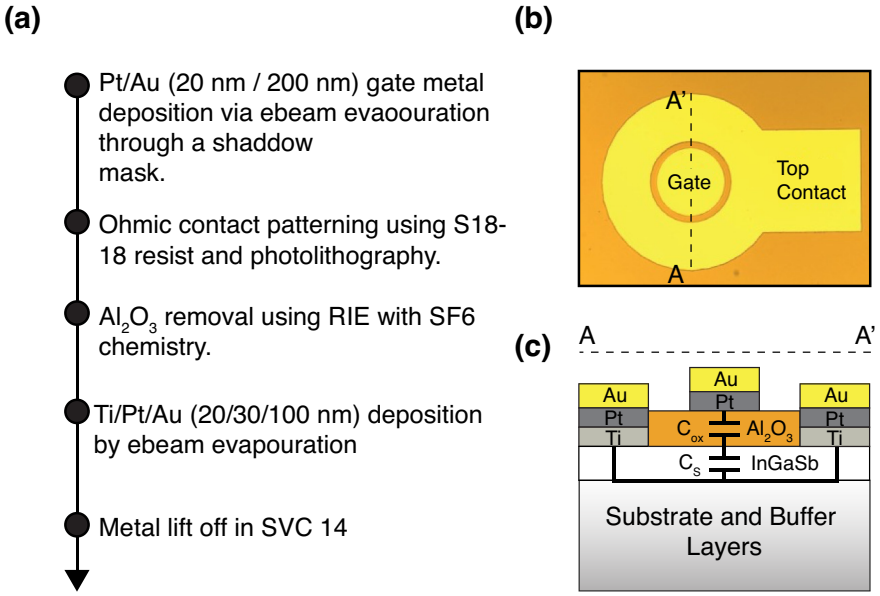


Figure 5.5: (a) Process flow used to fabricate metal oxide semiconductor capacitors (MOSCAPs) as part of the Power Series, which implement a top ohmic contact. (b) Plan view optical microscope image of a completed MOSCAP fabricated using the process flow in (a). (c) Schematic of the cross section of the MOSCAP structure shown in (b), illustrating the associated capacitances.

### 5.4.3 Electrical Characterisation

The impact of plasma power on the electrical properties of the interface was initially assessed, with and without a 15 minute FGA at 350°C, using room temperature, multifrequency, CV measurements, over a frequency range of 1 KHz to 1 MHz in a dark environment. It should be noted that a FGA would not normally be included in a device fabrication process post the formation of the ohmic contacts, due to the resulting degradation in contact resistance. This was deemed the most efficient methodology however to narrow a wide parameter space for subsequent investigation. As discussed in Section 5.2, the results were characterised qualitatively in terms of the following metrics: defined as $C_{mod} = [(C_{max,HF} - C_{min,HF})/C_{max,HF}] \times 100$ %; stretch out, quan-

tified by the maximise rate of change of capacitance with gate voltage, $[\frac{dC}{dV_g}]_{max}$; and frequency dispersion in accumulation, measured as the percentage change in accumulation capacitance per decade of frequency, $C_{\mathrm{LF,a}}$ - $C_{\mathrm{HF,a}})/(C_{\mathrm{LF,a}}) \times 100\%/N$, where $C_{\mathrm{LF,a}}$ is the low frequency accumulation capacitance, $C_{\mathrm{HF,a}}$ the high frequency accumulation capacitance, and $N$ the number of decades between the high and low frequencies used.

### Qualitative Metrics

No capacitance modulation was observed for samples which were exposed to ICP powers in excess of 250 W, for which high leakage current densities on the order of 10 $A/\mathrm{cm}^2$ were measured (not shown). Fig. 5.6 shows the room temperature CV measurements for p-type samples, processed with and without in-situ TMA pre-cleaning. For clarity, the following naming convention is adopted: samples which have been processed without in-situ TMA pre-cleaning are referred to as '$H_2$', whereas samples that have been subjected to TMA exposure prior to ALD are referred to as '$H_2$+TMA'. Additionally, the plasma power is incorporated such that the sample processed with TMA self cleaning and a 150 W $H_2$ plasma power is deemed the '150W-$H_2$+TMA' sample.

Prior to FGA, $C_{\mathrm{mod}}$ was found to degrade with $H_2$ plasma exposure, and decrease with increasing ICP power. The other metrics showed differing trends. $H_2$+TMA samples exhibited a degraded $C_{\mathrm{mod}}$ and stretch out in comparison to the equivalently processed $H_2$ samples. On visual inspection of the CV measurements shown in Fig. 5.6, $H_2$+TMA samples appear to exhibit substantially increased frequency dispersion in accumulation compared to $H_2$, however, this is somewhat misleading. It was found that the series resistance, $R_{\mathrm{s}}$, varied significantly depending on the inclusion of in-situ TMA pre-cleaning, which must terminate the semiconductor surface such that it increases the resistance of the metal-semicondcutor junction. The higher $R_{\mathrm{s}}$ of the $H_2$+TMA samples dominates the dispersion at frequencies more than $\sim$100 KHz, as evidenced by the heavily non-linear relationship between the measured capacitance and log frequency shown in Fig. 5.7. In order to correctly discern the qualitative impact of the various surface treatments on the oxide properties, the measured capacitances, $C_{\mathrm{m}}$, were corrected for $R_{\mathrm{s}}$ using Eqn. 5.1.

$$C_{\mathrm{c}} = \frac{(G_{\mathrm{m}}^2 + \omega C_{\mathrm{m}}^2)C_{\mathrm{m}}}{[G_{\mathrm{m}} - (G_{\mathrm{m}}^2 + \omega^2 C_{\mathrm{m}}^2)R_{\mathrm{s}}]^2 + \omega^2 + C_{\mathrm{m}}^2} \tag{5.1}$$

where $\omega$ is the angular frequency of the AC gate voltage, $G_{\mathrm{m}}$ the measured conductance and $R_{\mathrm{s}}$ was extracted in accumulation using:

$$R_s = \frac{G_{m,a}}{G_{m,a}^2 + \omega^2 C_{m,a}^2} \tag{5.2}$$

Following this procedure, frequency dispersion was extracted over a range of 1 kHz-100 kHz, which negated the effect of parasitic series inductance, $L_s$, which causes $C_m$ to increase at high frequency frequency. With these considerations, the frequency dispersion of $H_2$+TMA samples was in fact marginally lower than for $H_2$ samples processed with the same plasma power (other than the 0 W controls), as illustrated in Fig. 5.7 for 250 W samples.

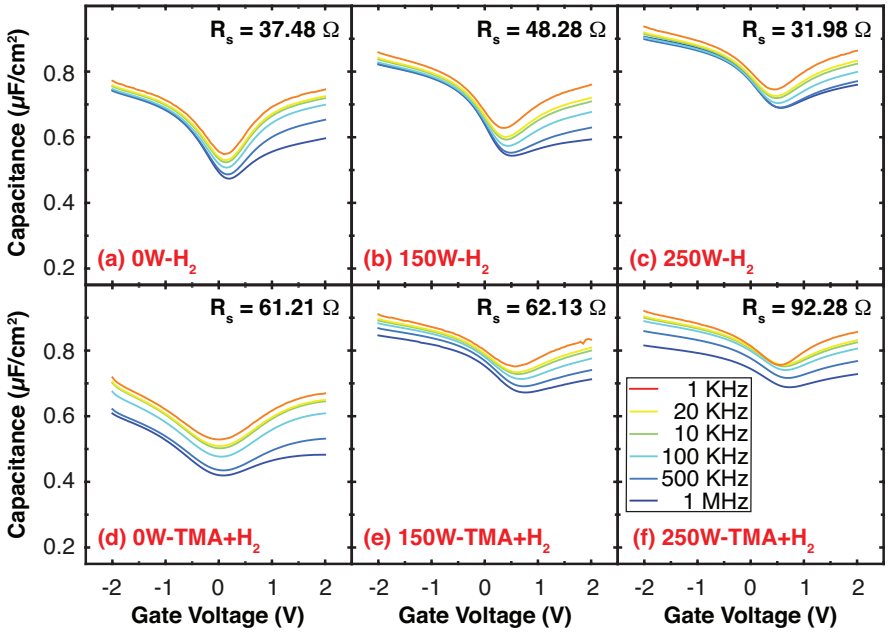


Figure 5.6: Room temperature CV measurements over a frequency range of 1 KHz to 1 MHz for p-type samples, processed with and without in-situ TMA pre-cleaning, pre FGA, exposed to $H_2$ plasma powers of (a,d) 0 W, (b,e) 150 W, and (c,f) 250 W.

Fig 5.8 shows the RT CV measurements for the p-type samples shown in Fig.5.6, processed with a FGA, and Fig 5.9 summarises the qualitative metrics for all p-type samples discussed thus far. The 250W-$H_2$ sample is not included as all devices had failed post the initial, pre FGA measurements. It was commonplace throughout these measurements that working capacitors would have initially low leakage currents, but fail during measurement, with the leakage current permantly increasing to compliance (50 mA). The combination of in-situ TMA cleaning, $H_2$ plasma exposure and FGA yield significantly improved metrics over the control, with the 150 W plasma treatment giving both the highest $C_{mod}$ and lowest stretch out. For this sample, the FGA increases $C_{mod}$ and $[dC/dV]_{max}$ by 122 % and 176 % respectively. Interestingly, post FGA, $R_s$ increased for $H_2$ samples, whereas it decreased for $H_2$+TMA. These results clearly demonstrate the improvement to the electrical properties the $In_{0.3}Ga_{0.7}Sb$-$Al_2O_3$ interface via in-situ $H_2$ plasma exposure in conjunction with TMA pre-cleaing; the mechanism of which appears to be

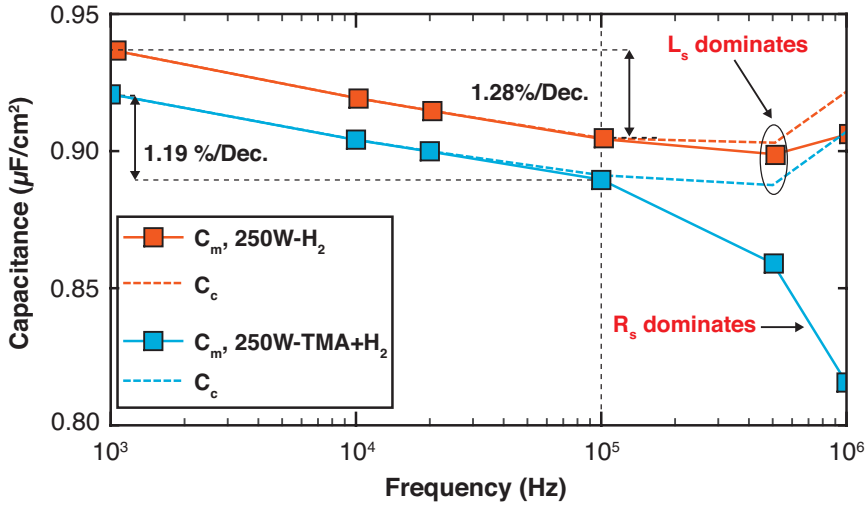


Figure 5.7: Frequency dispersion in accumulation for samples with different magnitudes of series resistance, $R_s$. Linear capacitance dispersion with log frequency is observed for both samples below 100 KHz, which is characteristic of border traps. At higher frequencies, $R_s$ dominates the admittance for the sample with higher $R_s$

critically dependent on the inclusion of a FGA.

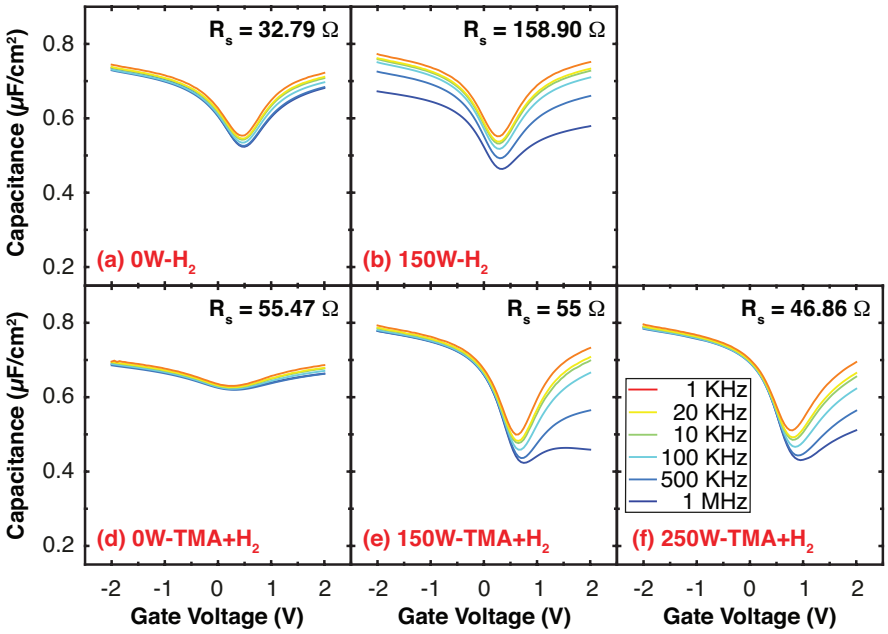


Figure 5.8: Room temperature CV measurements over a frequency range of 1 KHz to 1 MHz for p-type samples, processed with and without in-situ TMA pre-cleaning, post FGA, exposed to $H_2$ plasma powers of (a,d) 0 W, (b,e) 150 W, and (c,f) 250 W.

Fig. 5.10 shows the room temperature CVs for the n-type $H_2$+TMA samples pre and post FGA, exposed to plasma powers of 0, 150, and 250 W, with the corresponding qualitative metrics shown in Fig. 5.11. As with the p-type samples, the 150W-$H_2$+TMA n-type sample, post FGA,

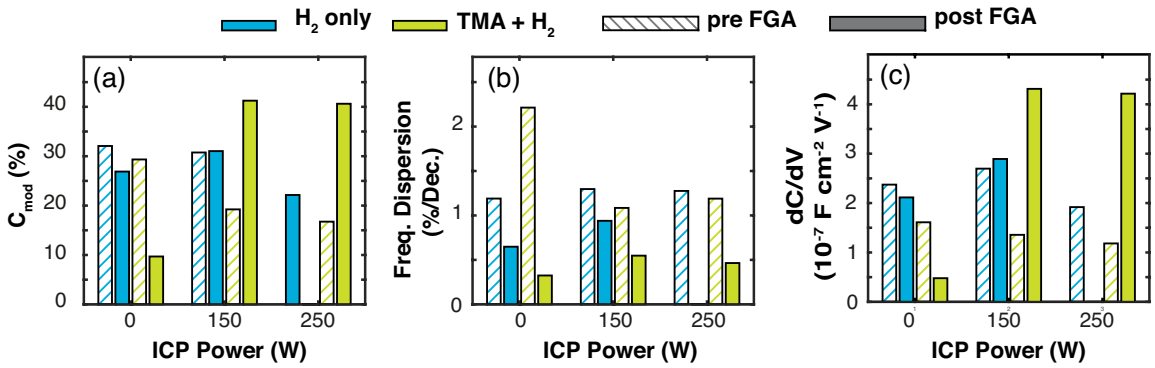


Figure 5.9: Comparison of modulation capacitance, Cmod (a), frequency dispersion in accumulation (b), and stretch out, dC/dV (c), for p-type samples as a function of $H_2$ plasma power, with and without in-situ TMA cleaning, pre and post FGA

exhibited the highest $C_{\text{mod}}$. Of all n-type samples post FGA, the 150W-$H_2$+TMA process also yielded the lowest frequency dispersion and stretch out.

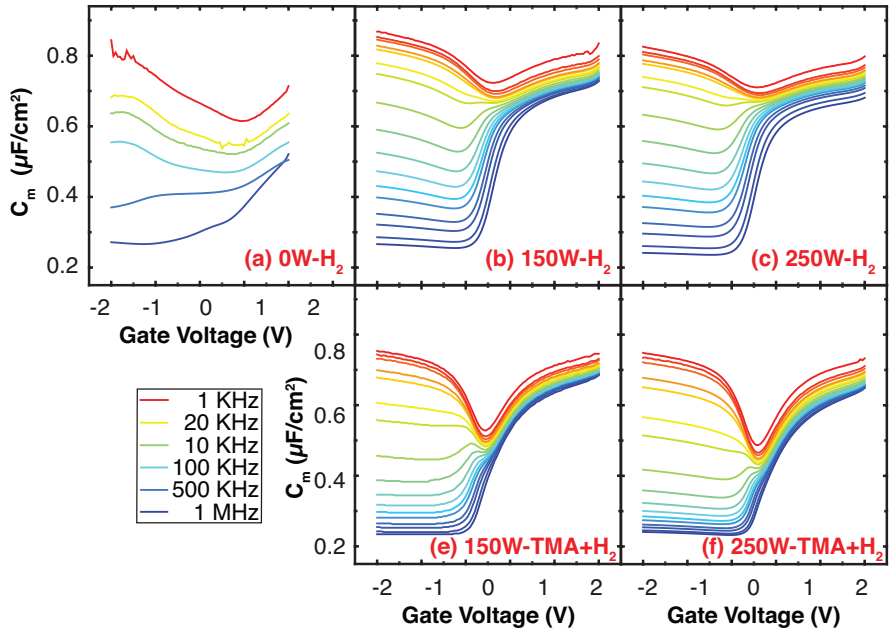


Figure 5.10: Room temperature CV measurements over a frequency range of 1 KHz to 1 MHz for n-type samples, processed with in-situ TMA pre-cleaning, pre and post FGA, exposed to $H_2$ plasma powers of 0 W, 150 W, and 250 W.

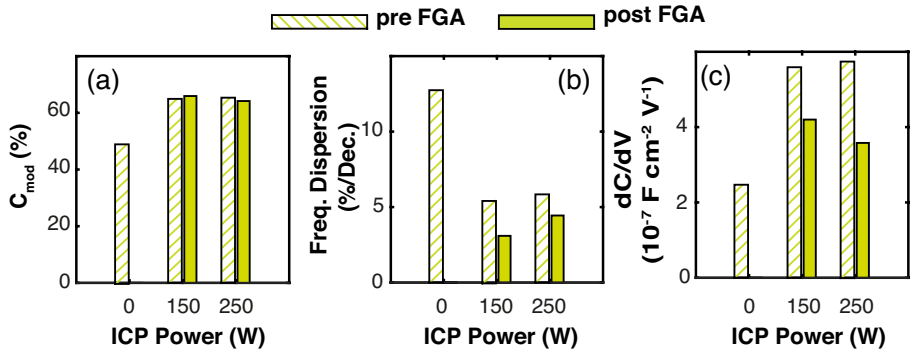


Figure 5.11: Comparison of modulation capacitance, $C_{\mathrm{mod}}$ (a), frequency dispersion in accumulation (b), and stretch out, dC/dV (c), for n-type samples pre and post FGA.

**Discussion of Fermi Level Pinning and Genuine Inversion**

Given the nominal doping concentration, the theoretical minimum capacitance for the Power Series samples, $C_{min,T}$, was calculated as follows:

- The maximum depletion width, $W_{\mathrm{D,max}}$,was calculated using Eqn. 5.3, where $\varepsilon_{\mathrm{InGaSb}}$ is the permittivity of InGaSb, k is Boltzmann's constant, T absolute temperature, $q$ the elementary charge, $n_{\mathrm{i}}$ the intrinsic carrier concentration, and $N_{\mathrm{D/A}}$ the nominal doping concentration, equal to either $N_{\mathrm{A}}$ or $N_{\mathrm{D}}$ for p and n-type samples respectively.

$$W_{\mathrm{D,max}} = \sqrt{\frac{4\varepsilon_{\mathrm{InGaSb}}\mathrm{kT}ln(N_{\mathrm{D/A}}/n_{\mathrm{i}})}{q^2 N_{\mathrm{D/A}}}} \tag{5.3}$$

- The corresponding minimum depletion capacitance, $C_{\mathrm{D,min}}$, was calculated as $C_{\mathrm{D,min}}$=$\varepsilon_{\mathrm{InGaSb}}$/ $W_{D,max}$.

