III-V Vertical Nanowire Transistor for Ultra-Low Power Applications

by

Xin Zhao

B.S. Physics
Peking University 2010

S. M. Materials Science and Engineering
Massachusetts Institute of Technology, 2012

Submitted to the Department of Materials Science and Engineering
in Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy in Materials Science and Engineering
at the
Massachusetts Institute of Technology

June 2017

©2017 Massachusetts Institute of Technology. All Rights Reserved.

Signature of Author ...............................................................................................................

Department of Materials Science and Engineering
February 16, 2017

Certified by ............................................................................................................................

Jesús A. del Alamo
Professor, Electrical Engineering and Computer Science
Thesis Supervisor

Certified by ............................................................................................................................

Silvija Gradečak
Associate Professor, Materials Science and Engineering
Thesis Reader

Accepted by ...........................................................................................................................

Donald R. Sadoway
Chairman, Departmental Committee on Graduate Studies
III-V Vertical Nanowire Transistor for Ultra-Low Power Applications

by

Xin Zhao

Submitted to the Department of Materials Science and Engineering on February 16, 2017 in Partial Fulfillment of the Requirements for the Degree of Doctor of Philosophy in Materials Science and Engineering

Abstract

Combining the superior carrier transport properties and flexible band structure engineering from III-V materials and ultimate scalability of vertical nanowire (VNW) device architecture, III-V VNW transistors are promising to extend Moore’s law further than any other device technology. In this thesis, III-V VNW transistor technology has been pioneered via a top down approach for logic applications in future ultra-low power systems.

Process flow and critical modules for sub-10 nm VNW transistors are developed from scratch. A novel dry etch technique based on BCl3/SiCl4/Ar chemistry for fabricating sub-20 nm III-V nanostructures with smooth, vertical sidewall and high aspect ratio (> 10) is developed. Digital etch (DE) is shown to mitigate the dry etch damage and reduce NW diameter below 10 nm in a controllable fashion while preserving the sidewall roughness and NW shape.

Top-down InGaAs VNW MOSFET is demonstrated for the first time. Record Ion of 224 μA/μm is obtained at Ioff = 100 nA/μm with Vdd = 0.5 V in third generation devices. With novel solvent-based, switching characteristics are observed in devices with diameter as small as 14 nm. The impact of the intrinsic source/drain asymmetry on the device electrical characteristics is studied in detail, highlighting the importance of uniform NW diameter.

The first experimental demonstration of III-V VNW TFETs with an InGaAs/InAs heterojunction fabricated by a top-down approach is introduced. Second generation TFETs demonstrate sub-thermal subthreshold characteristics over two orders of magnitude of current and a record high I60 in any experimental TFETs for Vds < 1 V at the time of device fabrication. The comparison of two generations of TFETs confirms oxide/semiconductor interface trap-assisted tunneling as the source of significant temperature dependence in the first device generation. Detailed analysis on the conductance-voltage characteristics on both generations of devices reveal a 100-120 mV/dec steepness of Urbach tails in the VNW TFETs.
Acknowledgments

First and foremost, I would like to express my deepest gratitude to my thesis advisor, Prof. Jesús del Alamo for his support and guidance over the years. I admire his dedication to excellence, enthusiasm, persistence, and insights. He is a role model for researchers, mentors and teachers.

I also want to thank Prof. Dimitri Antoniadis and Prof. Eli Yablonovitch for their sharp, witty and insightful comments and suggestions. They are the visionary engineer and scientist who have shaped the semiconductor industry. My sincere thanks to Prof. Fitzgerald for his unique view of the world that I find very inspiring and Prof. Silvija Gradečak for her guidance and patience as a thesis committee member.

This thesis would have not been possible without the support from my colleagues and friends. I would really miss the days that I fought side by side in the cleanroom with Alon Vardi, Tao Yu, Jianqiang Lin, Wenjie Lu, Luke Guo, Winston Chern, Yuhao Zhang, Lili Yu, Xu Zhang and Xiaowei Cai. Many thanks to del Alamo group members: Alex Guo, Shireen Warnock, Yufei Wu, Donghyun Jin, Ling Xia and Jorg Scholvin. I also appreciate the help from MTL and E3S colleagues, Chris Heidelberger, Daniel Peidra, Min Sun, Ahmad Zubair, James Teherani, Ryan Iutzi, Sameer Joglekar, Bin Lu, Han Wang, Feng Gao, Sapan Agarwal, Tianyao Xiao Xiabing Lou, Lan Wei, Wentao Wang, Laura Popa, Radhika Marathe, Omar Saadat, Mohamed Azize, Ujwal Radhakrishna, and Dong Seup Lee.

Finally, I would like to thank the staffs of MTL and SEBL, who have been devoting themselves to maintaining a world-class cleanroom, despite the numerous and often conflicting process needs. I would like to acknowledge Mark Mondol, Eric Lim, Bob Bicchieri, Donal Jamieson, Dennis Ward, Vicky Diadiuk, Kurt Broderick, Paul Tierney, Dave Terry, Paudely Zamora, Paul McGrath, Gary Riggott, Bernard Alamariu, Debroah Hodges-Pabon, Kristofor Payer, Mike Hobbs, Ryan O'Keefe, Tim Turner, and Bill Maloney.

Most of all, I cannot thank my parents enough for being always by on my side, in happiness or sorrow, in ecstasy or despair. I hope I have made you proud.

*This research was sponsored by the Center for Energy Efficient Electronics Sciences (NSF Award 0939514), Lam Research and SRC.*
Table of Contents

LIST OF FIGURES .................................................................................................................9

LIST OF TABLES ..........................................................................................................................18

CHAPTER 1 INTRODUCTION ........................................................................................................19

1.1 The Need for Ultra-Low Power Devices ...........................................................................19
1.1.1 The trade-off of energy and speed in electrical switches ...........................................19
1.1.2 Moore’s law of scaling .................................................................................................20
1.1.3 The need for ultra-low power devices .......................................................................23
1.1.4 The Limit of Si MOSFETs Technology .....................................................................25
1.2 III-V Vertical Nanowire Transistor Technology ...............................................................27
1.2.1 Motivation for III-V’s .................................................................................................29
1.2.2 The necessity of nanowire geometry .........................................................................31
1.2.3 The advantage of a vertical channel .........................................................................32
1.2.4 Top-down approach ...................................................................................................34
1.2.5 III-V vertical nanowire transistor technology ...........................................................35
1.3 Thesis Overview ...............................................................................................................36

CHAPTER 2 PROCESS DEVELOPMENT ..................................................................................38

2.1 Process Overview ..............................................................................................................39
2.2 Nanowire Formation ........................................................................................................45
2.2.1 Electron beam lithography .........................................................................................45
2.2.2 Reactive ion etching ..................................................................................................46
2.2.3 Digital etch ................................................................................................................53
2.3 Gate Stack .........................................................................................................................60
2.4 Backend Process ...............................................................................................................64
2.4.1 Planarization and etch back .......................................................................................64
2.4.2 Contact .......................................................................................................................69
2.5 Chapter Summary .............................................................................................................70

CHAPTER 3 INGAAS VERTICAL NANOWIRE MOSFETS ..................................................72

3.1 Introduction .......................................................................................................................72
3.2 Device Electrical Characteristics ....................................................................................74
3.2.1 First generation ...........................................................................................................74
3.2.2 Second generation .......................................................................................................80
3.2.3 Third generation ................................................................. 91
3.2.4 Benchmark .............................................................................. 94
3.3 Electrostatics .................................................................................. 95
3.4 Source/Drain Asymmetry ............................................................. 98
3.5 Roadblock: Top Contact ............................................................... 103
3.6 Chapter Summary ........................................................................... 105