- The theoretical minimum total capacitance, $C_{\mathrm{min,T}}$, was then calculated as the series combination of the oxide capacitance, $C_{\mathrm{ox}}$, and $C_{\mathrm{D,min}}$:

$$C_{\mathrm{min,T}} = \frac{C_{\mathrm{ox}}C_{\mathrm{D,min}}}{C_{\mathrm{ox}} + C_{\mathrm{D,min}}} \tag{5.4}$$

Estimating $C_{\mathrm{ox}}$ as the maximum measured capacitance in accumulation gives a theoretical minimum capacitance of $\sim 189\ nF/cm^2$. The minimum capacitance of both p-type 150W/250W-$H_2$+TMA samples is more than double the theoretical value, and therefore, the Fermi level is clearly pinned such that the surface cannot fully deplete. In contrast, the minimum measured capacitance of $\sim 230\ nF/cm^2$ for 150 and 250 W n-type $H_2$+TMA samples is only $\sim 30$ % higher than $C_{\mathrm{min,T}}$. Qualitatively, the flat capacitance responses for these samples at gate biases corresponding nominally to inversion is indicative of a genuinely inverted surface. Fig. 5.11 shows CV measurements for p and n-type 250W-$H_2$+TMA samples taken at 30 K using a Lakeshore cryogenics probe station and Keysight B1500 semiconductor parameter analyser. For the p-type sample, the minimal flat band shift and marginal decreases in accumulation capacitance demonstrate clearly that the Fermi level can indeed move freely into the valence band, and therefore, suggests that the n-type sample is genuinely inverted. For the n-type sample, the large positive flat band shift and increased frequency dispersion in accumulation indicate that the supposed accumulation region observed at room temperature was in fact false accumulation, and that the measured capacitance was due to high $D_{\mathrm{it}}$. This is expected given the room temperature p-type data showing the Fermi level of the interface resulting from the 250W-$H_2$+TMA process to be pinned somewhere above the valence band edge. The reduction of $C_{min}$ for the p-type sample with reducing temperature is attributed to the freezing out of interface traps. The large disparity

between $C_{min}$ at RT and at 30 K , which approaches $C_{\mathrm{min,T}}$ at low temperature, suggests significant movement of the Fermi level away from the valence band edge. This is in stark contrast to the p-type $H_2$+TMA for example, shown in Fig. 5.11, which clearly has very limited Fermi level movement.

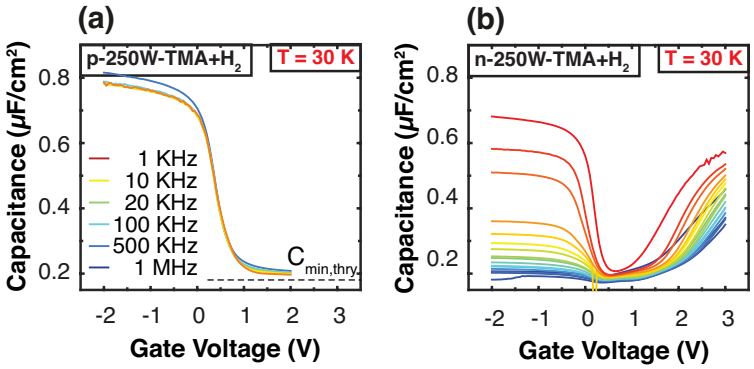


Figure 5.12: Capacitance-Voltage (CV) measurements at 30 K of (a) p and (b) n-type 250-W, $H_2$+TMA samples post FGA.

If the n-type 150W/250W $H_2$+TMA samples are indeed genuinely inverted, their minimum measured capacitance would correspond to a doping density of $N_{\mathrm{D}}$=$\sim 3.8 \times 10^{17} cm^{-3}$. This margin of error from the nominal value is higher than would be expected from error in the MBE process, which is commonly around $\pm$50%. An explanation for this is discussed in detail in Section 5.5.1, where it is conclusively shown that there is positive charge in the oxide which inverts the surface surrounding the gate. This results in a mechanism which masks the true high frequency response, and therefore, the true minimum capacitance corresponding to the actual doping concentration is not observed. Further, the existence of an inversion layer beyond the gate explains the persistence of the minority carrier response observed at low temperature in Fig.5.11 (b): minority carriers can be supplied to the inversion layer by diffusion from the externally inverted surface, and therefore do not require thermal generation and are not suppressed at low temperature.

To further validate the above, a novel impedance spectroscopy method was employed which has been shown to discern genuine surface inversion from false inversion due to $D_{\mathrm{it}}$ on InGaAs [36]. The methodology, which was derived from physics based AC simulations, involves the analysis of the frequency scaled conductance, $G_{\mathrm{m}}/\omega$, and $-\omega dC/d\omega$ as a function of frequency. The criteria for genuine surface inversion is as follows: for a given sample, the peak magnitude of both $G_{\mathrm{m}}/\omega$ and $-\omega dC/d\omega$ are equal and occur at the same frequency, and that frequency corresponds to the transition frequency, $\omega_m$, where the capacitance in inversion is half way between the maximum and minimum values, and the conductance in inversion is maximum. It is found that when applying this analysis to the n-type 250W-$H_2$+TMA sample, $G_{\mathrm{m}}/\omega$, and $-\omega dC/d\omega$ are in phase and of the same magnitude, and the frequency of their peak magnitude

correctly corresponds to the transition frequency, as highlighted in black in the CV (Fig. 5.10 (e) and GV (f) responses, indicative of genuine surface inversion. Unsurprisingly this is not the case for the p-type sample, which as expected is shown to be pinned.

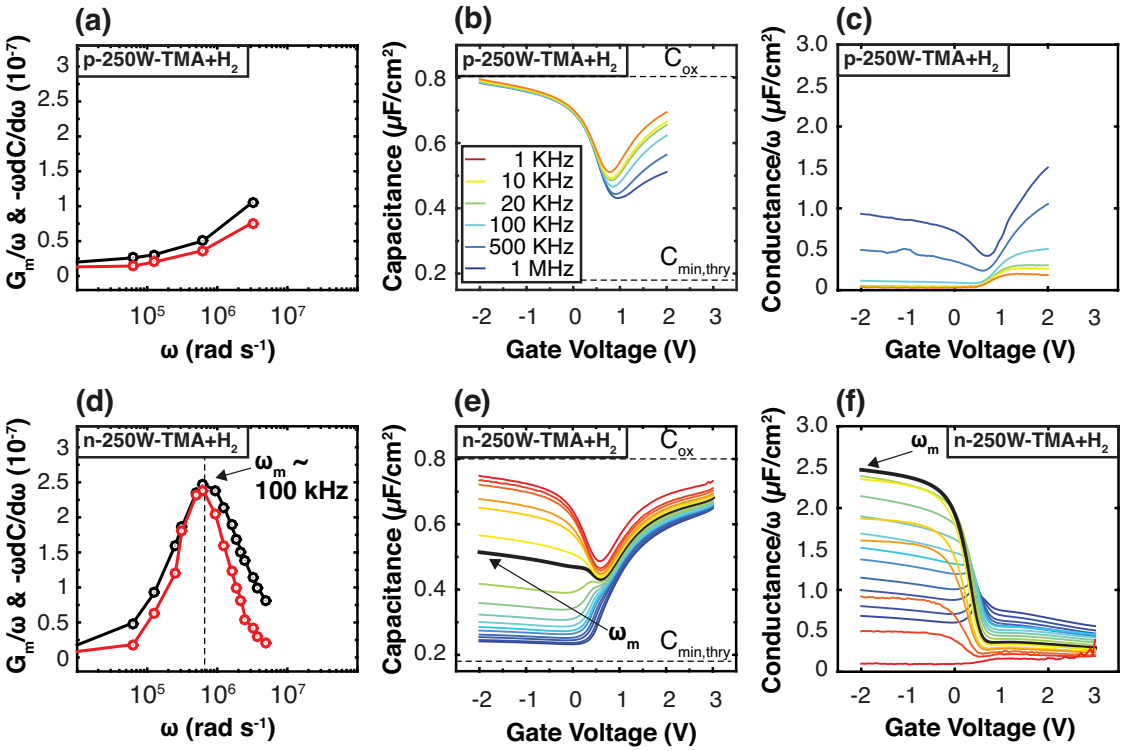


Figure 5.13: (a,d) $G_m/\omega$ and $-\omega dC/d\omega$ as a function of frequency for p and n-type 250W-$_2$+TMA samples. (b,c,e,f) Room temperature capacitance-voltage (CV) and conductance-voltage (GV) measurements for the aforementioned samples.

### 5.4.4 Chemical analysis

Bright field transmission electron microscope (TEM) images of the 0 W,150 W, and 250 W p-type $H_2$+TMA samples, post FGA, are shown in Fig.5.14 (a-c). The samples were prepared for TEM imaging at Tyndall National institute using focussed ion beam (FIB) thinning on a FEI 200 Workstation. The image shown for the 0 W sample was taken at TNI using a JEOL 20000FX microscope operated at 200 kV. The images of the 150 W and 250 W samples were taken at IBM Zurich Research Laboratory using a a double spherical aberration corrected JEOL ARM200F microscope also operated at 200 kV, which was further equipped with a liquid nitrogen free Si-drift detector enabling energy dispersive X-Ray (EDX) analysis of the interfaces. The images show the $Al_2O_3$ to be conformal and approximately 7-7.5 nm thick for all samples. The oxide capacitance for 150 W and 250 W samples was extracted from simulation (full details of which are discussed in Sections 5.6 to 5.8) as 0.8 $\mu$F/cm$^2$, giving a relative permittivity for $Al_2O_3$ of 6.69. Qualitative EDX line scans of the interface of the 150 W sample, shown in Fig. 5.14 (d), reveals the concentration of Al and O to vary throughout the thickness of the oxide, in addition to a $\sim 1.4$ *nm* thick region of intermixing of In,Ga and Sb with $Al_2O_3$, both of which would reduce the effective permittivity of the gate stack from that of the textbook $Al_2O_3$ value of 9.

X-Ray photoelectron spectroscopy (XPS) measurements were performed at the University of Liverpool to compare the chemical composition of the surface of the following samples:

- A virgin $In_{0.3}Ga_{0.7}Sb$ surface exposed to air to elucidate the composition of the native oxide formed.
- An HCl cleaned $In_{0.3}Ga_{0.7}Sb$ surface, which was processed with the ex-situ parameters as detailed in Fig. 5.4 and transferred to the XPS chamber within 1 minute.
- Sample 0W-$H_2$+TMA post FGA., which provided information as to the impact of in-situ TMA cleaning without preceeding $H_2$ plasma exposure. In order measure the composition of the InGaSb surface of this sample it was required to thin the oxide to approximately 4 nm using in-situ Ar sputtering. The Ar ion beam was operated at 500 eV and 10 $\mu$A flux which yielded a sputtering rate of 0.032 nm/min. XPS measurements were taken during the sputtering process which provided insight into the composition of the oxide as a function of depth. These measurements utilised a shallow take off angle (30°) to maximise the sensitivity to the surface composition, and minimise measurement of the bulk oxide.
- Sample 250W-$H_2$+TMA post FGA, which facilitated the combined impact of cleaning the sample surface with both in-situ $H_2$ plasma and TMA to be quantified. As with the 0W-$H_2$+TMA sample, this required the use of in-situ Ar sputtering, and incremental measurements during the sputtering process were also recorded.

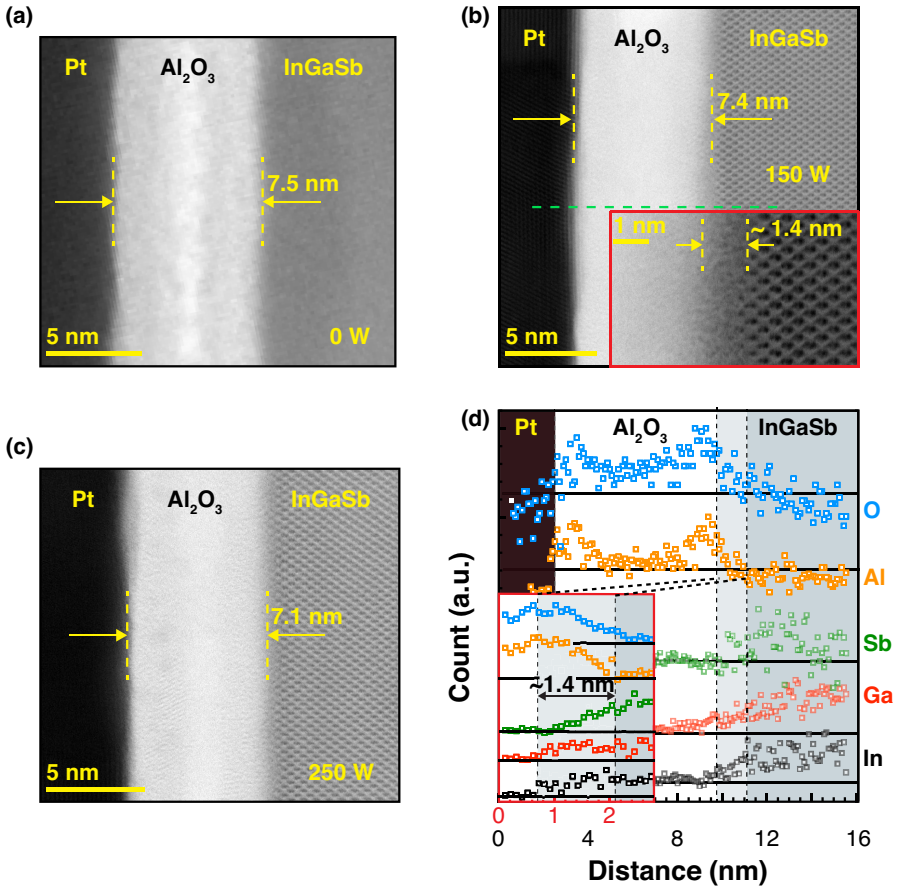


Figure 5.14: (a-c) TEM images of the $H_2$+TMA InGaSb-$Al_2O_3$ interface, treated with 0 W (a), 150 W (b) and (c) 250 W plasma power, post FGA. Inset to (b): higher magnification STEM image showing the existence of an interfacial layer approximately 1.4 nm thick. (d) Qualitative EDX line scan over the region marked in green in (b), revealing the oxide to be non-stoichiometric. Inset: High magnification EDX measurements over the interfacial region, showing intermixing / elemental diffusion of In,Ga and Sb with the oxide

Unmonochromatised Al K$\alpha$ (1486.6 eV) radiation was used as the photon source, which was operated at a power of 144 W. The spectrometer used was a Scienta SES200 hemispherical electron energy analyser which provided a resolution of 0.2 eV. The measured spectra were fitted with Gaussian-Lorentzian line shapes subsequent to a Shirley-type background subtraction and the binding energies were calibrating by setting the Au $4f_{7/2}$ core level (CL) at 84 eV (present due to the contacts on the sample).

For the native InGaSb surface, the XPS spectra shown in Fig. 5.15 shows the In $3d_{5/2}$ peak at 444.1 eV, which is in close agreement with reported values of 444.3 eV for InSb [42]. A further peak is observed at 444.9 eV corresponding to the In-O bond, and can be attributed to $In_2O_3$ which has been shown to exist at approximately +0.9 eV from the In $3d_{5/2}$ peak [43, 44]. The Ga $2p_{3/2}$ peak is evident at a BE of 1116.8 eV, in addition to a broad higher BE component at 1117.9 eV corresponding to the Ga-O bond. The energy range corresponding to the Sb $3d_{3/2}$

core level spectra was measured to avoid the overlap of O 1s and Sb $3d_{5/2}$ peaks. The Sb $3d_{3/2}$ peak was observed at 537.13 eV (values previously reported of 537.17 eV for InSb [45]), as well an associated oxide feature at an offset of +2.6 eV. Investigations into the oxide composition of the native GaSb surface have reported an Sb-O peak with a chemical shift from the bulk of ∼3.0 eV, which has been shown to comprise $Sb_2O_3$ and $Sb_2O_4$ components at +2.5 and +3.1 eV respectively [46].The oxide feature situated at +2.6eV from the Sb $3d_{3/2}$ line for the native InGaSb surface therefore appears to correspond predominantly to an $Sb_2O_3$ sub oxide phase.

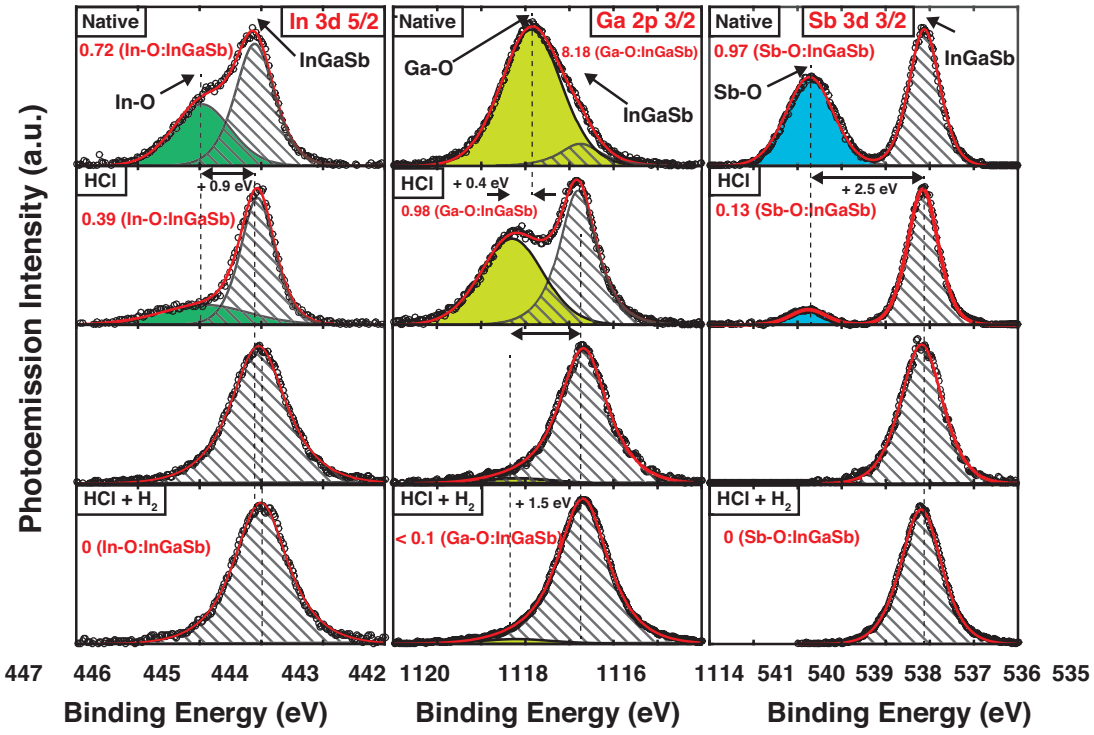


Figure 5.15: Deconvoluted XPS spectra of the In $3d_{5/2}$, Ga $2p_{3/2}$ and Sb $3d3/2$ peaks for: a native air exposed $In_{0.3}Ga_{0.7}Sb$ surface, an HCl cleaned $In_{0.3}Ga_{0.7}Sb$ surface, the $0W - H_2 + TMA$ sample, and the $250W - H_2 + TMA$.