CHAPTER 4 INGAAS/INAS HETEROJUNCTION VERTICAL
NANOWIRE TFETS ........................................................................... 107

4.1 Introduction ..................................................................................... 107
4.2 First Generation ............................................................................. 114
4.3 Temperature Dependence .............................................................. 122
  4.3.1 Output and subthreshold characteristics .................................. 122
  4.3.2 Conductance-voltage characteristics ..................................... 128
4.4 Second Generation .......................................................................... 135
4.5 Trap Assisted Tunneling and Urbach Tail Steepness ...................... 138
4.6 Benchmark ..................................................................................... 141
4.7 Chapter Summary ........................................................................... 143

CHAPTER 5 SUMMARY AND FUTURE WORK .............................. 145

5.1 Summary ....................................................................................... 145
5.2 Future Work ................................................................................ 148

APPENDIX: .......................................................................................... 151

A.1 Sample Process Flow ................................................................. 151
A.2 Recipes ......................................................................................... 153
A.3 Reactive Ion Etching ................................................................. 154
A.4 Publications ................................................................................. 155

REFERENCES .................................................................................... 156
List of Figures

Figure 1-1. (a) Schematic of the typical subthreshold characteristics (Id vs. Vgs) in a MOSFET. (b) Schematic of the energy-delay trade-off in a certain MOSFET technology due to the shape of the subthreshold characteristics. Given a target performance shown in the vertical line, the intersection highlighted in purple dot represents the lowest every consumption possible for an electronics system based on the chosen transistor technology and operating voltage.................20

Figure 1-2. Transistor count in integrated circuit chips as a function of time, reproduced from [2]. Progress shows exponential growth for decades.................................................................21

Figure 1-3. Progress in energy-delay trade-off across generations of technology and applications. The family of black solid lines represents the fundamental limit of different device technologies that have enabled mainframe to mobile computing. Moore’s law of scaling keeps pushing the limit of transistor technology as indicated by the green arrow, achieving higher performance, lower energy consumption at smaller footprint simultaneously in each successive technology node. The relentless improvement has led to numerous applications represented by the images, including personal computer, internet and smart phones. To materialize the vision of internet of things and deliver applications such as virtual reality and implantables highlighted in dotted circles, significant breakthroughs in device technologies are required in order to push the limit in energy-delay trade-off to the red dashed line. .........................................................22

Figure 1-4. (a) Power supply voltage (Vdd) and threshold voltage (Vt) vs. technology generation, reproduced from [12]. (b) Active and subthreshold power consumption in integrated circuits in various technology nodes, reproduced from [14]. The subthreshold leakage power grows exponentially as a result of Vt scaling. .................................................................25

Figure 1-5. Sketch of subthreshold characteristics of Si MOSFET (black solid line), III-V/Ge MOSFET (blue solid line) and a prototypical steep-slope device (green solid line) with the same off current. MOSFETs with novel channel materials such as III-Vs and Ge are expected to deliver higher performance at reduced supply voltage because of the superior carrier transport properties. Steep-slope devices including TFETs with sub-thermal subthreshold swing are promising device technologies to further reduce the power consumption. .............................28

Figure 1-6. Electron injection velocity in III-V HEMTs, reproduced from [11]. Electron injection velocity is shown for InGaAs and InAs HEMTs with different channel compositions and for silicon MOSFETs as a function of gate length.........................................................30
Figure 1-7. (a) MOSFET structural evolution, reproduced from [24]: 3D schematic (top) and cross-sectional schematic (bottom) of MOSFET structures with increasing electrostatic gate control of the channel from left to right. (b) DIBL as a function of effective gate length shows improvements in SCE as the gate number is increased from ETSOI to GAA, reproduced from [23].

Figure 1-8. (a) Schematic of vertical nanowire transistor architecture showing the contact length (L_c), spacer thickness (L_{spacer}) and gate length L_g. (b) Simulated performance of ring oscillators (RO) with FinFETs (FF), vertical nanowire FETs (VF) and two flavors of lateral nanowire FETs (L2 and L3) under 5 nm technology node design rules, reproduced from [26]. At each data point, device technology is optimized for minimum RO delay at a given supply voltage and fixed off current.

Figure 1-9. (a) InAs nanowire array growth from Au seed particles on InAs substrate by the VLS technique, reproduced from [30]. (b) Cross-sectional TEM image of InGaAs/InAlAs superlattice with 1 nm period and an atomically sharp interface. Material growth and imaging was done by Ryan Iutzi at MIT.

Figure 2-1. Schematics of the cross-section of MOSFET (left) and TFET (right).

Figure 2-2. Exemplary heterostructures of MOSFET and TFET used in this thesis.

Figure 2-3. Typical process flow for III-V VNW transistor fabrication.

Figure 2-4. Tilted SEM image of HSQ mask with 30 nm diameter pattern.

Figure 2-5. Tilted SEM image of GaAs (left) and GaSb (right) NWs fabricated with BCl3/N2 recipe at 40 °C.

Figure 2-6. Tilted SEM image of GaAs (left) and InGaAs (right) NWs fabricated with BCl3/N2 recipe at 40 °C.

Figure 2-7. 15 nm diameter InGaAs NW defined by optimized RIE technique with an aspect ratio greater than 15.

Figure 2-8. InGaAs pillars etched under different RIE conditions: (a) optimum parameters (see text); (b) through (e) one parameter different from (a). (b) Chamber pressure = 0.8 Pa (vs. 0.2 Pa), (c) RF platen power = 50 W (vs. 160 W, with substrate bias voltage of 120 V vs. 280 V), (d) SiCl4 flow rate = 0.25 SCCM (vs. 0.55 SCCM), (e) BCl3/Ar flow rate = 2/12 SCCM (vs. 7/7 SCCM).
Figure 2-9. Impact of substrate temperature on etching rate, profile, and surface roughness of InGaAs NWs: (a) 100 °C, (b) 175 °C, (c) 250 °C. .................................................................51

Figure 2-10. (a) 25 nm fins containing InP and InAlAs, etched with optimized BCl₃/Ar/SiCl₄ recipe. (b) Starting heterostructure used for etching.................................................................52

Figure 2-11. 10 nm fins etched with optimized BCl₃/Ar/SiCl₄ recipe. The starting heterostructure contains 20 nm of In₀.₂₈Ga₀.₆₂Sb on 200 nm Al₀.₆₅Ga₀.₃₅Sb buffer. ...............53

Figure 2-12. Etch rate per cycle of digital etch as a function of oxidation time under O₂ plasma, reproduced from [57] and fitting with Lukeš’ model [58].................................................................55

Figure 2-13. (a) D= 28 nm InGaAs NW fabricated by RIE. (b) Same NW as in (a) after 5 subsequent cycles of digital etch..................................................................................................56

Figure 2-14. Impact of digital etch on subthreshold and transconductance (inset) characteristics of D= 30 nm gate-all-around InGaAs NW-MOSFETs. 30 nm is the final device diameter. ......57

Figure 2-15. Fin structure in Figure 2-10 after 8 subsequent cycles of digital etch, demonstrating the uniformity of DE technique for a heterostructure containing In, Ga, Sb, Al and P..............58

Figure 2-16. Arsenide VNW array after 7 DE cycles in (a) 10% HCl in DI water, and (b) 10% HCl in IPA. The insets show close ups of the etched structures. The nanowire yield in (a) is 0% while the yield in (b) is 97%.................................................................59

Figure 2-17. InGaAs VNW with 5 nm diameter, 230 nm height obtained after 10 cycles of digital etch in H₂SO₄:methanol...................................................................................................60

Figure 2-18. (a) D = 33 nm InGaAs NW. (b) Same NW after 80 cycles of ALD Al₂O₃. .........61

Figure 2-19. D = 40 nm InGaAs NW covered by gate stack comprising 3 nm Al₂O₃ and 40 nm of sputtered W (measured on the planar surface). .................................................................62

Figure 2-20. CV characteristics of InGaAs MOSCAP with 3 nm Al₂O₃ dielectric: (left) before and (right) after FGA. .................................................................63

Figure 2-21. CV of InGaAs MOSCap with 0.5/3 nm Al₂O₃/LaAlO₃ dielectric after FGA with W (left) and WN (right) gate metal.................................................................64

Figure 2-22. Schematics illustrating the planarization and etch back process. .........................65
Figure 2-23. Tilted SEM images of the planarization and etch back process, each corresponding to one cartoon shown in Figure 2-12 as follows: (a) – (c), (b) – (d), (c) – (f), (d) – (h)..............67

Figure 2-24. SEM image of a D=30 nm NW MOSFET after 1st planarization and etch back (SOG and dielectric removed in BOE). The intrinsic region and the bottom n+ region of the nanowire are wrapped around by W. The top n+ region is exposed for ohmic contact formation. The roughness on the ground plane is caused by the erosion of W in BOE.........................68

Figure 2-25. SEM images of the via bottom after FOX dry etch without (left) and with (right) H2 recipe. Residue FOX is clearly seen on the right.................................................................69

Figure 3-1. Schematics (left column) and SEM images (right column) of III-V VNWs grown on Si substrate by VLS, SAE and template-assisted techniques [68].................................73

Figure 3-2. SEM images of III-V VNW MOSFETs fabricated by bottom-up approaches [69-71]. ........................................................................................................................................74

Figure 3-3. Starting heterostructure and first generation device design parameters..............75

Figure 3-4. D = 30 nm InGaAs NW after 10 cycles of digital etch used for first generation device demonstration. ..............................................................................................................75

Figure 3-5. Output characteristics of a D = 30 nm InGaAs single NW MOSFET.....................76

Figure 3-6. Subthreshold characteristics (left), transfer and gm characteristics at Vds = 0.5 V (right) of the device shown in Figure 3-5. .................................................................................................76

Figure 3-7. Output characteristics of a D = 50 nm InGaAs single NW MOSFET.....................77

Figure 3-8. Subthreshold characteristics (left), transfer and gm characteristics at Vds = 0.5 V (right) of the device shown in Figure 3-7. .................................................................................................77

Figure 3-9. Impact of digital etch on the subthreshold characteristics of D = 40 nm devices. This is the final device diameter ........................................................................................................78

Figure 3-10. Impact of digital etch on gate leakage current of D = 40 nm single NW MOSFETs. ...........................................................................................................................................79

Figure 3-11. Output characteristics of D = 40 nm InGaAs single NW MOSFETs fabricated without (left) and with (right) digital etch. .................................................................................................79
Figure 3-12. Key device figures of merit vs. final nanowire diameter for devices fabricated with and without digital etch (over 10 devices for each nanowire diameter are measured): (a) Linear ($V_{ds} = 0.05$ V) and saturated ($V_{ds} = 0.5$ V) subthreshold swing, (b) DIBL, (c) Normalized $g_m$ ($V_{ds} = 0.5$ V), and (d) Normalized $R_{on}$.

Figure 3-13. Starting heterostructure of 2nd generation devices and design parameters.

Figure 3-14. Output characteristics of a $D = 40$ nm InGaAs single NW MOSFET from G2.

Figure 3-15. Subthreshold and $g_m$ (inset, $V_{ds} = 0.5$ V) characteristics of the device shown in Figure 3-14.

Figure 3-16. Subthreshold (left) and output (right) characteristics of a $D = 20$ nm InGaAs single NW MOSFET from G2.

Figure 3-17. Comparison of devices with 30 and 40 nm diameters in terms of the statistics of key FOMs including $S_{lin}$, $S_{sat}$, DIBL, $g_{m, pk}$ and $R_{on}$.

Figure 3-18. Output characteristics of a $D = 40$ nm InGaAs array NW MOSFET.

Figure 3-19. Subthreshold characteristics (left) and $g_m$ characteristics at $V_{ds} = 0.5$ V (right) of the device shown in Figure 3-18.

Figure 3-20. Statistical comparison of single and array nanowire devices with 40 nm diameter in key FOMs including $S_{lin}$, $S_{sat}$, and $g_{m, pk}$.

Figure 3-21. 600 s sampling measurements in drain current as a function of time with 50 ms resolution, in both SNW (left) and ANW (right) devices.

Figure 3-22. Normalized power spectral density as a function of frequency for ANW and SNW devices. The red dotted line serves as a reference with 1/f dependence on the frequency.

Figure 3-23. Impact of RTA on subthreshold (left) and transconductance (right) characteristics of an exemplary $D = 40$ nm device.

Figure 3-24. Output characteristics of an exemplary $D = 40$ nm device before (black lines) and after (red lines) RTA.

Figure 3-25. Subthreshold, $g_m$ and output characteristics of one of the best devices in G2 ($D = 40$ nm) after RTA.
Figure 3-26. Starting heterostructure of 3\textsuperscript{rd} generation devices and the design parameters.

Figure 3-27. Subthreshold, $g_m$ and output characteristics of one of the best devices in G3 ($D = 40\text{ nm}$) after RTA.

Figure 3-28. (a) Subthreshold characteristics of a $D = 40\text{ nm}$ with $S_{\text{linear}} = 70\text{ mV/dec}$ (b) Subthreshold characteristics of a $D = 14\text{ nm}$ showing switching behavior at $V_{ds} = 0.5\text{ V}$.

Figure 3-29. Benchmark of $g_{m, pk}$ vs. minimum $S$ at $V_{ds} = 0.5\text{ V}$ for recently published InGaAs and InAs VNW MOSFETs. A trade-off is evident in terms of electrostatics and transport.

Figure 3-30. (a) $S_{\text{linear}}$ plotted versus $L_g / \lambda$ in recently published III-V VNW MOSFETs (symbols). The red line is the simulation [81]. The inset shows the degradation of gate efficiency $\eta$ due to the capacitance from channel charge, depletion region and interface traps ($C_n$, $C_d$ and $C_i$). (b) DIBL as a function of $L_g / \lambda$. The symbols are consistent with (a). The red line is the simulation from [81].

Figure 3-31. Subthreshold characteristics of an exemplary device in G2 measured with bottom electrode at the source (BES, black) vs. top electrode as the source (TES, red) (inset shows transconductance characteristics).

Figure 3-32. Output characteristics measured with BES vs. TES of the device shown in Figure 3-31.

Figure 3-33. The impact of source/drain asymmetry on the (a) ON and (b) subthreshold characteristics of G1 and G2 devices.

Figure 3-34. Improved NW profile in G2 (left) compared to G1 (right).

Figure 3-35. $g_{m, pk}$ plotted as a function of $1/R_{on}$ for G1 (black) and G2 (blue) devices.

Figure 4-1. (a). Schematic cross-section of n-type MOSFET and TFET (b). Schematic band diagram of n-type MOSFET and TFET. The solid red lines refer to the on state while the dashed blue lines represent the off state. The grey arrow highlights the current path when the device is turned on. In TFETs, the energy window for tunneling is $\Delta \Phi$.

Figure 4-2. Comparison of simulated n-type (right) and p-type (left) TFET transfer characteristics, reproduced from [21]. The numbers on the curves label the applied $V_{ds}$. The curves are shifted so that the gate voltage where the steepest sub-threshold slope occurs is at the
origin. The majority of the results predict steep SS and high on-current for a variety of materials and device architectures.

Figure 4-3. Comparison of experimental n-type (right) and p-type (left) TFET transfer characteristics, reproduced from [21]. The numbers on the curves label the applied $V_{ds}$. The curves are shifted so that the gate voltage where the steepest sub-threshold slope occurs is at the origin. The majority of the results show very poor subthreshold characteristics and low on-currents.