Cleaning the surface with HCl reduces all In, Ga and Sb sub oxides, with Ga-O remaining the most prominent. The Ga-O peak shifts by ∼ 0.4 eV to a higher BE of 1118.3 eV, in agreement with reported values of $Ga_2O_3$ [47]. The broad peak at a lower BE for the native surface is indicative of the presence of non-stoichiometric disordered $GaO_x$.

The In $3d_{5/2}$ and Sb $3d_{3/2}$ core level spectra for the 0W-$H_2$+TMA sample can be fitted well with single peaks corresponding the InGaSb substrate, indicating the complete removal of In and Sb sub oxides. The Ga $2p_{3/2}$ spectra shows $Ga_2O_3$ to have persisted, however, the Ga-O:InGaSb ratio has been reduced from 8.18 in the HCl only sample to < 0.8. The TMA therefore appears to offer the same self cleaning benefits for InGaSb as has been demonstrated for other III-Vs. It is interesting to note, however, that despite this, the CV measurements for this sample

demonstrated an almost entirely pinned Fermi level at the valence band edge. The XPS spectra for the 250W-$H_2$+TMA sample shows remarkably similar results to that of the 0W-$H_2$+TMA sample, with the only measurable difference being the marginal increasing in Ga-O:InGaSb ratio from $\sim 8$ % to $\sim 10$ %.

Analysis of the oxide composition offers insight into the disparity between the CV responses of these two samples despite their almost indistinguishable surface chemical composition. Fig. 5.16 shows the Al 2p CL spectra as function of etched depth for 0W-$H_2$+TMA and 250W-$H_2$+TMA samples. For both, the surface of the oxide is well fitted by single peaks corresponding to $Al_2O_3$, with the increased presence of $AlO_x$ detected as the oxide is thinned. Fig. 5.17 plots the ratio of the area of $AlO_x$ to $Al_2O_3$ for both samples as function of etch depth. Interestingly this shows that there is a higher concentration of $AlO_x$ towards the InGaSb interface for the 250 W sample than for the 0W. This appears to indicate that the improvement to the electrical properties of the InGaSb-$Al_2O_3$ interface with in-situ $H_2$ plasma exposure is not due to the reduction of native oxides on the surface, but in fact due to the formation of a different oxide composition at the interface.

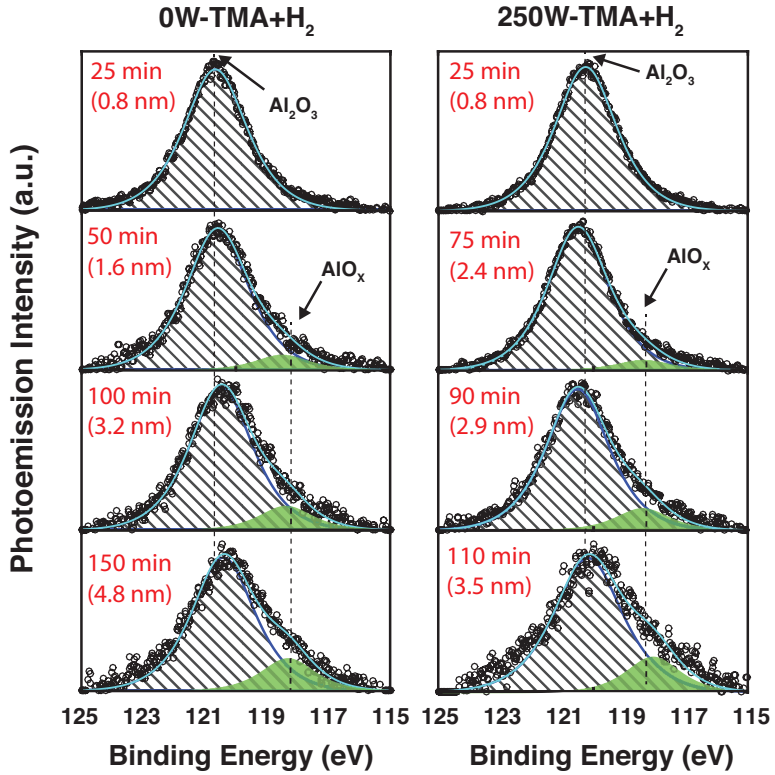


Figure 5.16: Deconvoluted XPS spectra of the Al $2_p$ peak measured on 0 W and 250 W samples, as a function of in-situ $Al_2O_3$ etch time.

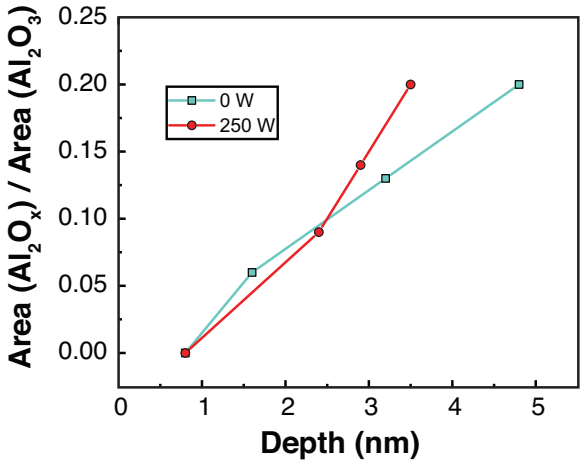


Figure 5.17: The ratio of the area of $AlO_x$ to $Al_2O_3$ for 0W and 250W $H_2$+TMA samples as a function of etch depth.

## 5.5 Impact of $H_2$ Plasma Exposure Time

Given the results of the Power Series, the 150W-$H_2$+TMA process including FGA was selected as the base process for investigation into the impact of exposure time. The parameter space for the Time Series is summarised in Fig. 5.18, which included $H_2$ plasma exposure times of 1, 10, 30 and 60 minutes, with the remaining parameters fixed at the values detailed in Section 5.4. The simplified MOSCAP process flow afforded by utilising a doped substrate and buffer layers is shown in Fig. 5.19, which comprised gate contact deposition through a shadow mask, post gate metal FGA, and the formation of a back contact using blanket metal deposition to the back of the sample. These samples were fabricated by Dr Uthayasankaran Peralagu and the corresponding RT qualitative metrics are shown in blue in Fig. 5.20.

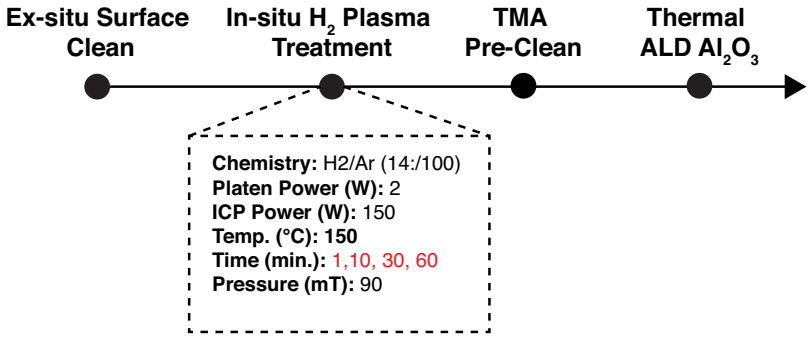


Figure 5.18: Experimental matrix for samples processed as part of the Time Series. The unspecified parameters are identical to those given in Fig. 5.4.

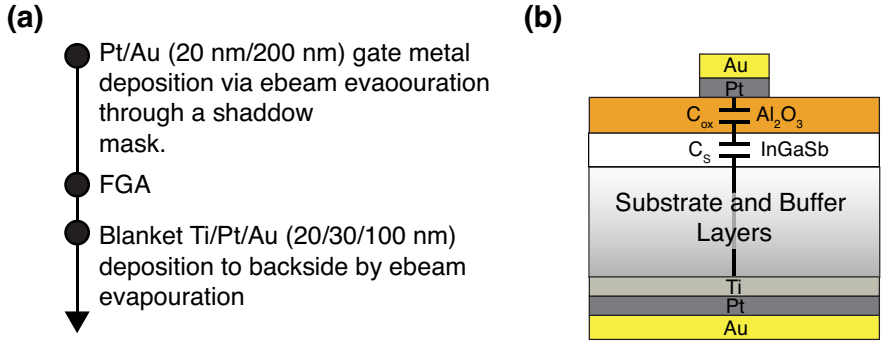


Figure 5.19: Experimental matrix for samples processed as part of the Time Series. The unspecified parameters are identical to those given in Fig. 5.4.

$C_{\text{mod}}$ and $dC/dV$ were found to increase with decreasing exposure time, with optimal values at 10 minutes. Frequency dispersion did not show a clear trend. Given the results of the Power Series indicating lower plasma powers to be beneficial, further samples were processed at a plasma power of 125 W, which was the lowest power that sustained a plasma at the given chamber conditions. It was intended to repeat the Time Series experiments at the this lower ICP power, however, the plasma was found to become unstable during longer runs, which limited

the data set to exposure times of 1 minute only. The RT CV measurements for p and n-type samples processed at 125 W for 1 minute are shown in Fig. 5.21 (a) and (b), with the corresponding qualitative metrics for the p-type sample included in Fig. 5.20, exhibiting the largest $C_{\text{mod}}$ (74.8 %), and lowest stretch out of all samples.

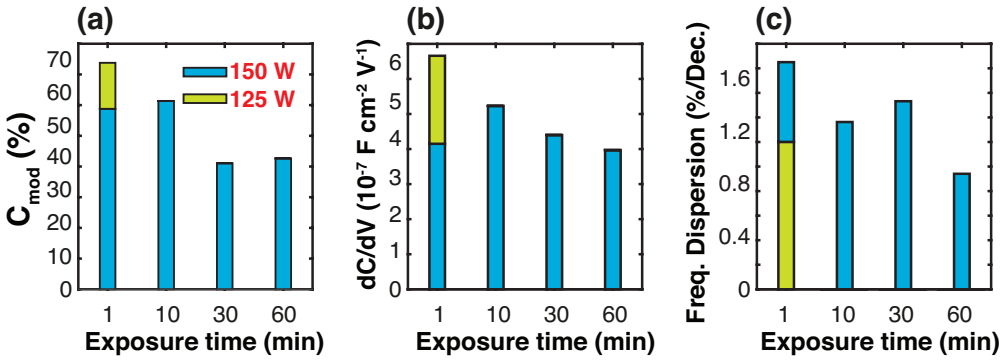


Figure 5.20: Comparison of modulation capacitance, Cmod (a), stretch out, dC/dV (b) and frequency dispersion in accumulation (c) Time Series samples.

### 5.5.1 Inversion Analysis

Assuming that the 1 MHz CV measurements yields the true high frequency response, and that the samples are fully depleted, the minimum measured capacitance of $\sim 0.2\ \mu$F/cm$^2$ corresponds to a doping density of $\sim 2.5 \times 10^{17}$ cm$^{-3}$ with an extracted oxide capacitance of 0.785 $\mu$F/cm$^2$. As discussed in Section 5.4, attributing a discrepancy between the measured and nominal theoretical minimum capacitance to a higher than nominal doping density is often questionable, with the possibility that the sample is in fact pinned such that the maximum depletion width is not reached, and the supposed inversion response is false inversion due to Dit [48]. In order to distinguish between genuine surface inversion and a false inversion response, variable temperature CV and GV measurements were analysed comprehensively by the following means: the novel impedance spectroscopy technique discussed in Section 5.4; extraction of the activation energy of minority carriers from an Arrenhius plot of equivalent parallel conductance in inversion; and, detailed in Serction 5.9, the extraction of the gate-voltage surface potential relationship from fitting the simulation of an equivalent circuit model to the experimental results.

As shown in Fig. 5.21, it is found that for both p and n-type samples, $G_{\text{m}}/\omega$, and $-\omega dC/d\omega$ are in phase and of the same magnitude, and the frequency of their peak magnitude correctly corresponds to the transition frequency, as highlighted in black in the CV (Fig.5.21 (c,d) and GV (e,f) responses, indicative of genuine surface inversion for both doping polarities.

Variable temperature measurements were undertaken at Tyndall National Institute using a microchamber probe station (Cascade Summit12971B) in a dark, dry air (dew point $<$ -65°C)

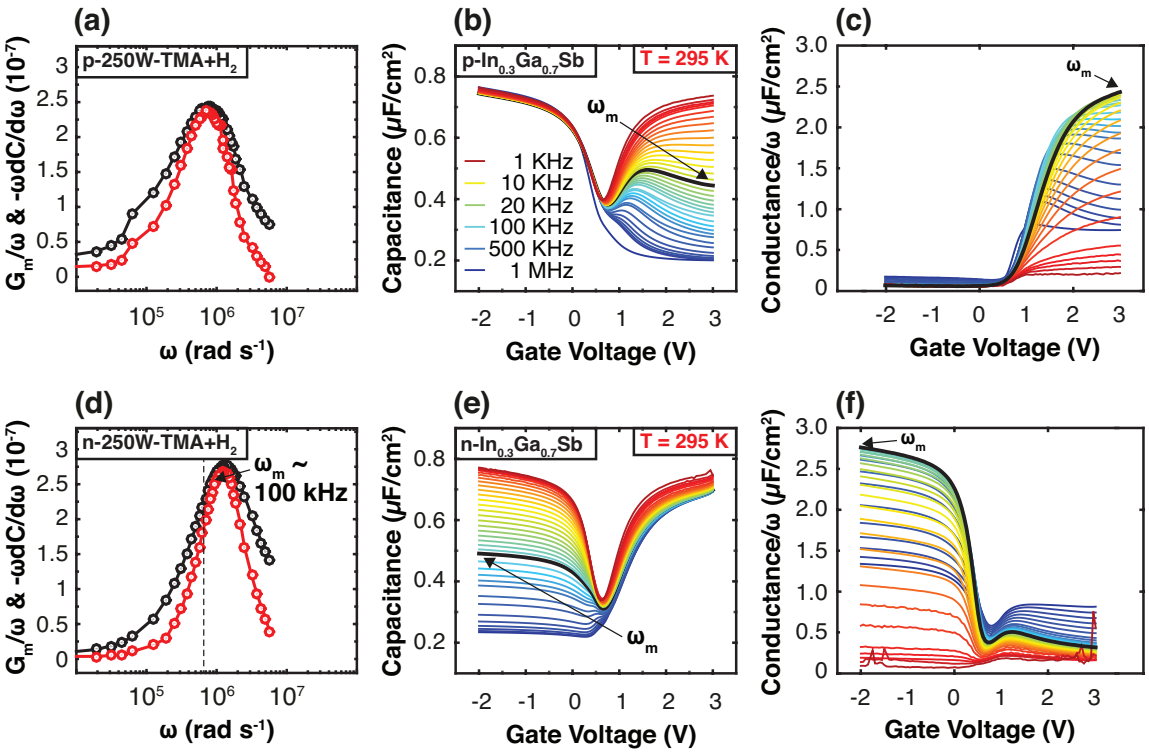


Figure 5.21: (a,d) $G_m/\omega$ and $-\omega dC/d\omega$ as a function of frequency for p and n-type samples treated with the optimal Time Series process. (b,c,e,f) Room temperature capacitance-voltage (CV) and conductance-voltage (GV) measurements for the aforementioned samples.

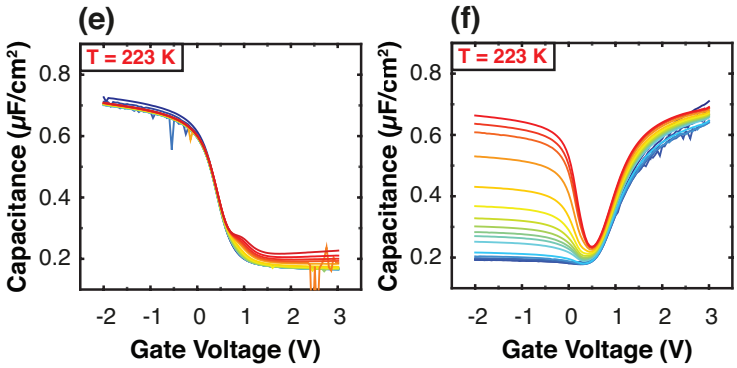


Figure 5.22: Low temperature CV measurements of p and n-type samples treated with the optimal Time Series process.

environment over a temperature range of RT to -50°C. As with all previous measurements, the data was recorded using a Keysight B1500 semiconductor analyser.

### The Activation Energy of Minority Carriers on Narrow Band Gap Materials

For the p-type sample, an Arrenhius plot of the equivalent parallel conductance in inversion (Vg = 3 V), $G_{\mathrm{I}}$, which is calculated from the measured capacitance ($C_{\mathrm{m}}$) and conductance ($G_{\mathrm{m}}$) via Eq. 5.5, versus 1/kT where k is Boltzmanns constant and T temperature, is shown in Fig. 5.23(a).

$$G_{\mathrm{I}} = \frac{\omega^2 C_{ox}^2 G_{\mathrm{m}}}{G_{\mathrm{m}}^2 + \omega^2 (C_{ox} - C_m)^2} \tag{5.5}$$

The extracted activation energy, $E_{\mathrm{A}}$, of 0.296 eV is in close agreement with half of the band gap energy ($E_{\mathrm{g}}/2 \sim 0.245$ eV, shown in Fig. 5.3(d)) indicating genuine surface inversion with minority carriers supplied via generation-recombination (G-R) in the bulk [49]. The magnitude of the discrepancy between the extracted $E_{\mathrm{A}}$ and $E_g/2$ is within the margin of error reported for both InGaAs [50] and Si [49] MOSCAPs. Furthermore, it is to be expected that $E_{\mathrm{A}}$ is higher than $E_{\mathrm{g}}/2$ for the InGaSb MOS system: the assignment of $E_{\mathrm{g}}/2$ for the activation energy of the G-R dominated regime is derived from the dependancy of $G_{\mathrm{I}}$ on intrinsic carrier concentration, $n_{\mathrm{i}}$, which can approximated as:

$$n_{\mathrm{i}} = \sqrt{N_{\mathrm{c}} N_{\mathrm{v}}} exp(\frac{E_{\mathrm{g}}}{2\mathrm{kT}}) \tag{5.6}$$

The deviation from this approximation increases, however, with reducing band gap. The right-hand Y-axis of Fig. 5.23(a) plots the ratio of $E_{\mathrm{A}}$ to $E_{\mathrm{g}}$ for InSb and GaSb, where $E_{\mathrm{A}}$ was calculated as $\frac{d}{d(1/\mathrm{kT})} ln(n_{\mathrm{i}})$, where the empirical relation of $n_{\mathrm{i}}(\mathrm{T})$ is known for InSb and GaSb from Refs. 51 and 52 respectively. As shown, $n_{\mathrm{i}}$ changes with temperature with an activation energy that his higher than $E_{\mathrm{g}}/2$ for both InSb and GaSb over this temperature range.