Figure 4-4. Schematic cross-section of n-type transverse (left) and inline (right) TFET. The arrows highlight electron flow.

Figure 4-5. Simulated energy band diagram along the NW for ON state (right) and OFF state (left). The self-consistent Poisson-Schrodinger solver implemented in Nextnano3 is used. The simulation was done with a 2D double-gate geometry to approximate the 3D NW channel, in order to speed up the simulation. The current model or interface states model are not included.

Figure 4-6. SEM image of a 50 nm SOG layer covering the p$^+$ source on which the gate stack was deposited. The purpose of this SOG layer was to reduce gate leakage and enable further EOT scaling.

Figure 4-7. Schematic of InGaAs/InAs heterojunction VNW TFET cross-section and heterostructure.

Figure 4-8. Subthreshold characteristics (left), transfer and $g_m$ characteristics at $V_{ds} = 0.3$ V (right) of one of the best performing devices.

Figure 4-9. Output characteristics in linear scale (left) and in semilog scale for positive and negative $V_{ds}$ (right) of the device shown in Figure 4-8.

Figure 4-10. Subthreshold characteristics of three single-NW TFETs at $V_{ds} = 0.3$ V illustrating device-to-device variability.

Figure 4-11. Output characteristics of the three devices shown in Figure 4-10 for positive and negative $V_{ds}$. Clear NDR is observed in all 3 devices.

Figure 4-12. Minimum subthreshold swing at $V_{ds} = 0.3$ V ($S_{sat}$), DIBL and $g_{mpk}$ at $V_{ds} = 0.3$ V for several devices plotted against $V_t$ at $V_{ds} = 0.3$ V.
Figure 4-13. DIBL and $g_{m,pk}$ at $V_{ds} = 0.3$ V plotted against minimum subthreshold swing at $V_{ds} = 0.3$ V ($S_{sat}$)............................................................................................................................................... 121

Figure 4-14. Benchmark of $g_{m,pk}$ vs. minimum $S$ at $V_{ds} = 0.5$ V for first generation InGaAs VNW MOSFETs and InGaAs/InAs VNW TFETs.......................................................................................................................... 121

Figure 4-15. Output characteristics in linear scale of a representative device at different temperatures.............................................................................................................................................. 123

Figure 4-16. Output characteristics in Figure 4-15 plotted in semilog scale................................. 123

Figure 4-17. $I_{d}-V_{ds}$ characteristics at $V_{gs} = 0.5$ V in semilog scale at different temperatures from 77 to 240 K. Negative/positive $V_{ds}$ region of a TFET corresponds to forward/reverse bias of a tunnel diode. Three regions, thermionic emission, Esaki and BTBT are observed.......................... 124

Figure 4-18. Drain current at $V_{ds} = 0.05$ V and $V_{gs} = 0.5$ V plotted as a function of the 1.5 power of the InAs bandgap at measured temperatures from 77 K to 240 K. ............................................................................ 124

Figure 4-19. $I_{d}-V_{ds}$ characteristics at $V_{gs} = -0.1$ V in semilog scale at different temperatures from 77 to 240 K. Compared to $V_{gs} = 0.5$ V characteristics in Figure 4-17, $V_{gs} = -0.1$ V also shows the thermionic emission and Esaki region. The critical difference is the presence of a thermally activated region ($V_{ds} > -0.2$ V). .................................................................................................................. 125

Figure 4-20. Subthreshold characteristics at $V_{ds}=0.05$ V and $T = 77$ to 300 K...................... 126

Figure 4-21. (a) Subthreshold swing at $V_{ds}=0.05$ V as a function of drain current, extracted at different temperatures. (b) Temperature dependence of average $S$ at $V_{ds}=0.05$ V. Red dashed line is the thermal limit of the subthreshold swing in a MOSFET. ................................................................. 127

Figure 4-22. Arrhenius plot of the subthreshold current at various $V_{gs}$ values versus inverse thermal energy. The slope of the linear relationship is the thermal barrier height. (b) Thermal barrier height extracted from the slopes in (a) as a function of $V_{gs}$................................................... 127

Figure 4-23. Temperature dependence in the subthreshold characteristics prevail in recent published III-V TFETs [96, 99, 118, 121, 127-130]....................................................................................................................................... 128

Figure 4-24. (a) Subthreshold characteristics at $V_{ds} = 0.05$ V of a representative device at various temperatures from 77 to 423 K. (b) – (d). Conductance voltage characteristics of the same device with various $V_{gd}$ biases at 77, 290 and 423 K. ......................................................................................... 130
Figure 4-25. Minimal $S_E$, $S_B$, extracted from NDR and backward diode regimes from GV characteristics and $S_S$ measured from subthreshold characteristics plotted as a function of temperature. The green dashed line represents the thermal limit. .............................................................. 132

Figure 4-26. (a) Output characteristics at 290 K in semilog scale. (b) Corresponding conductance versus $V_{ds}$ in semilog scale. A minimum slope of 160 mV/dec can be extracted from the NDR region. ................................................................................................................ 133

Figure 4-27. $I_d - V_{gs}$ characteristics at $V_{ds} = -0.25$ V and $T = 298$ K. The minimum slope is around 500 mV/dec................................................................................................................... 133

Figure 4-28. Distribution of minimum subthreshold swing ($S_S$) at $V_{ds} = 0.3$ V ($S_{sat}$), minimum $S_E$ at $V_{gd} = 0$ V versus $V_t$ at $V_{ds} = 0.3$ V. $S_E$ has much less variability compared to $S_S$. ............ 134

Figure 4-29. Room temperature output characteristics of an exemplary VNW TFET in linear scale (left) and in semilog scale (right). ......................................................................................................................... 136

Figure 4-30. Subthreshold and $g_m$ characteristics (top) and subthreshold swing vs. $I_d$ characteristics (bottom) of InGaAs VNW TFET. ......................................................................................................................... 137

Figure 4-31. Temperature dependence of subthreshold characteristics at $V_{ds} = 0.05$ V (top) and minimum $S$ at $V_{ds} = 0.05$ V as a function of temperature (bottom) for 2 generations of TFETs and G2 MOSFETs............................................................ 138

Figure 4-32. Schematic of band structure of the VNW TFETs illustrating the trap-assisted tunneling process. An electron can reach the conduction band from the valence band via a combination of tunneling and thermal emission................................................................. 140

Figure 4-33. $S$ vs $I_d$ in experimental TFETs with sub-thermal $S$ at RT and $V_{ds} < 1$ V (Adapted from [21])......................................................................................................................... 142

Figure 4-34. $I_{on}$ vs. $I_{off}$ at $V_{dd} = 0.3$ V ($V_{ds}=0.3$ V, $\Delta V_{gs} = 0.3$ V with exceptions marked next to data points) for recently published vertical NW TFETs containing III-V materials. All devices but the present ones are fabricated through bottom-up techniques................................................................. 143
List of Tables

Table 2-1. Yield of InGaAs VNWs after 7 cycles of DE ............................................................. 59
Chapter 1 Introduction

1.1 The Need for Ultra-Low Power Devices

1.1.1 The trade-off of energy and speed in electrical switches

The workhorse of the microelectronics revolution, silicon metal-oxide-semiconductor field-effect transistor (MOSFET), behaves as an electrical switch when it comes to logic operations. The Si MOSFET is essentially a three terminal device, where the current flowing from the source to drain terminal is controlled electrostatically by the voltage of a third gate terminal. The fundamental characteristics of a logic switch can be appreciated in Figure 1-1 (a) that shows the subthreshold characteristics ($I_d$ vs. $V_{gs}$ in semilog scale) of a typical n-type Si MOSFET (Si NMOS). Given a particular technology, this curve is largely fixed but the threshold voltage ($V_t$) is free to be tuned to some degree. The speed or performance on a device level is set by the current density level as the delay time is inversely proportional to the drive current density:

$$ t_{delay} = \frac{V_{dd}C}{I_{on}} \quad (1-1) $$

$t_{delay}$ is the delay time, inversely proportional to the transistor performance, $V_{dd}$ is supply voltage, $C$ is the total capacitance, and $I_{on}$ is the drive current. However, the energy consumption per switch action increases with drive current. As a result, the energy-delay curve of a particular Si MOSFET technology typically exhibits a trade-off, as shown schematically in Figure 1-1 (b). Each point in this curve is obtained by minimizing the device energy consumption for a given performance target, at a certain $V_{dd}$ and off state current $I_{off}$. $V_{dd}$ increases along the this curve.
This line also represents the ultimate limit of circuits or systems built based on the chosen technology: given a target performance shown in the vertical line, the lowest energy consumption and hence the lowest power required is ultimately limited by the intersect point.

![Figure 1-1. (a) Schematic of the typical subthreshold characteristics ($I_d$ vs. $V_{gs}$) in a MOSFET. (b) Schematic of the energy-delay trade-off in a certain MOSFET technology due to the shape of the subthreshold characteristics. Given a target performance shown in the vertical line, the intersection highlighted in purple dot represents the lowest energy consumption possible for an electronics system based on the chosen transistor technology and operating voltage.](image)

### 1.1.2 Moore’s law of scaling

Ever since Gordon Moore’s observation in 1965 about the scaling vs. time of transistor sizes became known as the Moore’s law [1], it has been regarded as the golden rule for the development of the semiconductor industry. The Si MOSFET has been following the geometrical scaling law and resulted in exponential increase in the number of transistors for more than 5 decades, as shown in Figure 1-2 that graphs the transistor count of integrated circuit chips vs. time [2]. The true magic of Si MOSFETs scaling lies in the simultaneous exponential improvement in energy, speed and cost when the transistor decreases in size following a
geometrical law – the perfect manifestation of the motto “smaller is better”. The family of solid black lines in Figure 1-3 sketch the improvement of Si MOSFET in both energy and performance with geometrical scaling in a fashion similar to Figure 1-1 (b). This magic has brought personal and mobile computing revolution, and even laid the foundation of an upcoming artificial intelligence revolution [3]. It has weaved electronics into the fabric of our everyday life by enabling products and services like personal computers, internet, smart phones, digital cameras and wearables, as all these applications (shown in pictures in Figure 1-3) are fundamentally enabled by the continuous improvement in the trade-off of energy and speed.

Moore’s Law – The number of transistors on integrated circuit chips (1971-2016)

Moore’s law describes the empirical regularity that the number of transistors on integrated circuits doubles approximately every two years. This advancement is important as other aspects of technological progress – such as processing speed or the price of electronic products – are strongly linked to Moore’s law.

Figure 1-2. Transistor count in integrated circuit chips as a function of time, reproduced from [2]. Progress shows exponential growth for decades.
Figure 1-3. Progress in energy-delay trade-off across generations of technology and applications. The family of black solid lines represents the fundamental limit of different device technologies that have enabled mainframe to mobile computing. Moore’s law of scaling keeps pushing the limit of transistor technology as indicated by the green arrow, achieving higher performance, lower energy consumption at smaller footprint simultaneously in each successive technology node. The relentless improvement has led to numerous applications represented by the images, including personal computer, internet and smart phones. To materialize the vision of internet of things and deliver applications such as virtual reality and implantables highlighted in dotted circles, significant breakthroughs in device technologies are required in order to push the limit in energy-delay trade-off to the red dashed line.
1.1.3 The need for ultra-low power devices

The relentless scaling of Si MOSFETs has changed almost every aspect of human society and the tremendous value of continuing the microelectronics revolution has been well recognized. In the future, microsystems are expected to be embedded in literally everything ranging from human brain to city infrastructure and stay connected to the cloud. The swarm of electronics connected to the cloud represents the paradigm of Internet of Things (IoT) [4]. To fully materialize this vision, however, significant breakthroughs in device technologies are required, as argued below.

Many envisioned sensing / edge / intelligent systems are subject to practical constraints in real-world applications, which limit the size and amount of energy storage (i.e., batteries) a sensing node can have. For example, implantables [5] including pacing, cochlear and retinal implants are necessary to have an operational lifetime of years as invasive surgery is generally required for implantation, posing a significant challenges on the power efficiency of the electronic systems. Ubiquitous wireless networks [6], essential to IoT infrastructure, embedded in building walls, vehicles, and sidewalks are also severely power constrained. Since energy storage elements such as batteries and capacitors are not experiencing nearly the same rate of scaling as integrated circuits, anatomic- and/or environmental-miniaturization of electronic systems require much more energy efficiency of constituent system components, down to nW or even lower [5, 7, 8]. Usually in these applications the performance can be delivered with current technologies but the bottleneck is to achieve the ultra-low power needed for acceptable performance. In Figure 1-3, these applications are highlighted in the dotted circles in the lower right corner. In the regard of low-power electronics, the innovations on the circuit and system
levels have made tremendous progress [9], enabling numerous applications as smart phones, wearables, etc. However, a particular technology sets the fundamentally limit (black lines in Figure 1-3) of how far the energy-performance trade-off can be pushed with novel circuits and systems. Potential applications beyond the limit of technology (dotted circle in Figure 1-3) will not be accessible to innovations on the circuit and system levels, and can only be enabled by better technologies. Improved device technologies also help exploit the potential of novel circuit topologies and system architectures in a more effective way. Si MOSFETs technologies are experiencing major difficulties to deliver adequate performance with ultra-low power constraints, as shown in the next section. To enable the exciting IoT vision, new device technologies are necessary to achieve ultra-low power operation with moderate performance needed for the envisioned sensing/edge/intelligent systems.

On the cloud front, power efficiency is also the primary consideration in servers, storage and network infrastructure systems. In 2014, data centers alone consumed about 1.8% of all the electricity in the United States [10], and is still increasing at an annual rate of 4%. In contrast, Si MOSFET scaling has recently entered an era of ‘power-constrained scaling’ as the power density dissipated by logic chips hit about $100 \text{W/cm}^2$, which is practically limited by packaging and cooling costs [11]. To fulfill the substantial performance requirement while further lowering the power density of logic chips, device innovations are fundamentally needed to help push cloud computing infrastructure toward the lower left corner of Figure 1.3.

In summary, the continuous scaling of Si MOSFETs for over five decades has led to an exciting vision of cloud-connected electronics embedded in everything. However, we are experiencing difficulties to realize both ultra-low power systems with adequate performance target and high-performance cloud computing infrastructure with acceptable power
consumption with Si MOSFET technology. To shed light on the possible methods to sustain the wide spectrum of applications, the limit of Si MOSFETs technology is first reviewed in the next section.

1.1.4 The Limit of Si MOSFETs Technology

The limit of Si MOSFET technology can be best understood by reviewing the challenges and solutions in the recent history of scaling. Figure 1-4 (a) [12] graphs the supply voltage of logic chips as a function of technology node. It reveals that $V_{dd}$ followed the exponential scaling rather well according to constant field scaling scheme until the 0.13 μm technology node, where it bent over with a much slower scaling factor. The reason behind the slower scaling of $V_{dd}$ is the non-scalable $V_t$, shown in Figure 1-4 (a).

Figure 1-4. (a) Power supply voltage ($V_{dd}$) and threshold voltage ($V_t$) vs. technology generation, reproduced from [12]. (b) Active and subthreshold power consumption in integrated circuits in various technology nodes, reproduced from [14]. The subthreshold leakage power grows exponentially as a result of $V_t$ scaling.