### Evidence of an Externally Inverted Surface

This analysis was found not to be applicable to the n-type sample as characteristics of an external inversion layer beyond the gated area were evident, where the InGaSb surface was inverted due to charge in the oxide. In such a case, the dominant mechanism over all temperatures by which minority carriers are supplied to the inversion layer is diffusion from the externally inverted surface, as depicted in the left hand side of the MOS schematic inset to Fig. 5.24(b) [49]. This mechanism results in the inversion response at a given frequency reducing with increasing gate area, as observed in Fig. 5.23(b) at a frequency of 1 MHz, due to the increased diffusion distance from the externally inverted region to the centre of the gate [53]. The existence of an external inversion layer is further validated by the fact that the minority carrier response of the

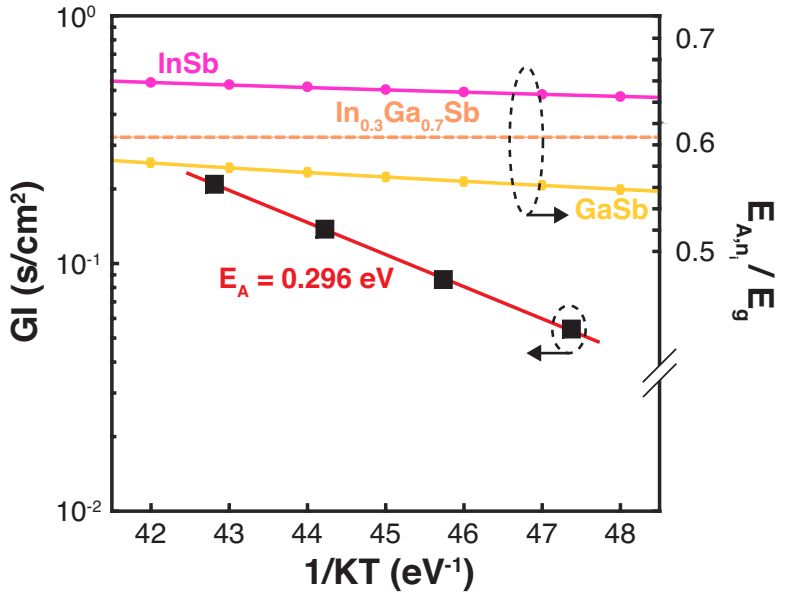


Figure 5.23: Left-hand Y-axis: Arrenhius plot of the equivalent parallel conductance in inversion ($Vg = 3\ V$),$G_{\mathrm{I}}$, against 1/kT. The experimental data points are shown as black squares, with the exponential trend line in red, showing an extracted activation energy, $E_{\mathrm{A}}$, of 0.296 eV. Right-hand Y-axis: the ratio of $E_{\mathrm{A}}$ to bandgap energy, $E_{\mathrm{g}}$, for InSb and GaSb, where $E_{\mathrm{A}}$ was calculated from known empirical relationships of $n_{\mathrm{i}}$(T), illustrating the deviation from an $E_{\mathrm{g}}/2$ dependance.

n-type sample is not suppressed at low temperature (Fig. 5.22(f)), as the minority carriers are not thermally generated, and that these characteristics are only observed for one doping polarity, as the same oxide charge present in the p-type sample would result in accumulation, not inversion. A further consequence of an external inversion layer is lateral AC current flow, where at high frequency the minority carriers can not follow the applied AC signal and thus the surface beneath the gate remains inverted and acts as a conductor through which AC current can flow laterally beyond the gate edge into the externally inverted surface, which behaves as a distributed R-C network [54] (depicted in the right hand side of the schematic inset to Fig. 5.23(b)). As the gate bias is pushed further into inversion, the coupling between the inversion layer beneath the gate and the external R-C network increases, and thus the measured capacitance increases [54]. It is observed that for all measured frequencies and temperatures, the n-type CV response features a distinct minimum at $V_{\mathrm{g}} \sim 0.4$ V, with $C_{\mathrm{m}}$ increasing with more negative $V_{\mathrm{g}}$. This is explained by the above mechanism masking the true high frequency response.

The observation of characteristics pertaining to an externally inverted surface is itself validation of genuine surface inversion, without which the gated region would not be coupled to the external inversion layer and these characteristics would not be observed.

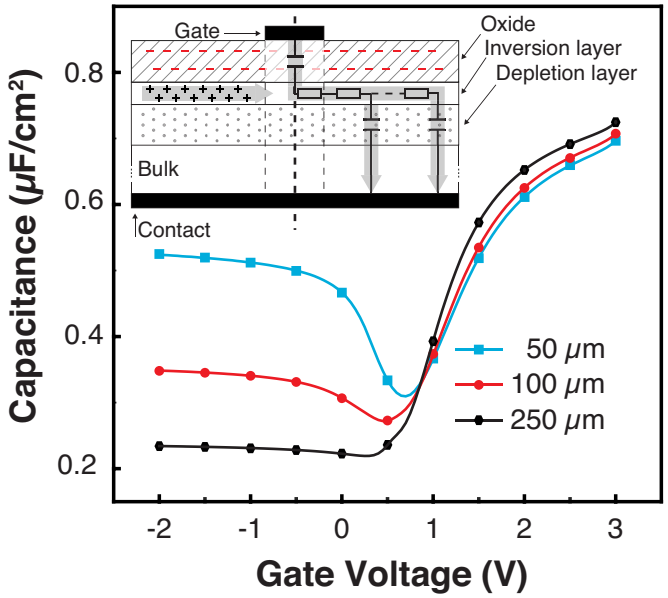


Figure 5.24: Room temperature capacitance-voltage (CV) measurements at a frequency of 1 MHz for the n-type sample shown in Fig. 5.21(b) for gate diameters of 50,100 and 250 $\mu m$. Inset: schematic of an n-type MOSCAP with negative charge in the oxide causing an external inversion layer beyond the gate.

## 5.6 Modelling the $In_{0.3}Ga_{0.7}Sb$-$Al_2O_3$ Interface I: Calculating the Semiconductor Charge

In order to circumvent the well documented issues associated with extracting $D_{it}$ on narrow band gap semiconductors [32], the experimental CV and GV data was simulated using an equivalent circuit model. The basis of the model is shown in Fig. 5.25 and comprises the oxide capacitance, $C_{ox}$, in series with the semiconductor capacitance, $C_s$(Fig. 5.25(a)), which in turn comprises the parallel combination of semiconductor depletion and inversion capacitances, $C_d$ and $C_{inv}$ respectively (Fig. 5.25(b)).

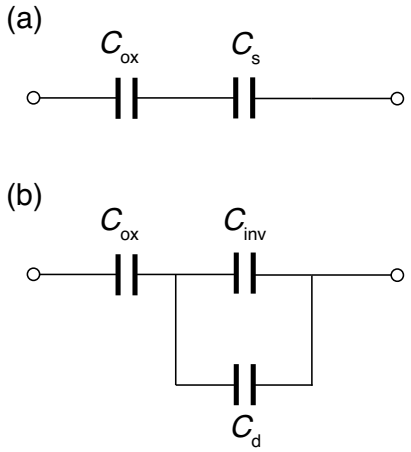


Figure 5.25: (a) The basis of the equivalent circuit model which is simply the oxide capacitance, $C_{ox}$, in series with the semiconductor capacitance, $C_s$. (b) $C_s$ comprises the parallel combination of the depletion and inversion capacitances, $C_d$ and $C_{inv}$ respectively.

$C_d$ and $C_{inv}$ can be calculated as per the textbook procedure [55], however, given the light effective mass of the conduction band of InGaSb, the Boltzmann approximation cannot be used and instead the Fermi-Dirac integral must be solved numerically. Further, the non-parabolicity of the conduction band should be accounted for and the population of the upper lying valleys considered. This procedure is summarised below, with a detailed discussion provided in Sections 5.6.1 and 5.6.2.

- The schematic band structure for $In_xGa_{1-x}Sb$ shown in Fig. 5.26 was calculated using Vegard's law [56] with the corresponding data for InSb and GaSb from Refs. [1] and [2] respectively, and bowing parameters from Ref. [57]. An InSb mole fraction of x=0.275 was used to set the energy gap of the $\Gamma$ valley to 490 meV in agreement with the measured PL data shown Fig 5.3 (d). Conduction band non-parabolicity was taken into account for the $\Gamma$ valley via the non-parabolicity factor $\alpha_\Gamma$, which was calculated from Eqn. 5.7, the expression for which is derived from k·p perturbation theory [37, 38, 58]:

$$\alpha_\Gamma = \frac{1}{E_{g,\Gamma}}\left(1 - \frac{m^*_\Gamma}{m_0}\right) \tag{5.7}$$

where $E_{g,\Gamma}$ is the bandgap energy between the HH maxima and Γ valley minima, $m^*_\Gamma$ is the effective mass of the Γ valley and $m_0$ is the electron rest mass.

- The potential dependant carrier concentration for each of the Γ, L, X and HH valleys was calculated via numerical integration of the product of the density of states of the given valley, $g_i$, where $i = \Gamma$, L, X, HH, and the Fermi-Dirac function, $F$, both of which are functions of energy, $E$. For electrons, the carrier concentration from each of Γ, L and X valleys was summed to yield the total electron concentration, as expressed by:

$$n(\psi(x)) = \sum_{\Gamma,L,X} \int_{E_{min,i}}^{\infty} g_i(E)F(E).dE \tag{5.8}$$

where, $n$ is the electron concentration, $\psi$ the electrostatic potential, $x$ the depth into the semiconductor and $E_{min,i}$ is the conduction minima for the $i$th valley.

- The calculated values for potential dependant electron and hole concentration were substituted into the Poisson equation, which was numerically integrated to calculate the electric field at the surface of the semiconductor, $E_s$.

- The charge components due to electrons, holes and ionised dopants were calculated using Gauss' law, which were then differentiated with respect to surface potential, $\psi_s$, to yield $C_d$ and $C_{inv}$.

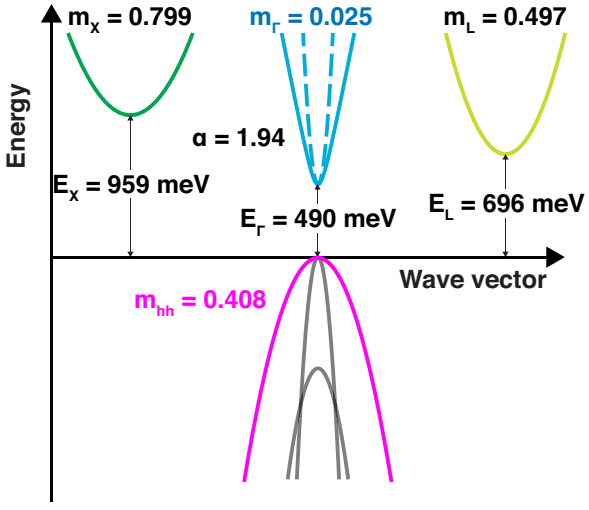


Figure 5.26: Schematic band structure of $In_xGa_{1-x}Sb$. An InSb mole fraction of x=0.275 to fit $E_{g,\Gamma}$ to 490 meV as measured by photoluminescence spectroscopy (PL), shown in Fig. 5.3(d).

### 5.6.1 Calculating the Total Charge Density, $Q_s(\psi_s)$

The following proceeds with an example of calculating the total electron concentration as a function of $\psi_s$. Details of calculating the hole concentration have been omitted as they are entirely equivalent, except that only one valley, the HH valley, was considered. Calculation of the integral shown in Eqn. 5.8 yields the following expression for the potential dependant electron concentration for a 3-d semiconductor with non-parabolic bands:

$$n(\psi(x)) = \frac{2}{\sqrt{\pi}} \sum_{i=\Gamma,L,X} N_i \times \int_0^\infty \frac{\sqrt{\eta}(1+\alpha_i\eta_i)(1+2\alpha_i\eta_i)}{exp[\eta_i - \frac{q\psi(x)}{k_BT} + \delta_i]} d_{\eta_i} \tag{5.9}$$

where $N_i$ is the effective density of states of the $i$th valley, given by

$$N_i = 2(\frac{(2\pi m_i^* kT)^{3/2}}{h^2}) \tag{5.10}$$

and $\eta_F = E_F - E_C/kT$, $m_i^*$ is the effective mass of the $i$th valley, $\kappa$ is Boltzmann's constant, $T$ temperature, and $h$ Planck's constant. For all valleys other than the $\Gamma$ valley, $\alpha$ was set to zero. Integrating the Poisson equation yields an expression for the electric field as a function of position, $E(x)$, which allows for the calculation of the electric field at the surface, $E_s$, where x=0 and $\psi = \psi_s$. Applying Gauss' Law gives the total charge in the semiconductor as given by Equation 5.11:

$$Q_s(\psi_s) = \varepsilon_s E_s = -Sign(\psi_s)\sqrt{2q\varepsilon_s \int p(\psi(x)) - n(\psi(x)) - N_A^- + N_D^+ . d\psi} \tag{5.11}$$

Substitution of Equation 5.9 into Equation 5.11, in addition to the equivalent expression for $p(\psi(x))$, allows $Q_s(\psi_s)$ to be solved and this is plotted as the black dashed line in Fig. 5.27.

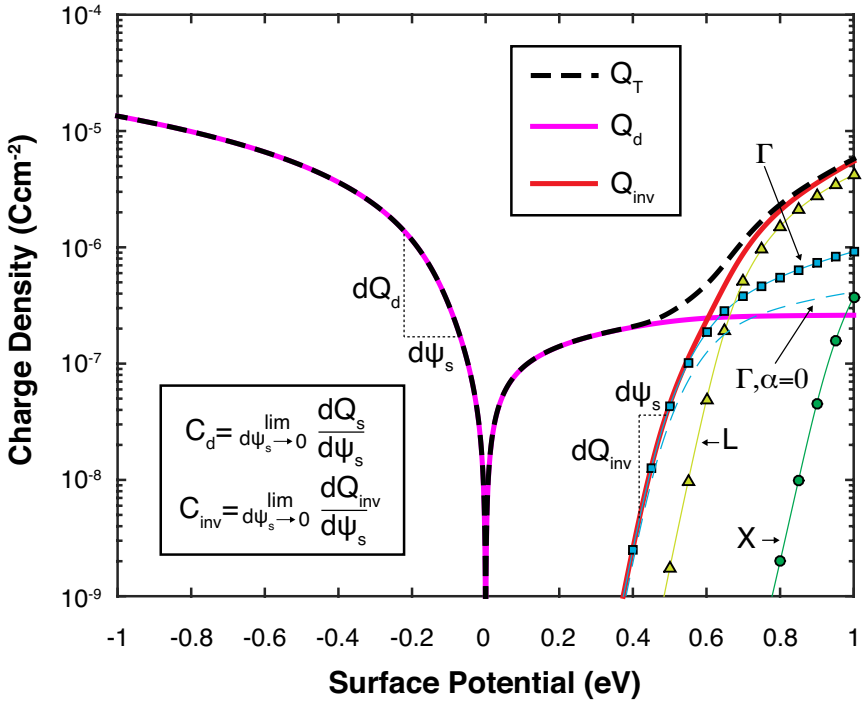


Figure 5.27: Calculated charge destiny as a function of surface potential, as calculated by considering the population of HH, Γ,L and X valleys, with carriers obeying Fermi-Dirac statistics, and a correction factor to account for the non-parabolicity of the Γ valley included. The components of $Q_T$, $Q_{\mathrm{d}}$ and $Q_{\mathrm{inv}}$, are depicted, as well as the individual components $Q_{\mathrm{inv}}$ due to the population of the valleys considered. The variation of inverted charge in the Γ valley with and without the non-parabolicity correction factor, $\alpha$, is also illustrated.

### 5.6.2 Separating Charge Components and Calculating $C_{\mathbf{d}}$ and $C_{\mathbf{inv}}$

The charge components of $Q_{\mathrm{s}}$, charge due to depletion, $Q_{\mathrm{d}}$, and charge due to inversion, $Q_{\mathrm{inv}}$, where $Q_s(\psi_s) = Q_{dep}(\psi_s)+Q_{inv}(\psi_s)$, can be calculated as follows. The inversion charge due to electrons (in the example of a p-type semiconductor) is given by the product of the elementary charge, $q$, and the integral of the electron density as a function of depth:

$$Q_{inv} = q\int n(\psi(x)).dx \tag{5.12}$$

Changing the variable of integration from $x$ to $\psi$ gives:

$$Q_{inv} = q\int n(\psi(x)).dx = q\int n(\psi)\frac{dx}{d\psi}.d\psi \tag{5.13}$$

$d\psi/dx$ is known from the numerical integration of the Poisson equation:

$$\frac{d\psi}{dx} = -E(x) = \sqrt{\frac{2}{\varepsilon_s}\int[p-n-N_A]} \tag{5.14}$$

which can be substituted into Equation.5.13 to give:

$$Q_{inv} = q\int \frac{n(\psi)}{\sqrt{\frac{2}{\varepsilon_s}\int[p-n-N_A]}}.d\psi \tag{5.15}$$

$Q_d$ can then be calculated by subtracting $Q_{inv}$ from $Q_s$. Fig. 5.27 plots the charge components $Q_{\mathrm{d}}$ and $Q_{inv}$ calculated in this manner as a function of $\psi_s$, as wells a the charge components due to each of the considered conduction band valleys: $Q_\Gamma$,$Q_L$ and $Q_x$. $C_{\mathrm{d}}$ and $C_{\mathrm{inv}}$ can then be calculated as a function of $\psi$ as the derivative of the charge components $Q_{\mathrm{d}}$ and $Q_{\mathrm{inv}}$.

$$C_{\mathrm{d}}(\psi) = \frac{dQ_{\mathrm{d}}(\psi)}{d\psi} \tag{5.16}$$

$$C_{\mathrm{inv}}(\psi) = \frac{dQ_{\mathrm{inv}}(\psi)}{d\psi} \tag{5.17}$$

## 5.7 Modelling the $In_{0.3}Ga_{0.7}Sb$-$Al_2O_3$ Interface II: incorporating $D_{it}$

The response due to interface traps is incorporated into the model by considering the Shockley-Read-Hall (SRH) statistics for the capture and emission of carriers with discrete interface traps distributed throughout the bandgap in energy [49]. The following is intended to provide an overview of the derivation of the equivalent circuit components which model the above mechanism, such that the origin of their analytical expressions are clear to the reader.

Fig 5.28 illustrates the possible processes by which carriers can be captured by and emitted from interface traps with a single discrete energy level, where the direction of the arrow indicates the direction of electron transition: (a) electron capture from the conduction band, (b) electron emission to the conduction band, (c) hole capture from the valence band (the transfer of an electron from a trap state to the valence band) and (d) hole emission to the valence band (the transfer of an electron from the valence band to the trap state). The rate by which carriers are captured by and emitted from traps is described by SRH theory [59]. The capture rate for carriers is simply proportional to their concentration and the density of trap states which are unoccupied and therefore available for capture. This is expressed by:

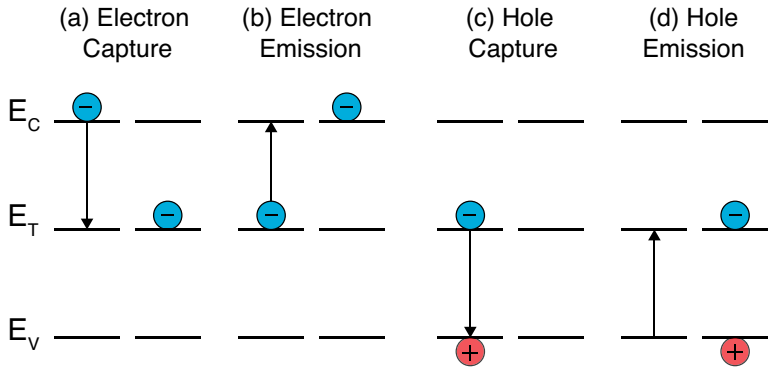


Figure 5.28: Illustration of the capture and emission of electrons and holes via Shockley-Read-Hall interface trap centres.

$$r_{c,n}(E_T) = c_n n_s D_{it}(1 - F_T(E_T)) \tag{5.18}$$

$$r_{c,p}(E_T) = c_p p_s D_{it} F_T(E_T) \tag{5.19}$$

where $r_{c,n}$ and $r_{c,p}$ are the rate of capture of electrons and holes respectively, $c_n$ and $c_p$ the electron and hole capture probabilities, $n_s$ and $p_s$ are the density of electrons and holes at the surface, $D_{it}$ is the density of interface traps and $F_T$ is the Fermi-Dirac function as a function of the

trap energy, $E_T$,corresponding to its probability of occupation. The electron capture probability coefficient is given by:

$$c_n = \sigma_n v_{th,n} \tag{5.20}$$

where $v_{th,n}$ is the thermal velocity of an electron, and $\sigma_n$ is a parameter called the 'electron capture cross section' which corresponds to the cross sectional area in space over which an electron can interact with a trap. The expression for $c_p$ is exactly equivalent.