To understand the flat threshold voltage curve, let us take a look at the transfer characteristics of a typical MOSFET shown in Figure 1-1 (a). Below the threshold voltage, the
current does not drop to zero directly, but rather assumes an exponential relationship with gate voltage. Its slope, denoted by the subthreshold swing (S), which is defined as the gate voltage required to change the drain current by a decade when the transistor is operated in the subthreshold region, has a fundamental limit at room temperature, which is non-scalable. MOSFETs rely on thermionic injection of carriers over an energy barrier and the carrier distribution over the thermal barrier obeys the Fermi-Dirac distribution, which can be approximated by the Boltzmann distribution. In the subthreshold region, this gives rise to the $\ln 10 \times \frac{kT}{q} = 60 \text{ mV/dec}$ limit [13]. The leakage current $I_{\text{off}}$ increases exponentially with linearly scaled $V_t$, as the slope is roughly constant below $V_t$. In order to control the leakage power, $I_{\text{off}}$ cannot be too high since the leakage power of a logic chip is proportional to the leakage current, as shown by the following equation:

$$P_{\text{diss}} = P_{\text{active}} + P_{\text{leak}} = \alpha f CV_d^2 + V_{dd} I_{\text{off}}$$

(1-2)

$P_{\text{diss}}, P_{\text{active}}$ and $P_{\text{leak}}$ refer to the total power dissipation, active power and leakage power consumption. As a result, $V_t$ cannot be scaled aggressively. However, even with a slowly varying $V_t$, the subthreshold leakage power consumption in the integrated circuit has grown exponentially and has become the dominating contribution to the total power dissipation, as shown in Figure 1-4 (b) [14] where the active and leakage power are shown as a function of technology node. Tri-gate technology was introduced partly because of its improved subthreshold slope to control leakage power [15]. However, tri-gate technologies cannot beat the 60 mV/dec thermal limit of S and the supply voltage cannot be scaled to very low values as a result. As shown in Eq. 1-2, the dissipation power cannot reach ultra-low levels because it is
proportional to the square of the supply voltage. In this sense, the thermal limit of $S$ represents the fundamental challenge to achieve adequate performance with ultra-low supply voltage.

With the poorly scalable $V_t$, $V_{dd}$ cannot continue to scale down because of the loss of performance due to the reduction of gate overdrive ($V_{dd} - V_t$). As a result of the slower scaling of supply voltage, the active power of the logic chips no longer goes down but rises up instead as shown in Figure 1-4 (b), eventually causes overheating issues, especially in high-density, high-performance circuits like the CPUs and GPUs. However, even with a slower scaling $V_{dd}$, the gate overdrive is still decreasing. To compensate for the drive current loss, transport-enhanced channel engineering has been widely adopted, with the most well-known technology being strained-silicon [16]. Despite the great success of strained-silicon technology in the last decade, the quest for high device density in advanced technology nodes makes continued enhancement of strain engineering increasingly difficult. Mechanical strain and associated performance gain has started to diminish due to aggressive transistor pitch scaling. To continue Moore’s Law of scaling to meet the requirement of high-performance computing, it is imperative to find an effective way to enhance carrier transport in scaled dimensions while reducing the power consumption.

### 1.2 III-V Vertical Nanowire Transistor Technology

One of the most promising directions to deliver high performance logic chips while enabling power reduction is to replace the Si channel in a MOSFET with new materials such as InGaAs and Ge which have better carrier transport properties [11, 17]. The improved transfer characteristics expected from these new channel materials are shown in the blue line in Figure 1-5, potentially achieving higher drive current at a lower supply voltage. However, MOSFETs
with new channel materials cannot beat the 60 mV/dec thermal limit for subthreshold swing and are therefore still quite limited to achieve ultra-low power applications. In this regard, steep-subthreshold device technologies which has the potential to overcome the thermal limit of 60 mV/dec for S are being actively pursed, with the Tunnel FET (TFET) being one of the most promising candidates due to its similarity in device technology to the MOSFET. The anticipated transfer characteristics are shown in the green line in Figure 1-5, offering much better performance at very low supply voltage. Nonetheless, most TFET concepts suffer from inherent trade-offs between steep S and drive current [18] and it is of paramount difficulty to achieve high drive current with steep S. Consequently, the TFET faces great challenges in applications demanding high-performance.

Figure 1-5. Sketch of subthreshold characteristics of Si MOSFET (black solid line), III-V/Ge MOSFET (blue solid line) and a prototypical steep-slope device (green solid line) with the same off current. MOSFETs with novel channel materials such as III-Vs and Ge are expected to deliver higher performance at reduced supply voltage because of the superior carrier transport properties. Steep-slope devices including TFETs with sub-thermal subthreshold swing are promising device technologies to further reduce the power consumption.
In this thesis, III-V vertical nanowire transistor technology is proposed and investigated as a single technology platform for both ultra-low power and high-performance applications based on III-V MOSFETs and TFETs. The motivation behind III-V materials, vertical channel and nanowire geometry are given in this section.

1.2.1 Motivation for III-V’s

1.2.1.1 Transport enhancement

III-V materials, especially InGaAs and InAs are known for their superior electron transport properties, which can transfer to better device technology to continue Moore’s law for high-performance computing applications. For modern deeply scaled transistors, the drive current, $I_{on}$ is equal to:

$$I_{on} = v_{inj} \times Q_{inj} \equiv v_{inj} \times C_{inv} \times (V_{dd} - V_t)$$ (1-3)

$(V_{dd} - V_t)$ is the gate overdrive and $C_{inv} \times (V_{dd} - V_t)$ gives the inversion charge density at the virtual source. For a given charge density, the drive current is proportional to the source injection velocity, $v_{inj}$ [19]. Figure 1-6 graphs the injection velocity extracted from InGaAs and InAs high-electron-mobility-transistors (HEMTs) and Si MOSFETs [11]. The $v_{inj}$ values of InGaAs and InAs are more than twice that of comparable silicon MOSFETs at less than half the voltage due to the lighter electron effective masses, which promises higher performance than Si MOSFETs while lowering the power consumption. This extraordinary transport characteristics has led to the demonstration of InAs quantum-well MOSFETs with peak transconductance ($g_{m,pk}$) as high as 3.45 mS/µm at a drain bias of 0.5 V [20].
1.2.1.2 Band structure engineering

A major concern for TFET is the disadvantage of low current levels, as mentioned above briefly. To obtain higher current, the tunnel probability in a TFET [18] shown below needs to be increased:

$$T_{WK} = \exp\left(-\frac{m_{eff}^2 E_b}{2\sqrt{2\epsilon_0|\phi|}}\right)$$  \hspace{1cm} (1-4)

According to this equation, the tunnel probability is inversely proportional to the exponential function of the electron effective mass $m_{eff}$ and tunnel barrier height $E_b$. Therefore, small effective mass and tunnel barrier height are highly preferred for high-performance TFETs. III-V
materials are widely recognized for their direct bandgap and small effective masses [21], 0.023 \( m_0 \) for bulk InAs as compared to 0.19 \( m_0 \) for Si. Furthermore, III-Vs embrace flexible and powerful band structure engineering capabilities, which has resulted in an extremely high peak-to-valley current ratio over 600 in GaAs/InGaAs/InAs quantum-well resonant tunneling diodes [22]. The ability to achieve staggered and broken-gap heterojunctions in junctions such as AlGaSb/InAs and InAs/GaSb with very small tunnel barrier height is demonstrated to boost the TFET ON current greatly by shortening the tunneling distance.

1.2.2 The necessity of nanowire geometry

Scalability is the primary concern in determining device technologies. As discussed earlier, the challenge in MOSFET scaling is maximizing performance (i.e. drain current density) at reduced voltage. Footprint scaling demands shrinking all dimensions, including the gate length, in a harmonious way. With the ever-shrinking device dimensions and gate length in particular, short-channel effects (SCEs) in MOSFETs such as drain-induced-barrier-lowering (DIBL) and degradation of subthreshold swing hurts leakage power, performance and device variability [23]. Mitigating short-channel effects as the gate length scales down, requires enhanced gate control of the channel. This has dictated looking beyond the planar bulk or thick SOI MOSFET to ultra-thin body MOSFET and eventually to so-called multi-gate structures such as the Double-gate or FinFET, Tri-gate MOSFET and the Gate-All-Around Nanowire MOSFET in a horizontal or vertical configuration, as shown by the evolution of MOSFET architecture in Figure 1-7 (a) (reproduced from [24]). For the same channel length and Si body thickness, GAA offers much better electrostatics than double-gate (DG), tri-gate (TG) or extremely thin silicon-on-insulator (ETSOI) architectures, evidenced by the DIBL vs effective
gate length ($L_{\text{eff}}$) shown in Figure 1-7 (a), reproduced from [23]. Therefore, a GAA architecture with a nanowire channel is regarded as the ultimate scalable technology [23, 25].

![MOSFET structural evolution](image)

Figure 1-7. (a) MOSFET structural evolution, reproduced from [24]: 3D schematic (top) and cross-sectional schematic (bottom) of MOSFET structures with increasing electrostatic gate control of the channel from left to right. (b) DIBL as a function of effective gate length shows improvements in SCE as the gate number is increased from ETSOI to GAA, reproduced from [23].

### 1.2.3 The advantage of a vertical channel

The advantage of aligning the direction of current flow perpendicular to the wafer surface in a MOSFET, e.g. a vertical channel, has been attractive to researchers in recent years as an option to continue Moore’s law [26-29]. When combined with the outstanding SCEs of the GAA geometry, vertical nanowire (VNW) transistors, schematically show in Figure 1-8 (a), have the potential to carry Moore’s law further into the future than any other device structure [26]. The fundamental advantage is the uncoupling of footprint scaling on the plane of the wafer from gate length ($L_g$) scaling, which takes place along the vertical dimension. Gate length, spacer thickness ($L_{\text{spacer}}$) and contact length ($L_c$) scaling can now be relaxed from the huge pressure of contacted gate pitch scaling. Consequently, as shown in Figure 1-8 (b) [26], the performance – energy trade-off of vertical NW (VF) is greatly improved compared to lateral
NW (L2 & L3) and FinFET (FF) architecture. Furthermore, [29] shows that a 30% area reduction is enabled in SRAMs with vertical channel transistors.

Figure 1-8. (a) Schematic of vertical nanowire transistor architecture showing the contact length ($L_c$), spacer thickness ($L_{\text{spacer}}$) and gate length $L_g$. (b) Simulated performance of ring oscillators (RO) with FinFETs (FF), vertical nanowire FETs (VF) and two flavors of lateral nanowire FETs (L2 and L3) under 5 nm technology node design rules, reproduced from [26]. At each data point, device technology is optimized for minimum RO delay at a given supply voltage and fixed off current.

More importantly, the vertical channel is capable of fully leveraging the potential of the band structure engineering capabilities of III-Vs. The transport now takes place in the direction of the epitaxial growth direction which enables band structure engineering along carrier transport direction, which is essential for tunnel barrier height engineering for III-V TFETs. For III-V MOSFETs, this opens up a whole class of brand new opportunities, e.g. reducing leakage current by drain and source bandgap engineering, strain engineering to boost carrier density and velocity at the same time, etc.
1.2.4 Top-down approach

III-V vertical nanowires have been demonstrated via direct nanowire growth or bottom-up approach with a number of methods, including vapor-liquid-solid (VLS), selective area epitaxial (SAE) and template-assisted growth techniques. Figure 1-9 (a) [30] shows a tilted SEM image of InAs nanowire array with 60 nm diameter and 1.3 μm length, grown on InAs substrate with VLS technique using Au as the catalytic particle. As a comparison, top-down approach relies on plasma etching of planar wafers with lithographically patterned mask patterns to obtain structures. While bottom-up techniques offer great promise for future Si integration, top-down approach benefits from much more powerful and established planar...
epitaxial techniques that can produce extremely complicated heterostructures with atomically sharp interfaces, and is therefore an excellent research vehicle for technology development and device physics exploration. Figure 1-9 (b) presents a cross-sectional TEM image of an InGaAs/InAlAs superlattice with 1 nm period and atomically sharp interfaces\(^1\), as a showcase of the capability of the planar epitaxial technique. In general, top-down approach which has been perfected with high reproducibility and yield dominates the industry and can leverage the accumulated industry expertise and infrastructure, as opposed to relative new bottom-up techniques. In this thesis, top-down techniques to define III-V nanowires are developed.

### 1.2.5 III-V vertical nanowire transistor technology

Vertical nanowire transistor technology has the potential to carry Moore’s law through to the end due to its ultimate scalability. With the insertion of III-Vs as the channel material, both high-performance computing and ultra-low power systems can potentially be realized by III-V MOSFETs and TFETs, respectively. In addition, the combination of MOSFETs and TFETs enables tremendous circuit and system design flexibility, which can significantly boost their performance [31]. Although TFETs and MOSFETs are needed, the only major difference between them are the channel materials and these two devices share most of the process technologies in common. This means that the development of vertical nanowire transistor technology can enable very broad applications from having both TFETs and MOSFETs while adding only a fraction of the costs and complexity.

\(^1\) Material growth and TEM imaging were done by Ryan Iutzi at MIT.
1.3 Thesis Overview

With the upcoming internet of things revolution, the Si MOSFET is facing increasing difficulty in fueling the wide-range of application space spanning from ultra-low power systems to high-performance computing infrastructure. Recognizing the unique advantage of III-V VNW transistor technology, the goal of this thesis is to demonstrate III-V VNW MOSFETs and TFETs and examine the potential of this technology for future ultra-low power and high-performance computing applications. The thesis is organized in the following way.

Chapter 2 introduces the fabrication techniques and process integration for III-V vertical nanowire transistor technology. An overview of the process flow is given in the beginning, followed by a detailed discussion of the most critical process module, III-V nanowire formation with sub-10 nm diameter resolution and high aspect ratio. Gate stack module, comprised of atomic layer deposition and gate metal sputtering is introduced next. The chapter concludes with a thorough documentation of the backend process module, including planarization and etch back, via etch and contact formation.

Chapter 3 focuses on the results of top-down InGaAs VNW MOSFETs. Three generations of devices are fabricated and the results of electrical characterization are carefully studied, followed by performances benchmark with other published III-V VNW MOSFETs. The electrostatics of III-V VNW MOSFETs are investigated in detail and an estimation of the interface trap density is provided. The comparison of first two generations of devices enables the first experimental study of source/drain asymmetry intrinsic to vertical nanowire transistor architectures. The major show stopper, the top semiconductor/metal contact is discussed.
Chapter 4 summarizes the device demonstrations of top-down InGaAs/InAs heterojunction VNW TFETs. The device structure of the first generation TFET and its DC characteristics at room temperature are presented. The temperature dependence is then analyzed in detail by examining the output, subthreshold and conductive-voltage characteristics. A second generation TFET which suppresses the temperature dependence and obtains sub-thermal switching is introduced. The comparison of the first and second generation TFETs establishes oxide-semiconductor interface trap-assisted tunneling as a major thermal leakage mechanism that overshadows the true tunneling steepness, and also enables the extraction of the band edge steepness in VNW TFETs. Devices are benchmarked with other published III-V VNW TFETs, revealing record performance achieved with top-down approach in this thesis.