Conversely, the emission rate is proportional to the concentration of occupied traps, and the density of unoccupied states in the semiconductor, as given by:

$$r_{e,n}(E_T) = e_n D_{it} F_T(E_T) \tag{5.21}$$

$$r_{e,p}(E_T) = e_p D_{it}(1 - F_T(E_T)) \tag{5.22}$$

where $e_n$ and $e_p$ are the electron and hole emission probabilities, which are dependant on the DOS of the conduction and valence bands respectively, as well as the distance of the trap from the band edge to which it is emitting.

$$e_n = \sigma_n v_{th,n} n_s \frac{1 - F_T}{F_T} \tag{5.23}$$

If one considers a trap of donor type, (this is an arbitrary choice, acceptors are equally applicable) it can exist in one of two states, neutral or positive, by participating in the following reactions:

$$D_{it}^{+} + e^{-} \rightleftharpoons D_{it}^{0} \tag{5.24}$$

$$D_{it}^{0} + h^{+} \rightleftharpoons D_{it}^{+} \tag{5.25}$$

where $D_{it}^{+}$ and $D_{it}^{0}$ denote the density of positively charged and neutral interface traps respectively. The rate with which the positive charge density changes due the changing state of interface traps, therefore, is simply the sum of the electron and hole capture and emission rates:

$$\frac{dD_{\mathrm{it}}^{+}}{dt} = r_{\mathrm{e,n}}(t) + r_{\mathrm{c,p}} - (r_{\mathrm{c,n}} + r_{\mathrm{e,p}}) \tag{5.26}$$

The total admittance of the MOS capacitor is the ratio of the total ac current (or equivalently charge variation with time) to the ac gate voltage. The equivalent circuit in Fig. 5.29 (a) is derived in this way by expressing the total charge variation in the semiconductor as the sum of the charge variation due to single-level interface traps (as given by Eqn. 5.26), depletion charge and inversion charge, where

$$G_{\mathrm{n}} = \frac{q^2}{\mathrm{kT}} D_{\mathrm{it}} \tau_{\mathrm{n}} F_{\mathrm{T}} (1 - F_{\mathrm{T}}) \tag{5.27}$$

$$G_{\mathrm{p}} = \frac{q^2}{\mathrm{kT}} D_{\mathrm{it}} \tau_{\mathrm{p}} F_{\mathrm{T}} (1 - F_{\mathrm{T}}) \tag{5.28}$$

$$C_{\mathrm{T}} = \frac{q^2}{\mathrm{kT}} D_{\mathrm{it}} F_{\mathrm{T}} (1 - F_{\mathrm{T}}) \tag{5.29}$$

and $\tau_{\mathrm{p}}$ and $\tau_{\mathrm{n}}$ are given by

$$\tau_p = (p_{\mathrm{s}} V_{\mathrm{th,p}} \sigma_{\mathrm{p}})^{-1} \tag{5.30}$$

$$\tau_n = (n_{\mathrm{s}} V_{\mathrm{th,n}} \sigma_n)^{-1} \tag{5.31}$$

This circuit is somewhat intuitive: a carrier is captured through a resistance, the magnitude of which increases with increasing distance from the band edge from which the carrier is being captured, and the trapped carrier is represented by a capacitance which stores its charge. This circuit can be extended to include interface traps distributed throughout the band gap in energy by summing a distributed network of branches, each of which corresponds to a unique discrete trap energy and has a capacitance and capture resistances specific to the energy of the trap represented. In order to sum the associated trap admittance components, it is required to first perform a Wye-Delta transformation so that the admittances can be added in parallel. This transformation is depicted in Fig. 5.30 (a) for each branch of Fig. 5.29 (b), where

$$C_{\mathrm{Tn},L} = \frac{j\omega C_{\mathrm{T},L} G_{\mathrm{n},L}}{j\omega C_{\mathrm{T},L} + G_{\mathrm{n},L} + G_{\mathrm{p},L}} \tag{5.32}$$

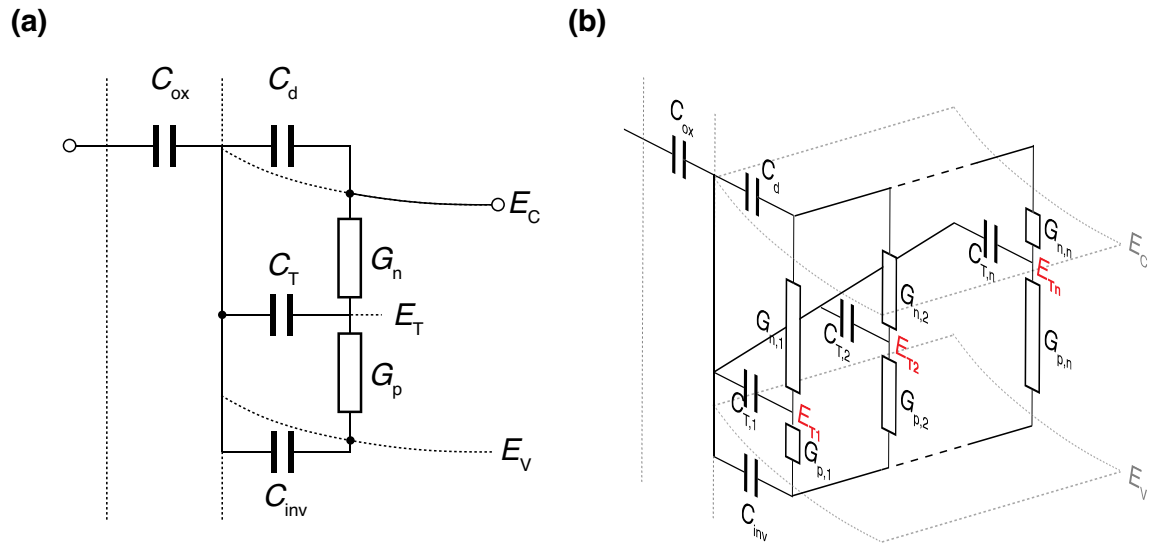


Figure 5.29: (a) Schematic diagram of the equivalent circuit of a single-level discrete interface trap. (b) Extension of (a) by incorporating multiple branches of discrete energy levels distributed throughout the bandgap.

$$C_{\mathrm{Tp},L} = \frac{j\omega C_{\mathrm{T},L} G_{\mathrm{p},L}}{j\omega C_{\mathrm{T},L} + G_{\mathrm{n},L} + G_{\mathrm{p},L}} \tag{5.33}$$

$$G_{\mathrm{it},L} = \frac{G_{\mathrm{n},L} G_{\mathrm{p},L}}{j\omega C_{\mathrm{T},L} + G_{\mathrm{n},L} + G_{\mathrm{p},L}} \tag{5.34}$$

The subscript '$L$' denotes the $L$th trap level. These admittances can now simply be summed to from the lumped equivalent circuit shown in Fig. 5.31 (b).

Given that individual trap energies are not distinguishable experimentally, the discrete interface traps can be treated as continuous in energy, which allows their integration across the bandgap as opposed to the summation of discrete branches. Converting to integrals, the lumped equivalent circuit of Fig.5.30 (b) can be represented by the circuit of Fig.5.31 (a) where

$$C_{\mathrm{Tn}} = qD_{\mathrm{it}}\tau_{\mathrm{n}} \times \int_0^1 (1-F)[[j\omega F(1-F)] + F\tau_{\mathrm{p}}^{-1} + (1-F)f\tau_{\mathrm{n}}^{-1}]^{-1}.dF \tag{5.35}$$

$$C_{\mathrm{Tp}} = qD_{\mathrm{it}}\tau_{\mathrm{p}} \times \int_0^1 (1-F)[[j\omega F(1-F)] + F\tau_{\mathrm{n}}^{-1} + (1-F)f\tau_{\mathrm{p}}^{-1}]^{-1}.dF \tag{5.36}$$

$$G_{\mathrm{it}} = qD_{\mathrm{it}}\tau_{\mathrm{n}}^{-1}\tau_{\mathrm{p}}^{-1} \times \int_0^1 [j\omega F(1-F)] + F\tau_{p}^{-1} + (1-F)f\tau_{\mathrm{n}}^{-1}]^{-1}.dF \tag{5.37}$$

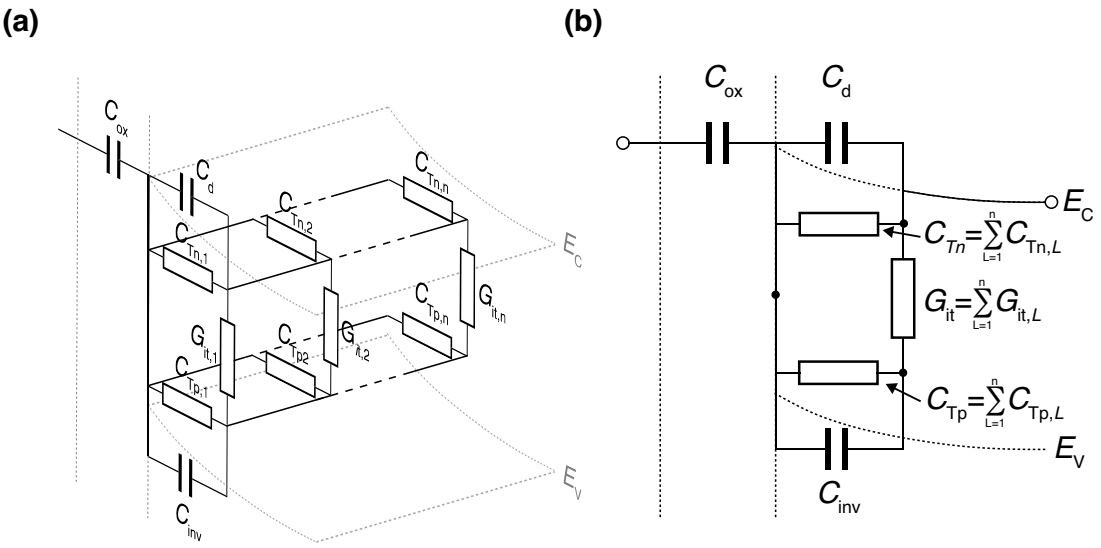


Figure 5.30: (a) Schematic diagram the equivalent circuit shown in Fig.5.29 (a) following a Wye-Delta transformation for each branch. (b) The lumped circuit equivalent of (a) by summing the admittances in parallel.

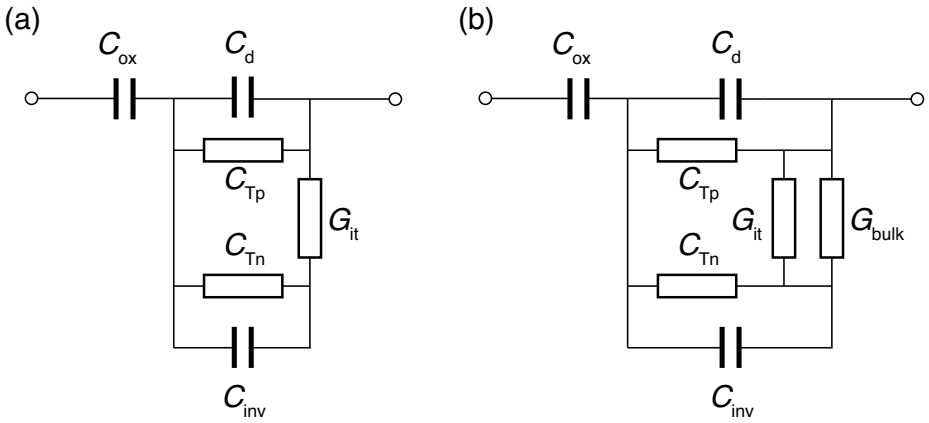


Figure 5.31: The equivalent circuit model from Fig. 5.25 (b) with the inclusion of admittances (a) $C_{\mathrm{Tp}}$, $C_{\mathrm{Tn}}$ and $G_{\mathrm{it}}$ to account for the response due to $D_{\mathrm{it}}$, and $G_{\mathrm{bulk}}$ to account for minority carriers supplied to to the inversion region from the bulk.

The addition of $G_{\mathrm{bulk}}$ in Fig. 5.31 (b) accounts for the supply of minority carriers to the inversion layer via either G-R in the depletion region or diffusion from the bulk, and its value is fitted to experimental data.

## 5.8 Modelling the $In_{0.3}Ga_{0.7}Sb$-$Al_2O_3$ Interface III: incorporated Border Traps

The response due to border traps is incorporated via the distributed bulk oxide trap model of Yuan *etal*. [34]. This model is based on carriers tunnelling in and out of the oxide as they are captured and emitted by border traps distributed throughout the thickness of the dielectric [34]. As with interface traps, a trapping event has an associated conductance and capacitance, $G_{\mathrm{BT}}$ and $C_{\mathrm{BT}}$ respectively. $C_{\mathrm{BT}}$ has exactly the same form as $C_{\mathrm{T}}$ given in Eqn. 5.29 and as is equal to

$$C_{\mathrm{BT}}(\mathrm{E,x}) = \frac{q^2}{\mathrm{kT}} N_{\mathrm{BT}} F_{\mathrm{T}} (1 - F_{\mathrm{T}}) \tag{5.38}$$

where $N_{\mathrm{BT}}$ is the density of border traps per volume per energy ($\mathrm{cm}^{-3}\ \mathrm{eV}^{-1}$). $G_{\mathrm{BT}}$ is related to $C_{\mathrm{BT}}$ by the trap time constant:

$$\frac{C_{\mathrm{BT}}(E,x)}{G_{\mathrm{BT}}(E,x)} = \tau(\mathrm{x}) = (p_{\mathrm{s}} V_{\mathrm{th,p}} \sigma_{\mathrm{p}})^{-1} e^{2\kappa_{\mathrm{p}} x} \tag{5.39}$$

which incorporates the tunnelling attenuation factor, $\kappa$, as calculated from

$$\kappa_{\mathrm{p}} = \frac{\sqrt{2m^*(E - E_{\mathrm{v}}^{\mathrm{ox}})}}{\hbar} \tag{5.40}$$

Using 5.38,5.39 and 5.40, the expression for the total admittance associated with the trapping of a carrier at a given depth in the oxde and at a given energy, $Y_{BT}$, can be derived. As with interface traps, this expression can then be integrated over energy to yield the total admittance response due to border traps at a given depth into the oxide, which yields:

$$Y_{BT}(x) = \frac{q^2 N_{BT} ln(1 + j\omega\tau(x))}{\tau(x)} \tag{5.41}$$

The distributed border trap model divides the oxide thickness into an finite number of incremental capacitances, $\Delta C_{\mathrm{ox}}$, with each increment connected to corresponding depth dependant $Y_{\mathrm{BT}}$ admittance, as shown in Fig. 5.32 (b).

The final equivalent circuit used to model the InGaAsb-$Al_2O_3$ interface is shown in Fig. 5.33, which utilised the Yuan model for border trap interactions with both majority and minority carriers. Finally, series resistance, $R_{\mathrm{s}}$, and parasitic series inductance, $L_{\mathrm{s}}$, were included to accurately model the experimental dispersion. $R_{\mathrm{s}}$ was extracted from measured data using Eqn. 5.2 and $L_{\mathrm{s}}$

was used as a fitting parameter. An example of experimental dispersion, and the components due to border traps, $R_s$ and $L_s$ is shown in Fig. 5.33 (b). Excluding $L_s$, this equivalent circuit model has been demonstrated previously to be able to accurately reproduce experimental data from InAs, InGaAs and InP MOS capacitors [37, 38].

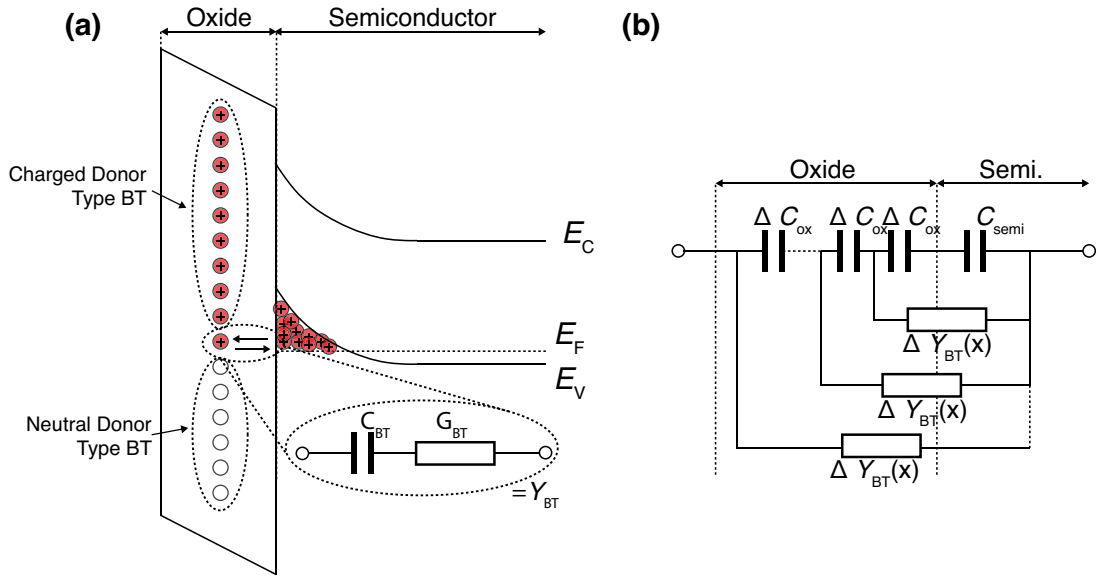


Figure 5.32: (a) Illustration of the trapping of inverted carriers by border traps. (b) The corresponding equivalent circuit.

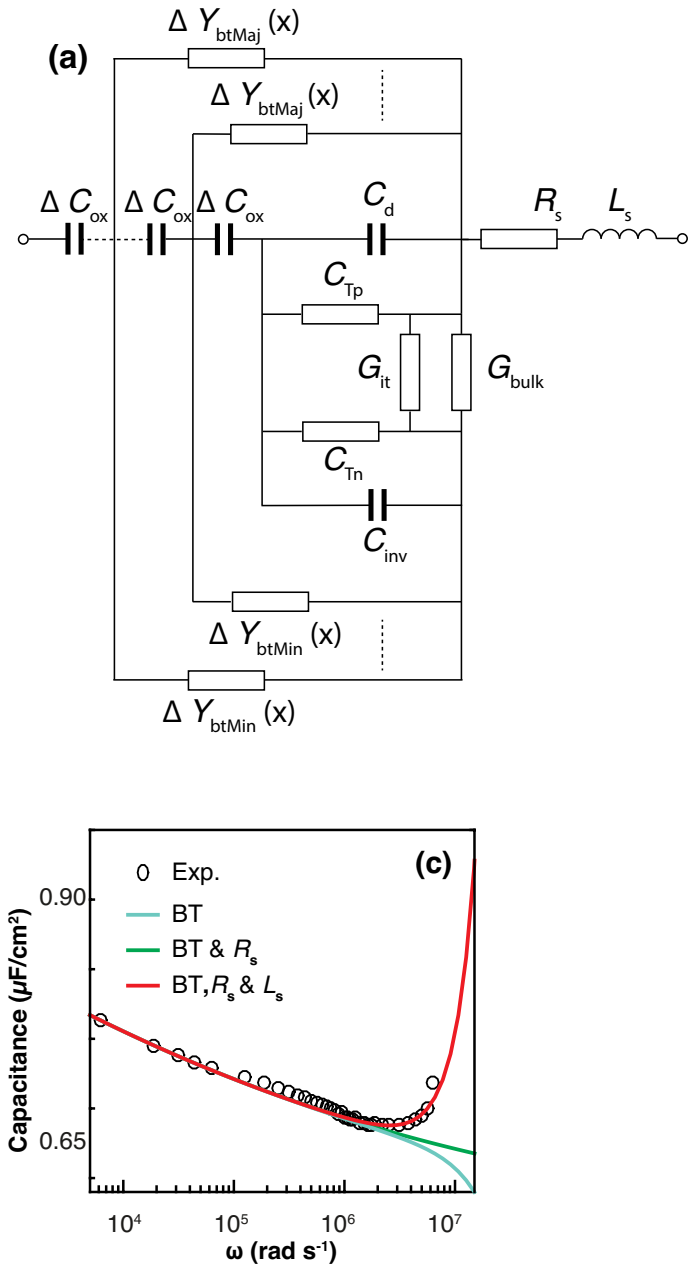


Figure 5.33: (a) The complete equivalent circuit model used to simulate the InGaSb-$Al_2O_3$ interface. (b) The impact of $R_s$, $L_s$ and BT on frequency dispersion in accumulation.

## 5.9 Simulation Results

Excellent fits to the experimental CV and GV data for both p and n-type samples was achieved, shown in Fig. 5.34, with interface trap and border trap parameters common to both and shown in Fig. 5.35. The best fit was achieved with interface trap hole capture cross sections that decayed away from the valence band edge, the values for which are also included in Fig. 5.35. The doping concentrations were extracted as $N_D = 1.7 \times 10^{17}$ cm$^{-3}$ and $N_A$=2.5 $\times 10^{17}$ cm$^{-3}$. It is not surprising the these values are higher than the nominal doping concentration, as it is likely the dopants from the highly doped buffer layer will have diffused into the capacitor layer. Furthermore it has been shown that $H_2$ p-dopes antimonides which may explain the discrepancy between p and n-type. Low $D_{it}$ across the band gap was extracted, with a minimum value of $1.73 \times 10^{12}$ cm$^{-2}$ eV$^{-1}$ located 110 meV below the conduction band edge. The border trap distribution was fitted with two gaussians centred close to the band edges. Border trap densities were extracted with peak magnitudes of $3 \times 10^{19}$ cm$^{-3}$ eV$^{-1}$ near the valence band edge, and $6.5 \times 10^{19}$ cm$^{-3}$ eV$^{-1}$ near the conduction band edge. The extracted $V_g$-$\psi_s$ relationship is plotted in Fig. 5.36 and shows an unpinned Fermi level that can move into both valence and conduction bands.

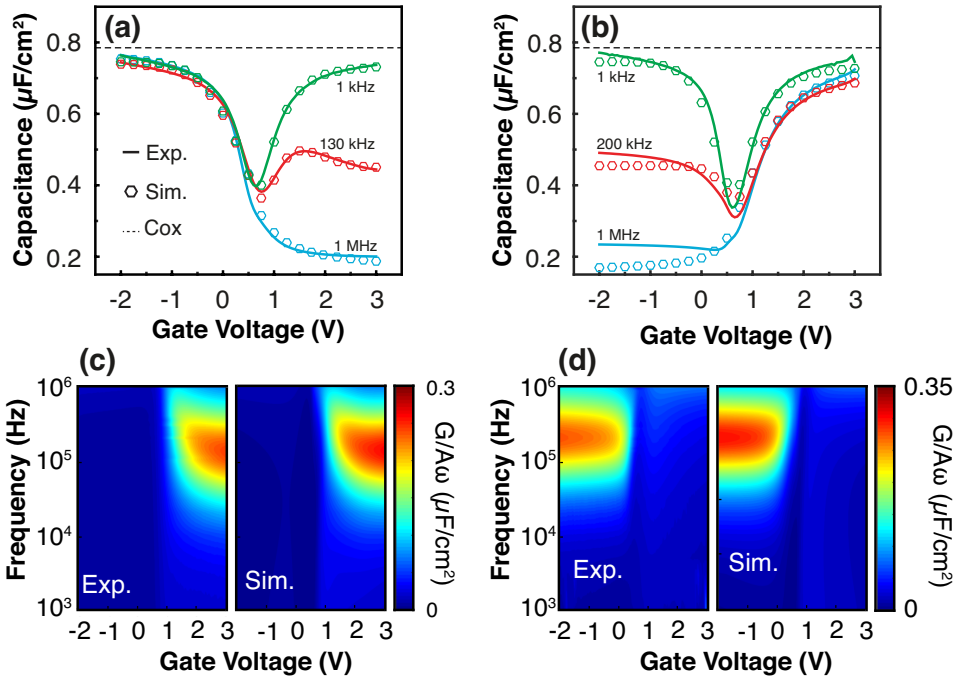


Figure 5.34: Comparison between experimental and simulated multifrequency, room temperature, CV (a,c) and GV (b,d) responses for p-type and n-type capacitors processed with the optimal $H_2$ plasma process.

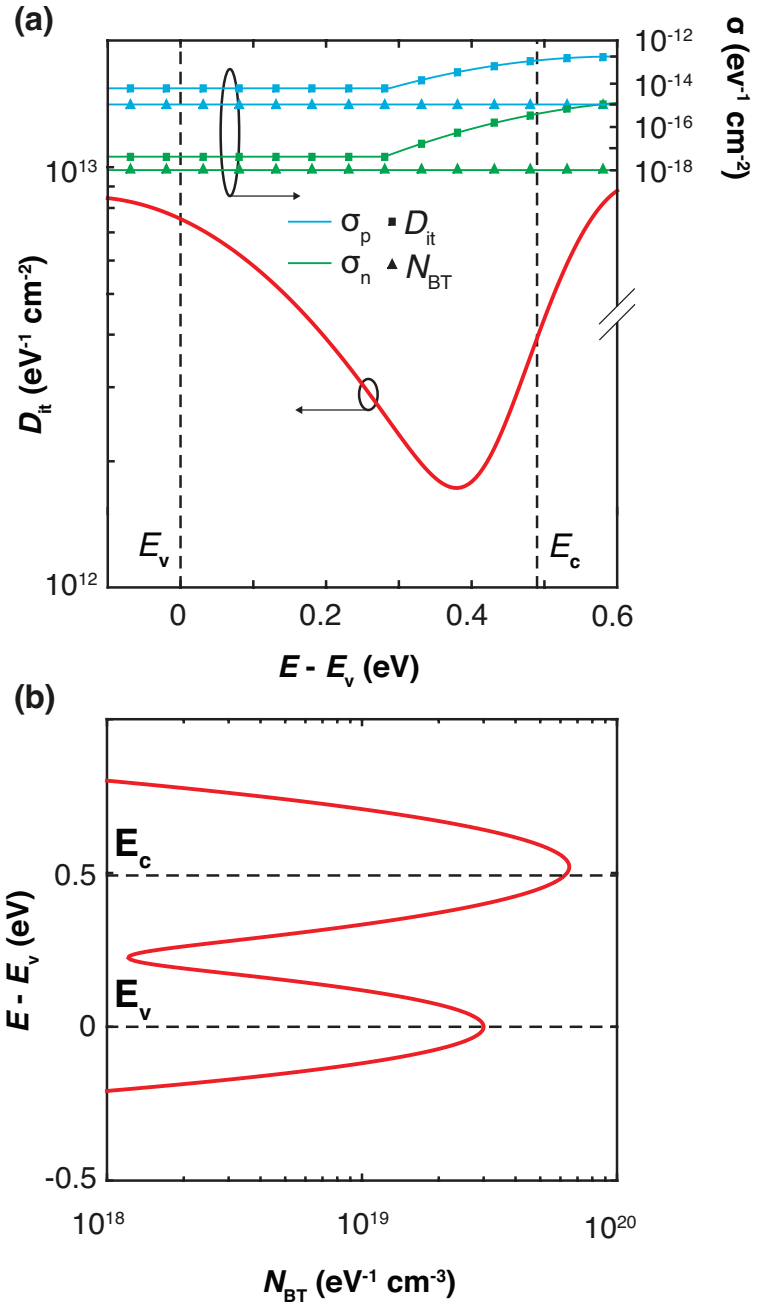


Figure 5.35: Input parameters to the full interface state model: (a) interface state density, $D_{it}(E)$, and (b) border trap density, $N_{bt}(E,x)$.

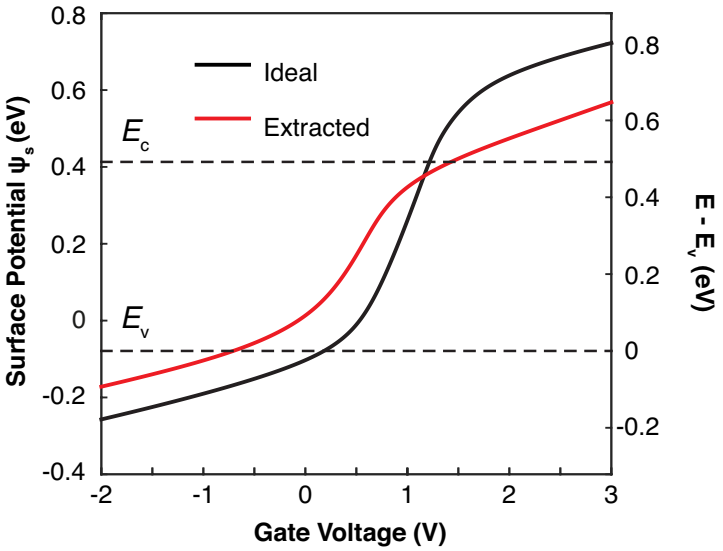


Figure 5.36: Comparison of the extracted $V_g$-$\psi_s$ relationship with the ideal.

## 5.10 Chapter summary and future work

### 5.10.1 Summary

In summary, a systematic array of experiments has been undertaken to assess the impact of $H_2$ plasma cleaning on the $In_{0.3}Ga_{0.7}Sb$-$Al_2O_3$ interface. The parameter space was explored in terms of plasma power, exposure time, the inclusion of in-situ TMA pre cleaning, and annealing in forming gas. XPS measurements found ex-situ HCl cleaning to effectively reduce all In,Ga and Sb sub oxides. In-situ TMA exposure was found to completely remove In and Sb sub oxides, and reduce the Ga-O content to $< 8$ % (Ga-O:InGaSb). Interestingly, the addition of $H_2$ plasma cleaning was found to marginally increase the Ga-O content to $\sim 10\%$, however, offered significant improvements to the electrical properties of the interface. Analysis of the oxide composition found that it varied depending on the surface cleaning treatment and the $AlO_x$ content increased significantly towards the semiconductor interface of $H_2$ plasma cleaned samples, which appears to be the origin of improvement to the electrical properties. Further optimisation of the process yielded significantly improved qualitative electrical properties, and comprehensive analysis in terms of variable temperature measurements, extraction of the activation energy (for the p-type sample) and the observation of characteristics which evidence peripheral inversion unequivocally demonstrated that the interface of this optimal process was unpinned and that a genuine inversion response was demonstrated for both p and n-type MOSCAPs. Simulation of the interface for the optimal process validated the above findings, and demonstrated an unpinnned Fermi level which could move freely across the entire bandgap. Interface state and border trap densities were extracted, with a minimum $D_{\text{it}}$ of $1.73 \times 10^{12}\text{cm}^{-2}$ $\text{eV}^{-1}$ located at $\sim$110 meV below the conduction band edge and peak border trap densities approximately aligned with the valence and conduction band edges of $3 \times 10^{19}$ $\text{cm}^{-3}$ $\text{eV}^{-1}$ and $6.5 \times 10^{19}$ $\text{cm}^{-3}$ $\text{eV}^{-1}$ respectively.

### 5.10.2 Future Work

With the optimal gate stack process demonstrated in this chapter capable of yielding an unpinned InGaSb-$Al_2O_3$ interface, subsequent work should implement this gate stack module within a MOSFET process for both p and n-type devices. The junctionless MOSFET architecture is proposed for this intial endeavour due to its processing simplicity, which will introduce as few additional, potentially detrimental processing steps to that of the optimal gate stack process in isolation. To this end, a total of 6 wafers have been grown by Dr Matthew Steer using MBE, the layer structures for which are shown in Figs. 5.37 and 5.38. The structure of each of these wafers is equivalent to that discussed in Chapter 4 for the InGaAs Junctionless FinFET; with a wide bandgap buffer utilised to confine a thin channel. Both lattice matched and strained (1.2%

compressive) channel variants were grown on GaAs substrates for both p and n-type channels (Fig. 5.37). The channel thickness was chosen as 20 nm to ensure that it was below the pseudomorphic critical thickness [60]. Furthermore, this channel thickness should be sufficiently thin such that planar devices can be fabricated, simplifying the process further. A channel doping concentration of $1 \times 10^{18}$ cm$^{-3}$ was chosen as a trade-off between on-state performance and the ability to turn the device off; a decision which erred towards off-state performance and a relatively low doping concentration as poor subthreshold performance plagues the antimonide FET literature [61]. Further lattice matched wafers were grown, yet on a Si substrate, shown in Fig. 5.38.

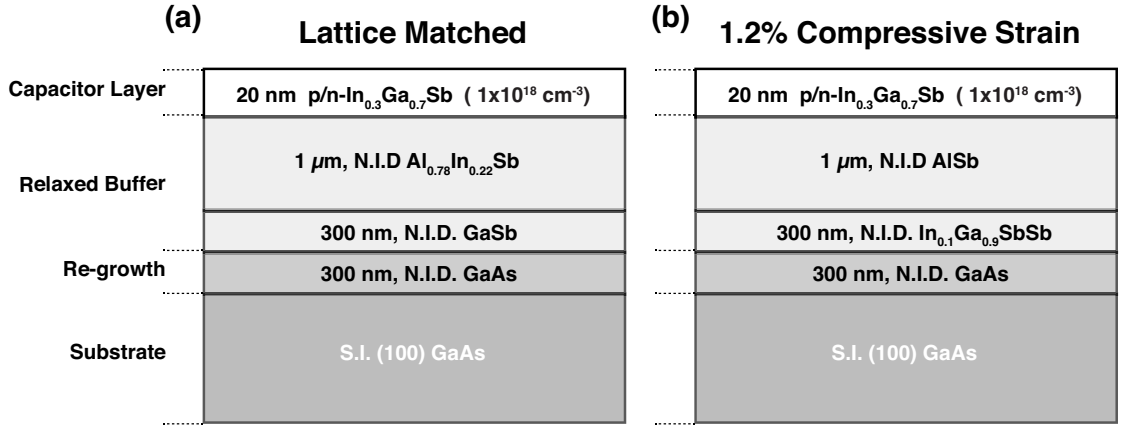


Figure 5.37: Schematic diagram of the epitaxial layer structures grown by MBE on III-V substrates (GaAS) for the fabrication of InGaSb junctionless MOSFETs. (a) Utilises an AlInSb buffer which is lattice matched to the channel. (b) Utilises an AlSb buffer layer to compressively strain the InGaSb channel. Both p and n-type variants were grown.

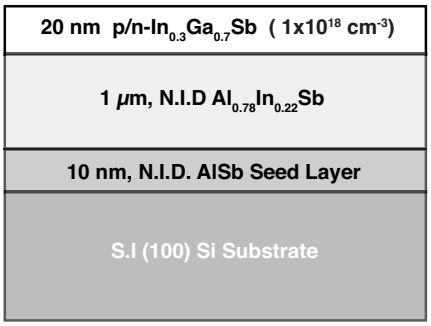


Figure 5.38: Schematic diagram of the epitaxial layer structures grown by MBE on Si substrates for the fabrication of InGaSb junctionless MOSFETs. Both p and n-type variants were grown.

The proposed fabrication process is shown in Fig. 5.39. It utilises a wrap around gate architecture which removes the requirement to isolate the devices by forming a mesa [62]. Should the above process and material yield working devices, the immediate focus should be to fabricate and demonstrate a common channel InGaSb CMOS inverter, the potential for which is the most desirable asset of InGaSb.

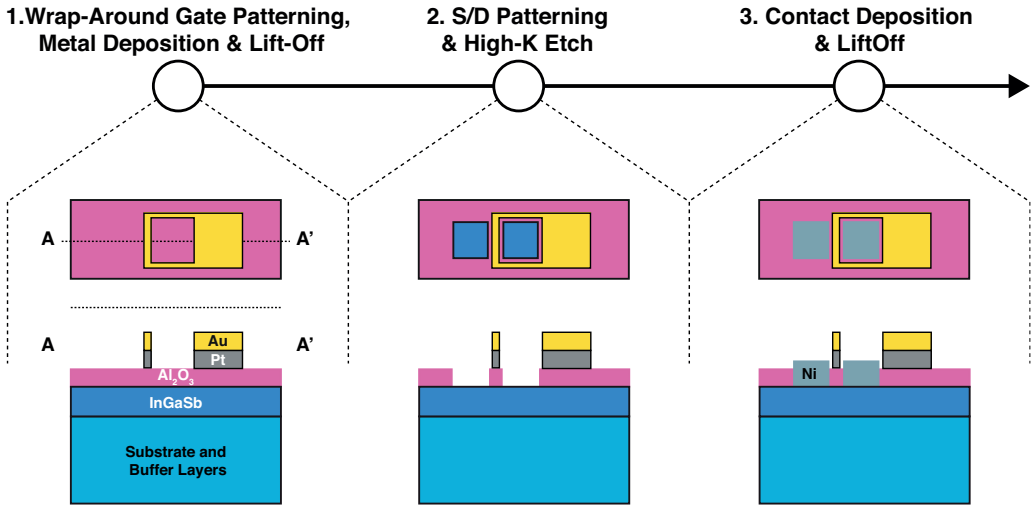


Figure 5.39: Proposed fabrication process flow for the fist iteration of InGaSb junctionless FETs.

In addition to the fabrication of InGaSb MOSFETs, further investigation should be undertaken into the physical understanding of the improvements in electrical properties of the InGaSb-$Al_2O_3$ interface due to the inclusion of $H_2$ plasma cleaning. To this end, further XPS measurements should taken for the optimal sample with the results compared to those measured thus far. In particular, an assessment of the impact of the various cleaning processes on Sb-Sb dimers should be researched as a possible source of improvement: for III-V arsenides, it has been shown that reducing the presence of As-As dimers at dielectric interface is critical in reducing $D_{it}$ [63].

### 5.10.3 A Proposed Novel MOS Test Structure

Confirming genuine surface inversion for the optimal p-type sample proved greatly problematic and necessitated comprehensive analysis requiring variable temperature measurements and simulation. This was not the case for the n-type sample, however, and genuine surface inversion was unambiguously observed due to the happenstance that charge in the oxide inverted the surface beyond the gate. It is therefore proposed that a novel MOS test structure could be fabricated that purposefully induces peripheral inversion around the gate, facilitating the identification of genuine surface inversion for any sample by simply measuring 2 or more devices with different diameters. The proposed test structure to achieve this and corresponding process flow is shown in Fig. 5.40. This requires only 2 additional processing steps to that of a normal MOS capacitor, does not require accurate alignment, and is greatly simpler than fabricating a device. The test structure would work by applying a DC bias to the outer gate, labelled 'DC gate', in order to invert the surface of the semiconductor surrounding the inner gate, labelled 'AC gate'. Otherwise the MOSCAP would simply be measured as normal. It is additionally thought that this test structure would be applicable for any bandgap semiconductor. This has interesting ramifications for wide bandgap materials, such as GaN, where a minority carrier response cannot otherwise be observed at room temperature at experimental frequencies. With this test structure, however, minority carriers could readily diffuse in and drift out of the inversion layer in phase with an applied AC bias.

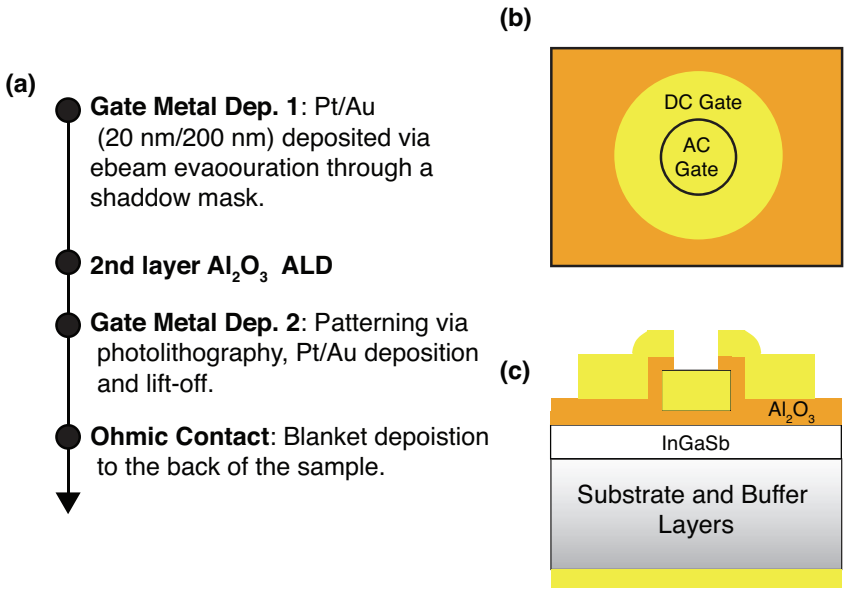


Figure 5.40: Proposed novel test structure which explicitly measures an inversion response and bandgap semiconductor.

# Bibliography


[1] "InSb Electrical Properties", IOFFE, 2018. [Online]. Available: http://www.ioffe.ru/SVA/NSM/Semicond/InSb/electric.html

[2] "GaSb Electrical Properties", IOFFE, 2018. [Online]. Available: http://www.ioffe.ru/SVA/NSM/Semicond/GaSb/electric.html

[3] Z. Yuan, A. Nainani, A. Kumar, X. Guan, B. R. Bennett, J. B. Boos, M. G. Ancona, and K. C. Saraswat, "InGaSb: Single channel solution for realizing III-V CMOS," in *Symposium on VLSI Technology*, 2012, pp. 185–186.

[4] B. R. Bennett, M. G. Ancona, J. B. Boos, and B. V. Shanabrook, "Mobility enhancement in strained p-InGaSb quantum wells," *Applied Physics Letters*, vol. 91, no. 4, pp. 10–13, 2007.

[5] Z. Yuan, A. Kumar, C.-Y. Chen, A. Nainani, B. R. Bennett, J. B. Boos, and K. C. Saraswat, "Antimonide-Based Heterostructure p-Channel MOSFETs With Ni-Alloy Source/Drain," *IEEE Electron Device Letters*, vol. 34, no. 11, pp. 1367–1369, 2013.

[6] K. Nishi, M. Yokoyama, H. Yokoyama, T. Hoshi, H. Sugiyama, M. Takenaka, and S. Takagi, "High hole mobility front-gate InAs/InGaSb-OI single structure CMOS on Si," in *Symposium on VLSI Technology*, 2015, pp. T174–T175.

[7] J. Nah, H. Fang, C. Wang, K. Takei, M. H. Lee, E. Plis, S. Krishna, and A. Javey, "III-V complementary metal-oxide-semiconductor electronics on silicon substrates," *Nano Letters*, vol. 12, no. 7, pp. 3592–3595, 2012.

[8] A. Nainani, T. Irisawa, Z. Yuan, Y. Sun, T. Krishnamohan, M. Reason, B. R. Bennett, J. B. Boos, M. G. Ancona, Y. Nishi, and K. C. Saraswat, "Development of high-k dielectric for antimonides and a sub 350°C III-V pMOSFET outperforming Germanium," in *International Electron Devices Meeting*, 2010, pp. 6.4.1–6.4.4.

[9] A. Nainani, T. Irisawa, Z. Yuan, B. R. Bennett, J. B. Boos, Y. Nishi, and K. C. Saraswat, "Optimization of the $Al_2O_3$/GaSb Interface and a High-Mobility GaSb pMOSFET," *IEEE Transactions on Electron Devices*, vol. 58, no. 10, pp. 3407–3415, 2011.

[10] A. Nainani, Z. Yuan, A. Kumar, B. R. Bennett, J. B. Boos, and K. C. Saraswat, "III-Sb MOSFETS: Opportunities and Challenges," *ECS Transactions*, vol. 45, no. 4, pp. 91–96, 2012.

[11] K. Takei, M. Madsen, H. Fang, R. Kapadia, S. Chuang, H. S. Kim, C.-H. Liu, E. Plis, J. Nah, S. Krishna, Y.-L. Chueh, J. Guo, and A. Javey, "Nanoscale InGaSb Heterostructure Membranes on Si Substrates for High Hole Mobility Transistors," *Nano Letters*, vol. 12, no. 4, pp. 2060–2066, 2012.

[12] A. Greene, S. Madisetti, P. Nagaiah, V. Tokranov, M. Yakimov, R. Moore, and S. Oktyabrsky, "(Invited) InGaSb MOSFET Channel on Metamorphic Buffer: Materials, Interfaces and Process Options," *ECS Transactions*, vol. 53, no. 1, pp. 149–160, 2013.

[13] A. Nainani, Y. Sun, T. Irisawa, Z. Yuan, M. Kobayashi, P. Pianetta, B. R. Bennett, J. Brad Boos, and K. C. Saraswat, "Device quality Sb-based compound semiconductor surface: A comparative study of chemical cleaning," *Journal of Applied Physics*, vol. 109, no. 11, p. 114908, 2011.

[14] L. Zhao, Z. Tan, R. Bai, N. Cui, J. Wang, and J. Xu, "Effects of Sulfur Passivation on GaSb Metal-Oxide-Semiconductor Capacitors with Neutralized and Unneutralized $(NH_4)_2S$ Solutions of Varied Concentrations," *Applied Physics Express*, vol. 6, no. 5, p. 056502, 2013.

[15] D. M. Zhernokletov, H. Dong, B. Brennan, M. Yakimov, V. Tokranov, S. Oktyabrsky, J. Kim, and R. M. Wallace, "Surface and interfacial reaction study of half cycle atomic layer deposited $HfO_2$ on chemically treated GaSb surfaces," *Applied Physics Letters*, vol. 102, no. 13, p. 131602, 2013.

[16] Z. Y. Liu, B. Hawkins, and T. F. Kuech, "Chemical and structural characterization of GaSb(100) surfaces treated by HCl-based solutions and annealed in vacuum," *Journal of Vacuum Science & Technology B*, vol. 21, no. 1, pp. 71–77, 2003.

[17] Z. Tan, L. Zhao, J. Wang, and J. Xu, "Improved Properties of $HfO_2/Al_2O_3$/GaSb MOS Capacitors Passivated with Neutralized $(NH_4)_2S$ Solutions," *ECS Solid State Letters*, vol. 2, no. 8, pp. 61–62, 2013.

[18] U. Peralagu, I. M. Povey, P. Carolan, J. Lin, R. Contreras-Guerrero, R. Droopad, P. K. Hurley, and I. G. Thayne, "Electrical and physical characterization of the $Al_2O_3$/p-GaSb interface for 1%, 5%, 10%, and 22% $(NH_4)_2S$ surface treatments," *Applied Physics Letters*, vol. 105, no. 16, p. 162907, 2014.

[19] L. B. Ruppalt, E. R. Cleveland, J. G. Champlain, S. M. Prokes, J. Brad Boos, D. Park, and B. R. Bennett, "Atomic layer deposition of $Al_2O_3$ on GaSb using in situ hydrogen plasma exposure," *Applied Physics Letters*, vol. 101, no. 23, pp. 1–6, 2012.

[20] E. R. Cleveland, L. B. Ruppalt, B. R. Bennett, and S. M. Prokes, "Effect of an in situ hydrogen plasma pre-treatment on the reduction of GaSb native oxides prior to atomic layer deposition," *Applied Surface Science*, vol. 277, pp. 167–175, 2013.

[21] L. B. Ruppalt, E. R. Cleveland, J. G. Champlain, B. R. Bennett, and S. M. Prokes, "Integration of atomic layer deposited high-k dielectrics on GaSb via hydrogen plasma exposure," *AIP Advances*, vol. 4, no. 12, p. 127153, 2014.

[22] M. Barth, G. Bruce Rayner, S. McDonnell, R. M. Wallace, B. R. Bennett, R. Engel-Herbert, and S. Datta, "High quality $HfO_2$/p-GaSb(001) metal-oxide-semiconductor capacitors with 0.8 nm equivalent oxide thickness," *Applied Physics Letters*, vol. 105, no. 22, p. 222103, 2014.

[23] H. D. Trinh, Y. C. Lin, M. T. Nguyen, H. Q. Nguyen, Q. V. Duong, Q. H. Luc, S. Y. Wang, M. N. Nguyen, and E. Yi Chang, "Demonstrating 1 nm-oxide-equivalent-thickness $HfO_2$/InSb structure with unpinning Fermi level and low gate leakage current density," *Applied Physics Letters*, vol. 103, no. 14, pp. 2–7, 2013.

[24] E. Y. Chang, H. Trinh, and Y. Lin, "Studying of InSb MOS capacitors for post CMOS application," in *28th Symposium on Microelectronics Technology and Devices*, 2013, pp. 1–3.

[25] H.-Y. Chou, V. V. Afanas'ev, M. Houssa, A. Stesmans, L. Dong, and P. D. Ye, "Electron band alignment at the interface of (100)InSb with atomic-layer deposited $Al_2O_3$," *Applied Physics Letters*, vol. 101, no. 8, p. 082114, 2012.

[26] C. H. Hou, M. C. Chen, C. H. Chang, T. B. Wu, C. D. Chiang, and J. J. Luo, "Effects of Surface Treatments on Interfacial Self-Cleaning in Atomic Layer Deposition of $Al_2O_3$ on InSb," *Journal of the Electrochemical Society*, vol. 155, no. 9, p. G180, 2008.

[27] A. Kadoda, T. Iwasugi, K. Nakatani, K. Nakayama, M. Mori, K. Maezawa, E. Miyazaki, and T. Mizutani, "Characterization of $Al_2O_3$/InSb/Si MOS diodes having various InSb thicknesses grown on Si(111) substrates," *Semiconductor Science and Technology*, vol. 27, no. 4, p. 045007, 2012.

[28] H. S. Kim, I. Ok, M. Zhang, F. Zhu, S. Park, J. Yum, H. Zhao, J. C. Lee, P. Majhi, N. Goel, W. Tsai, C. K. Gaspe, and M. B. Santos, "A study of metal-oxide-semiconductor capacitors on GaAs, $In_{0.53}Ga_{0.47}As$, InAs, and InSb substrates using a germanium interfacial passivation layer," *Applied Physics Letters*, vol. 93, no. 6, 2008.

[29] H. D. Trinh, Y. C. Lin, E. Y. Chang, C. T. Lee, S. Y. Wang, H. Q. Nguyen, Y. S. Chiu, Q. H. Luc, H. C. Chang, C. H. Lin, S. Jang, and C. H. Diaz, "Electrical characteristics

of $Al_2O_3$/InSb MOSCAPs and the effect of postdeposition annealing temperatures," *IEEE Transactions on Electron Devices*, vol. 60, no. 5, pp. 1555–1560, 2013.

[30] H. D. Trinh, Y. C. Lin, M. T. Nguyen, H. Q. Nguyen, Q. V. Duong, Q. H. Luc, S. Y. Wang, M. N. Nguyen, and E. Yi Chang, "Demonstrating 1 nm-oxide-equivalent-thickness HfO2/InSb structure with unpinning Fermi level and low gate leakage current density," *Applied Physics Letters*, vol. 103, no. 14, pp. 2–7, 2013.

[31] R. Engel-Herbert, Y. Hwang, and S. Stemmer, "Comparison of methods to quantify interface trap densities at dielectric/III-V semiconductor interfaces," *Journal of Applied Physics*, vol. 108, no. 12, pp. 0–15, 2010.

[32] K. Martens, C. O. Chui, G. Brammertz, B. De Jaeger, D. Kuzum, M. Meuris, M. M. Heyns, T. Krishnamohan, K. Saraswat, H. E. Maes, and G. Groeseneken, "On the correct extraction of interface trap density of MOS devices with high-mobility semiconductor substrates," *IEEE Transactions on Electron Devices*, vol. 55, no. 2, pp. 547–556, 2008.

[33] L. Terman, "An investigation of surface states at a silicon/silicon oxide interface employing metal-oxide-silicon diodes," *Solid-State Electronics*, vol. 5, no. 5, pp. 285 – 299, 1962.

[34] Y. Yuan, B. Yu, J. Ahn, P. C. McIntyre, P. M. Asbeck, M. J. W. Rodwell, and Y. Taur, "A distributed bulk-oxide trap model for $Al_2O_3$ InGaAs MOS devices," *IEEE Transactions on Electron Devices*, vol. 59, no. 8, pp. 2100–2106, 2012.

[35] H. C. Lin, G. Brammertz, K. Martens, G. De Valicourt, L. Negre, W. E. Wang, W. Tsai, M. Meuris, and M. Heyns, "The Fermi-level efficiency method and its applications on high interface trap density oxide-semiconductor interfaces," *Applied Physics Letters*, vol. 94 *(15)*, no. 15, p. 153508, 2009.

[36] S. Monaghan, E. O'Connor, R. Rios, F. Ferdousi, L. Floyd, E. Ryan, K. Cherkaoui, I. M. Povey, K. J. Kuhn, and P. K. Hurley, "Capacitance and conductance for an MOS system in inversion, with oxide capacitance and minority carrier lifetime extractions," *IEEE Transactions on Electron Devices*, vol. 61, no. 12, pp. 4176–4185, 2014.

[37] A. S. Babadi, E. Lind, and L. E. Wernersson, "Modeling of n-InAs metal oxide semiconductor capacitors with high-k gate dielectric," *Journal of Applied Physics*, vol. 116, no. 21, 2014.

[38] G. Brammertz, A. Alian, D. H. C. Lin, M. Meuris, M. Caymax, and W. E. Wang, "A combined interface and border trap model for high-mobility substrate metal-oxide-semiconductor devices applied to $In_{0.53}Ga_{0.47}As$ and InP capacitors," *IEEE Transactions on Electron Devices*, vol. 58, no. 11, pp. 3890–3897, 2011.

[39] M. L. Huang, Y. C. Chang, C. H. Chang, Y. J. Lee, P. Chang, J. Kwo, T. B. Wu, and M. Hong, "Surface passivation of III-V compound semiconductors using atomic-layer-deposition-grown $Al_2O_3$," *Applied Physics Letters*, vol. 87, no. 25, pp. 1–3, 2005.

[40] C. L. Hinkle, A. M. Sonnet, E. M. Vogel, S. McDonnell, G. J. Hughes, M. Milojevic, B. Lee, F. S. Aguirre-Tostado, K. J. Choi, H. C. Kim, J. Kim, and R. M. Wallace, "GaAs interfacial self-cleaning by atomic layer deposition," *Applied Physics Letters*, vol. 92, no. 7, pp. 2006–2009, 2008.

[41] A. D. Carter, W. J. Mitchell, B. J. Thibeault, J. J. M. Law, and M. J. W. Rodwell, "Al2O3 growth on (100) In0.53Ga 0.47as initiated by cyclic trimethylaluminum and hydrogen plasma exposures," *Applied Physics Express*, vol. 4, no. 9, pp. 2000–2002, 2011.

[42] P. D. King, T. D. Veal, M. J. Lowe, and C. F. McConville, "Surface electronic properties of clean and S-terminated InSb(001) and (111)B," *Journal of Applied Physics*, vol. 104, no. 8, pp. 1–8, 2008.

[43] R. Giulian, D. J. Manzo, J. B. Salazar, W. Just, A. M. H. de Andrade, J. R. Schoffen, L. A. B. Niekraszewicz, J. F. Dias, and F. Bernardi, "Structural and electronic characterization of antimonide films made by magnetron sputtering," *Journal of Physics D: Applied Physics*, vol. 50, no. 7, p. 075106, 2017.

[44] W. K. Liu, W. T. Yuen, and R. A. Stradling, "Preparation of insb substrates for molecular beam epitaxy," *Journal of Vacuum Science & Technology B*, vol. 13, no. 4, pp. 1539–1545, 1995.

[45] R. Tessler, C. Saguy, O. Klin, S. Greenberg, E. Weiss, R. Akhvlediani, R. Edrei, and A. Hoffman, "Oxide-free InSb (100) surfaces by molecular hydrogen cleaning," *Applied Physics Letters*, vol. 88, no. 3, pp. 1–3, 2006.

[46] S. McDonnell, B. Brennan, E. Bursa, R. M. Wallace, K. Winkler, and P. Baumann, "Gasb oxide thermal stability studied by dynamic-xps," *Journal of Vacuum Science & Technology B*, vol. 32, no. 4, p. 041201, 2014.

[47] R. Suri, D. J. Lichtenwalner, and V. Misra, "Interfacial self cleaning during atomic layer deposition and annealing of HfO2 films on native (100)-GaAs substrates," *Applied Physics Letters*, vol. 96, no. 11, 2010.

[48] R. Engel-Herbert, Y. Hwang, and S. Stemmer, "Comparison of methods to quantify interface trap densities at dielectric/III-V semiconductor interfaces," *Journal of Applied Physics*, vol. 108, no. 12, pp. 1–4, 2010.

[49] E. H. Nicollian and J. Brews, *MOS (metal oxide semiconductor) physics and technology*, ser. Metal Oxide Semiconductor Physics and Technology. Wiley, 1982.

[50] 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, pp. 2009–2012, 2011.

[51] M. Oszwalldowski and M. Zimpel, "Temperature dependence of intrinsic carrier concentration and density of states effective mass of heavy holes in insb," *Journal of Physics and Chemistry of Solids*, vol. 49, no. 10, pp. 1179 – 1185, 1988.

[52] D. Martin and C. Algora, "Temperature-dependent GaSb material parameters for reliable thermophotovoltaic cell modelling," *Semiconductor Science and Technology*, vol. 19, no. 8, p. 1040, 2004.

[53] E. O'Connor, K. Cherkaoui, S. Monaghan, D. O'Connell, I. Povey, P. Casey, S. B. Newcomb, Y. Y. Gomeniuk, G. Provenzano, F. Crupi, G. Hughes, and P. K. Hurley, "Observation of peripheral charge induced low frequency capacitance-voltage behaviour in metal-oxide-semiconductor capacitors on Si and GaAs substrates," *Journal of Applied Physics*, vol. 111, no. 12, 2012.

[54] E. H. Nicollian and A. Goetzberger, "Lateral ac current flow model for metal-insulator-semiconductor capacitors," *IEEE Transactions on Electron Devices*, vol. 12, no. 3, pp. 108–117, 1965.

[55] S. M. Sze and K. K. NG, *Physics of Semiconductor Devices*, 3rd ed. Wiley, 2006.

[56] L. Vegard, "Die Konstitution der Mischkristalle und die Raumfullung der Atome," *Zeitschrift fur Physik*, vol. 1, no. 1, pp. 17–26, 1921.

[57] I. Vurgaftman, J. R. 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.

[58] V. Ariel-Altschul, E. Finkman, and G. Bahir, "Approximations for Carrier Density in Non-parabolic Semiconductors," *IEEE Transactions on Electron Devices*, vol. 39, no. 6, pp. 1312–1316, 1992.

[59] W. Shockley and W. T. Read, "Statistics of the Recombinations of Holes and Electrons," *Phys. Rev.*, vol. 87, pp. 835–842, 1952.

[60] H. J. Gossmann, B. A. Davidson, G. J. Gualtieri, G. P. Schwartz, A. T. MacRander, S. E. Slusky, M. H. Grabow, and W. A. Sunder, "Critical layer thickness and strain relaxation measurements in GaSb(001)/AlSb structures," *Journal of Applied Physics*, vol. 66, no. 4, pp. 1687–1694, 1989.

[61] W. Lu, I. P. Roh, D. . Geum, S. . Kim, J. D. Song, L. Kong, and J. A. del Alamo, "10-nm fin-width ingasb p-channel self-aligned finfets using antimonide-compatible digital etch," in *IEEE International Electron Devices Meeting*, 2017, pp. 17.7.1–17.7.4.

[62] R. Oxland, X. Li, S. W. Chang, S. W. Wang, T. Vasen, P. Ramvall, R. Contreras-Guerrero, J. Rojas-Ramirez, M. Holland, G. Doornbos, Y. S. Chang, D. S. Macintyre, S. Thoms, R. Droopad, Y. C. Yeo, C. H. Diaz, I. G. Thayne, and M. Passlack, "InAs FinFETs With $H_{fin} = 20$ nm Fabricated Using a Top-Down Etch Process," *IEEE Electron Device Letters*, vol. 37, no. 3, pp. 261–264, 2016.

[63] M. Houssa, E. Chagarov, and A. Kummel, "Surface Defects and Passivation of Ge and III-V Interfaces," *MRS Bulletin*, vol. 34, no. 7, pp. 504–513, 2009.

# Chapter 6

# Conclusion

The practical limitations of scaling transistors in strict accordance with Dennard scaling theory have resulted in increased power dissipation in logic circuits with each progressive technology generation. With chip power consumption now approaching 200 W/cm$^2$, further increases cannot be tolerated and the supply voltage must be reduced in order to facilitate further scaling. Utilising III-V compound semiconductors instead of Si for logic devices could facilitate the reduction of supply voltage while simultaneously improving on-state performance, however, the semiconductor industry will not implement such a disruptive change unless the performance gain is substantial and the associated fabrication costs are economical. A multitude of challenges exist in order to meet these requirements and this thesis has reported on two novel technologies which pertain to realising an all III-V CMOS technology in light of these issues.

In recent years, rapid advances have been made with III-V n-type MOSFETs which, when current is normalised by gated perimeter, now do significantly outperform Si devices at equivalent gate lengths. This is not the case when current is normalised by fin pitch, however, which corresponds to the drive current available for a given semiconductor area. In this circumstance, the high aspect ratio of current non-planar Si technology trumps the intrinsic advantage of III-Vs in terms of transport properties. In order to complete, III-V devices must too exploit the vertical dimension, and in the interest of economic viability, they must do so while utilising Si CMOS compatible process modules, which in turn necessitates a 'top-down' fabrication process. Should this be achieved, the problems faced by III-V n-type devices are the same as those faced by Si: how do you in practice continue to scale devices while maintaining yield, reliability and cost? Chapter 4 addressed all of these points and presented the design, fabrication, measurement and analysis of high aspect ratio, top-down InGaAs junctionless FinFETs with fin top widths scaled to less than 10 nm.

The junctionless architecture is a promising candidate for the ultimately scaled device as it removes the requirement to form extremely abrupt, high doping concentration gradients between

source, drain and channel regions, which is an ever increasing fabrication challenge as gate lengths decrease. To enable its fabrication, a low damage, highly anisotropic ICP-RIE dry etch process was developed. The optimal devices fabricated demonstrated excellent performance with a minimum subthreshold slope, SS, of 73 mV/Dec. and an on-current, $I_{on}$, normalised by gated perimeter, of 80.51 $\mu$A/$\mu$m (measured at 0.5 V from $I_{off}$ = 100 nA/$\mu$m and $V_{\mathrm{d}}$= 0.5 V) which is competitive with state of the art III-V junctionless devices of the same gate length. With $I_{on}$ re-normalised to fin width, the advantage of employing high aspect ratio fins is explicit, and, with a corresponding $I_{on}$ of 371.8 $\mu$A/$\mu$m, the devices fabricated in this work demonstrated the highest current per device channel width of any III-V junctionless device at any gate length.

The etch process developed yielded highly vertical sidewalls with a significantly lower DC bias than anything demonstrated in the literature for etching III-Vs at this feature size. The low damage properties of this process were evident given the excellent performance demonstrated, and it is clear that an etched 'top-down' III-V device can indeed compete with those fabricated using 'bottom-up' approaches. Further, this etch process has reach beyond the junctionless architecture, which is currently a relatively immature technology, and should prove a useful tool in the fabrication of future In-containing non-planar III-V devices, junctionless or otherwise.

The simplified processing and high performance yielded by the junctionless FinFETs fabricated indicate the promise of this architecture for the ultimately scaled device, and further work should be undertaken to scale gate lengths to those demonstrated by state of the art Si devices in order to fully assess its performance and viability. To this end, a process flow was proposed in Section 4.8.2 which may enable the realisation of devices with substantially shorter gate lengths than those demonstrated.

For this initial endeavour into the fabrication of III-V junctionless devices, $In_{0.53}Ga_{0.47}As$ was utilised due to its relative technological maturity in comparison to other III-Vs, however, the knowledge gleaned from the work presented should serve as a useful starting point in the fabrication of equivalent devices which utilise other In-containing compounds, such as InGaAs with differing In composition, InAs, and indeed, In-antimonides which are discussed further below.

Of course, a high performance III-V n-type device needs a complementary equivalent, and in the interest of processing simplicity and economic viability, both p and n-type devices would ideally be fabricated from the same material system, and better yet, the same material. To date, III-V p-type devices have demonstrated poor performance and Ge is currently the most technological mature, high hole mobility material. Chapter 5 addressed the need for a high performance III-V p-type device by investigating the in-situ $H_2$ plasma cleaning of the $In_3Ga_{0.7}Sb$-$Al_2O_3$ interface, which is the first to report doing so. Unlike other III-Vs, InGaSb has excellent transport prop-

erties for both electrons and holes and could therefore facilitate III-V CMOS with a common channel material for both device polarities; provided, of course, that it is possible to from a gate stack with sufficient electrical properties. To this end, a systematic array of experiments was undertaken to assess the impact of $H_2$ plasma cleaning on the $In_{0.3}Ga_{0.7}Sb$-$Al_2O_3$ interface. The optimal process, which included ex-situ HCl surface clean, in-situ $H_2$ plasma and TMA pre-cleaning, and a post gate metal forming gas anneal, was unequivocally demonstrated to yield a fully unpinnned MOS interface with both n and p-type MOSCAPs explicitly demonstrating a genuine minority carrier responses. This result has the important ramification that this gate stack process is indeed applicable to both p and n-type MOSFETs, and therefore, a unified gate stack process could be used in the fabrication of InGaSb CMOS devices. This has the potential to greatly simplify the fabrication, and therefore reduce the cost, of high mobility CMOS manufacture.

# Appendix A

# Junctionless FinFET Fabrication Process

## A.1 Lithography Level 0 (EBL): Markers

- Resist application:
  - Sample cleaning: 50°C acetone[1], 5 min. → IPA, 5 min. → $N_2$ blow dry
  - Dehydration bake: 180°C oven, 10 min.
  - Spin Resist: PMMA 8% (2010), 5k rpm, 60 s (nominal thickness $\sim$ 200 nm)
  - Soft bake: 154°C hotplate, 2min.
  - Spin Resist: PMMA 4% (2041), 5k rpm, 60 s nominal thickness $\sim$ 115 nm)
  - Soft bake: 154°C hotplate, 2min.
- EBL exposure:
  - Global & Cell Makers (CD $\geq$ 50 $\mu$m ):
    - Dose: 440 $\mu$C/cm$^2$
    - Beam current/spot size: 32 nA/24 nm
    - VRU/BSS: 16/20 nm
    - Prox. Correction: Custom PEC file, simulated using PENELOPE

[1] All processes were carried out at RT unless otherwise stated.

- Resist development:
  - Wet development: 23°C MIBK:IPA (2.5:1), 1 min. → IPA, 30 s → $N_2$ blow dry
  - Ash: $O_2$ plasma, 80 W, 2 min.
- Metal lift-off:
  - Metal deposition: Ti/Au (10 nm/100 nm), ebeam evaporation
  - Lift-off: 50°C acetone, 2 hours → IPA, 5 min. → $N_2$ blow dry

## A.2 Lithography Level 1 (EBL): Fin, S/D & contacts

- Resist application:
  - Sample cleaning: 50°C acetone, 5 min. → IPA, 5 min. → DI water, 5 min. → $N_2$ blow dry
  - Resist adhesion promoter: SurPass 3000, 1 min. → DI water, 1 min. → $N_2$ blow dry
  - Spin Resist: HSQ:MIBK (3:1), 5k rpm, 60 s (nominal thickness $\sim$ 75 nm)
  - Soft bake: 90°C hotplate, 2min.
- EBL exposure:
  - Fin ( CD = 10 nm, 15 nm & 20 nm, all patterns = 5 nm wide):
    - Dose: 12,000 $\mu$C/cm$^2$ (10 nm wide fin), 18,700 $\mu$C/cm$^2$ (15 nm wide fin), 30,400 $\mu$C/cm$^2$ (20 nm wide fin)
    - Beam current/spot size: 1 nA/4 nm
    - VRU/BSS: 2/2.5 nm
    - Prox. Correction: No correction
  - S/D (CD = 30 nm)
    - Dose: 800 $\mu$C/cm$^2$
    - Beam current/spot size: 1 nA/4 nm
    - VRU/BSS: 2/2.5 nm
    - Prox. Correction: No correction

  - Contacts (CD = 2$\mu$m)
    - Dose: 615 $\mu$C/cm$^2$
    - Beam current/spot size: 32 nA/24 nm
    - VRU/BSS: 16/20 nm
    - Prox. Correction: Custom PEC file, simulated using PENELOPE [2]

- Resist development:
  - PEB: 95°C hotplate, 2 min.
  - Wet development: 30 s 23°C TMAH (25%) → 30 s DI water → 30 s DI water (fresh beaker) → $N_2$ blow dry

## A.3 Lithography Level 3 (EBL): Bond Pads

- Resist application:
  - Sample cleaning: 50°C acetone, 5 min. → IPA, 5 min. → DI water, 5 min. → $N_2$ blow dry
  - Resist adhesion promoter: SurPass 3000, 1 min. → DI water, 1 min. → $N_2$ blow dry
  - Spin Resist: HSQ:MIBK (3:1), 5k rpm, 60 s (nominal thickness ∼ 75 nm)
  - Soft bake: 90°C hotplate, 2min.
- EBL exposure:
  - Bond Pads ( CD = 20 $\mu$m):
    - Dose: 550 $\mu$C/cm$^2$ (20 nm wide fin)
    - Beam current/spot size: 32 nA/24 nm
    - VRU/BSS: 16/20 nm
    - Prox. Correction: Custom PEC file, simulated using PENELOPE
- Resist development:
  - PEB: 95°C hotplate, 2 min.

[2] PENELOPE is an open-source software available at: http://pypenelope.sourceforge.net/

  - Wet development: 30 s 23°C TMAH (25%) → 30 s DI water → 30 s DI water (fresh beaker) → $N_2$ blow dry

- Etch:
  - Fin etch: refer to Table 4.3 for details.
  - HSQ etch: ICP-RIE, $CF_4$, 50 SCCM, 2000 W ICP Power, 15 W Platen Power, 20 mT, 60°C, 1 min.
  - Digital Etch: 3 cycles, each cycle comprising:
    * Plasma oxidation: RIE, $O_2$, 10 SCCM, 10 W, 50 mT, RT, 2 min.
    * Oxide wet etch: $H_2SO_4$: DI $H_2O$ (1:1), 30 s → IPA, 15 s → $N_2$ blow dry
- High-k deposition: $Al_2O_3$ / $HFO_2$ (nominally 1 nm/3nm) bi-layer deposited via ALD at TNI (Refer to Chapter 4, Ref. 24 for further details)

## A.4 Lithography Level 4 (EBL): Gate contact

- Resist application:
  - Sample cleaning: 50°C acetone, 5 min. → IPA, 5 min. → $N_2$ blow dry
  - Dehydration bake: 180°C oven, 10 min.
  - Spin Resist: PMMA 8% (2010), 5k rpm, 60 s (nominal thickness $\sim$ 200 nm)
  - Soft bake: 154°C hotplate, 2min.
  - Spin Resist: PMMA 4% (2041), 5k rpm, 60 s nominal thickness $\sim$ 115 nm)
  - Soft bake: 154°C hotplate, 2min.
- EBL exposure:
  - Gate & Gate Bond Pad (CD $\geq$ 50 $\mu$m ):
    * Dose: 440 $\mu C/cm^2$
    * Beam current/spot size: 32 nA/24 nm
    * VRU/BSS: 16/20 nm
    * Prox. Correction: Custom PEC file, simulated using PENELOPE

- Resist development:
  - Wet development: 23°C MIBK:IPA (2.5:1), 1 min. → IPA, 30 s → $N_2$ blow dry
  - Ash: $O_2$ plasma, 80 W, 2 min.
- Metal lift-off:
  - Metal deposition: Ti/Au (10 nm/100 nm), ebeam evaporation
  - Lift-off: 50°C acetone, 2 hours → IPA, 5 min. → $N_2$ blow dry

## A.5 Lithography Level 5 (EBL): S/D contact

- Resist application:
  - Sample cleaning: 50°C acetone, 5 min. → IPA, 5 min. → $N_2$ blow dry
  - Dehydration bake: 180°C oven, 10 min.
  - Spin Resist: PMMA 8% (2010), 5k rpm, 60 s (nominal thickness $\sim$ 200 nm)
  - Soft bake: 154°C hotplate, 2min.
  - Spin Resist: PMMA 4% (2041), 5k rpm, 60 s nominal thickness $\sim$ 115 nm)
  - Soft bake: 154°C hotplate, 2min.
- EBL exposure:
  - Contacts/Bond Pads (CD = 100 $\mu$m ):
    - Dose: 440 $\mu$C/cm$^2$
    - Beam current/spot size: 32 nA/24 nm
    - VRU/BSS: 16/20 nm
    - Prox. Correction: Custom PEC file, simulated using PENELOPE
- Resist development:
  - Wet development: 23°C MIBK:IPA (2.5:1), 1 min. → IPA, 30 s → $N_2$ blow dry
  - Ash: $O_2$ plasma, 80 W, 2 min.
- Metal lift-off:

- : Oxide etch: 30 s in-situ Ar exposure
- Metal deposition: Ni/Ti/Pd/Au (4/15/15100 nm), ebeam evaporation
- Lift-off: 50°C acetone, 2 hours $\rightarrow$ IPA, 5 min. $\rightarrow$ $N_2$ blow dry

# Appendix B

# MOSCAP Fabrication Process

The following details the MOSCAP fabrication process for both the Power and Time series samples, subsequent to $Al_2O_3$ deposition. Full details of the $Al_2O_3$ ALD process itself are given in Chapter 5, Sections 5.4 and 5.5.

## B.1 Gate metal deposition

- Metal deposition: Pt/Au (20 nm/200 nm), ebeam evaporation through a shadow mask

## B.2 Lithography Level 1, Power Series Only (photolithography):

- Resist application:
  - Sample cleaning: 50°C acetone[1], 5 min. → IPA, 5 min. → $N_2$ blow dry
  - Spin Resist: Microposit S1818, 4k rpm, 30 s (nominal thickness $\sim 2\ \mu$m)
  - Soft bake: 90°C oven, 30 min.
- Photolithography exposure: hard contact, Al gap 45 $\mu$m, 6 s
- Resist development:
  - Wet development: Microdev: DI water (1:1), 45 s. → DI water, 30 s → $N_2$ blow dry

[1] All processes were carried out at RT unless otherwise stated.

  - Ash: $O_2$ plasma, 80 W, 2 min.

- $Al_2O_3$ etch: RIE, $SiCl_4$, 25 SCCM, 100 W, 8 mT, RT, 8 min.
- Metal lift-off:
  - Metal deposition: Ti/Pt/Au (30 nm/50 nm/100 nm), ebeam evaporation
  - Lift-off: 80°C SVC14, 2 hours → IPA, 5 min. → $N_2$ blow dry

## B.3 Backside Metal Deposition, Time Series Only:

- Metal deposition: Ti/Pt/Au (30 nm/50 nm/100 nm), ebeam evaporation to the back of the sample

# Appendix C

# InGaSb-$Al_2O_3$ MOS Equivalent Circuit Model Parameters

| | Parameter | p-InGaSb | n-InGaSb |
|---|---|---|---|
| Pt | $\phi_m$ (eV) | 5.2 | |
| $Al_2O_3$ | $m_h^*$ ($m_0$) | 0.28 | |
| | $m_e^*$ ($m_0$) | 0.28* | |
| | $E_g$ (eV) | 6.65 | |
| | $\phi_{Al_2O_3}$ (eV) | 2.58 | |
| | $\kappa_h$ ($nm^{-1}$) | 8.23 | |
| | $\kappa_e$ ($nm^{-1}$) | 3.46 | |
| | $T_{ox}$ (nm) | 7.4 | |
| $In_xGa_{1-x}Sb$ | x | 0.275 | |
| | $E_g$ (meV) | 490 | |
| | $m_{hh}^*$ ($m_0$) | 0.408 | |
| | $m_{e,\Gamma}^*$ ($m_0$) | 0.025 | |
| | $N_{A/D}$ ($10^{17}$ $cm^{-3}$) | 2.5 | 1.7 |
| | $\phi_s$ (eV) | 4.2 | |
| | $G_{bulk}$ ($S/cm^2$) | 0.75 | 1.2 |
| Extrinsic | $R_s$ ($m\Omega$ $cm^2$) | 3.3 | 21 |
| | $L_s$ ($nH/cm^2$) | 1† | 2† |

* $0.28m_0$ also used for $m_h^*$ for $Al_2O_3$ due to lack of available data.

† $L_s$ was found to increase with increasing gate diameter. p and n-type data was measured on capacitors with gate diameters of 100 $\mu$m and 250 $\mu$m respectively.

Table C.1: InGaSb-$Al_2O_3$ MOS equivalent circuit model parameters.