Chapter 5 concludes the thesis with a summary of its contributions and suggests future directions to improve the process, MOSFETs and TFETs. A novel device concept, superlattice-source NW-FET is introduced in the end.
Chapter 2 Process Development

As discussed in Chapter 1, the goal of this thesis is to demonstrate III-V vertical nanowire transistors (MOSFETs and TFETs) fabricated via a top-down approach. Cross-section of an exemplary MOSFET and a TFET are shown schematically on the left and right in Figure 2-1. Except for the nanowire semiconductor, the rest of the structures are almost identical in the two devices, echoing with the previous argument that a single technology platform can support a broad application space by providing both MOSFETs and TFETs. To ensure scalability for future ultra-scaled technology nodes, sub-10 nm diameter nanowires [26] in different III-V materials with high quality surface need to be prepared. In this chapter, a generic process flow is described first, followed by a detailed discussion of the critical process modules. Processes that are different from this generic process flow is specified as necessary. The fabrication technology developed in this chapter forms the basis for device demonstration in the rest of this thesis.

Figure 2-1. Schematics of the cross-section of MOSFET (left) and TFET (right).
2.1 Process Overview

MOSFET and TFET fabrication starts with planar wafers with different heterostructures. MOSFET heterostructures are grown by metal-organic chemical vapor deposition (MOCVD) on 2-inch semi-insulating (100) InP substrates by Prof. Fitzgerald’s group at MIT. Molecular beam epitaxy (MBE) is used by IntelliEpi Inc. to grow TFET heterostructures on 2-inch semi-insulating (100) InP substrates. Typical heterostructure designs for MOSFET and TFETs are shown in Figure 2-2. From bottom to top, the MOSFET structure comprises of a 175 nm InP buffer, a 300 nm heavily n+ doped In0.53Ga0.47As bottom contact layer, an undoped 80 nm In0.53Ga0.47As channel layer and a 70 nm heavily n+ doped In0.53Ga0.47As top contact layer. In essence, the MOSFET structure features an n+/i/n+ doping profile and In0.53Ga0.47As lattice matched to InP. The top contact layer can function as either source or drain terminal. An improved heterostructure for better top contact was also used, as discussed in Chapter 3.

In contrast, the TFET heterostructure uses p+/i/n+ doping profile and the tunnel junction is formed between the p+ and intrinsic layers, enabling an n-type transistor design. From bottom to top, the structure consists of a 400 nm undoped In0.52Al0.48As buffer, 280 nm heavily p+ doped source contact layer, a 20 nm p++ layer serving as the source of the tunnel junction, 2 nm intrinsic InAs/8 nm In0.7Ga0.3As “notch” layer followed by a 50 nm undoped In0.53Ga0.47As channel, and a composite drain contact layer comprised of n+ In0.53Ga0.47As, In0.7Ga0.3As and InAs. The design principles of MOSFET and TFET heterostructures are discussed in Chapters 3 and 4, respectively. Different heterostructures have been used for process development, which are identified along the discussion.
### MOSFET

<table>
<thead>
<tr>
<th>Material</th>
<th>Thickness (nm)</th>
<th>Dopant</th>
<th>Level (/cm³)</th>
<th>Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>In₀.₅₃Ga₀.₄₇As</td>
<td>70</td>
<td>Si</td>
<td>6×10¹⁹</td>
<td>n⁺</td>
</tr>
<tr>
<td>In₀.₅₃Ga₀.₄₇As</td>
<td>80</td>
<td>-</td>
<td>UID</td>
<td>i</td>
</tr>
<tr>
<td>In₀.₅₃Ga₀.₄₇As</td>
<td>300</td>
<td>Si</td>
<td>6×10¹⁹</td>
<td>n⁺</td>
</tr>
<tr>
<td>InP</td>
<td>175</td>
<td>-</td>
<td>UID</td>
<td>i</td>
</tr>
<tr>
<td>InP</td>
<td>-</td>
<td>-</td>
<td>UID</td>
<td>i</td>
</tr>
</tbody>
</table>

### TFET

<table>
<thead>
<tr>
<th>Material</th>
<th>Thickness (nm)</th>
<th>Dopant</th>
<th>Level (/cm³)</th>
<th>Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>In₀.₅₃Ga₀.₄₇As</td>
<td>5</td>
<td>Si</td>
<td>6×10¹⁹</td>
<td>n⁺</td>
</tr>
<tr>
<td>InₐAs</td>
<td>2</td>
<td>Si</td>
<td>6×10¹⁹</td>
<td>n⁺</td>
</tr>
<tr>
<td>In₀.₅₃Ga₀.₄₇As</td>
<td>8</td>
<td>Si</td>
<td>6×10¹⁹</td>
<td>n⁺</td>
</tr>
<tr>
<td>In₀.₅₃Ga₀.₄₇As</td>
<td>40</td>
<td>Si</td>
<td>6×10¹⁹</td>
<td>n⁺</td>
</tr>
<tr>
<td>In₀.₅₃Ga₀.₄₇As</td>
<td>50</td>
<td>-</td>
<td>UID</td>
<td>i</td>
</tr>
<tr>
<td>InₐAs</td>
<td>8</td>
<td>-</td>
<td>UID</td>
<td>i</td>
</tr>
<tr>
<td>InₐAs</td>
<td>2</td>
<td>-</td>
<td>UID</td>
<td>i</td>
</tr>
<tr>
<td>InₐAs</td>
<td>20</td>
<td>C</td>
<td>1×10¹⁰</td>
<td>p⁺</td>
</tr>
<tr>
<td>InₐAs</td>
<td>280</td>
<td>C</td>
<td>5×10¹⁹</td>
<td>p⁺</td>
</tr>
<tr>
<td>In₀.₅₃Al₀.₄₇As</td>
<td>400</td>
<td>-</td>
<td>UID</td>
<td>i</td>
</tr>
<tr>
<td>InP</td>
<td>-</td>
<td>-</td>
<td>UID</td>
<td>i</td>
</tr>
</tbody>
</table>

---

Figure 2-2. Exemplary heterostructures of MOSFET and TFET used in this thesis.

Major fabrication steps are shown in the process flow in Figure 2-3. The left and right column of images show the top and side (cross-section) view after the processing steps. The starting wafer is cleaved into 1.2 × 1.2 cm² pieces. This is followed by soft ultrasonic in water to remove major particles, if any, and solvent cleaning with acetone/ methanol/ isopropanol. 5/60 nm Ti/Au is then deposited by PMMA (PMMA A8 from MicroChem Inc.) lift-off exposed by electron beam lithography, serving as alignment marks for EBL steps later. After deposition of 2-3 nm Si₃N₄ as the adhesion layer, HSQ, a high-resolution negative-tone resist with high dry etch resistance is first exposed via electron beam lithography (EBL) in the shape of dots with
Figure 2-3. Typical process flow for III-V VNW transistor fabrication.
different diameters in the range of a few tens of nanometers and developed in 25% TMAH in water solution. With HSQ serving as the hard mask (Step 1 in Figure 2-3), NWs are then formed by reactive ion etching (RIE) with BCl3/Ar/SiCl4 chemistry and the unmasked semiconductor is etched away (Step 2 in Figure 2-3). As there is no etch stopper for the selected chemistry, the etching depth is controlled by timing the process and is 200 – 250 nm typically. A digital etch (DE) technique is then performed to further thin down the nanowire and improve the sidewall quality after receiving RIE damage (Step 3 in Figure 2-3).

In order to achieve a uniform coverage of gate dielectric and gate metal, ALD High-κ dielectric and sputtered gate metal are chosen due to the excellent conformality as well as the ability to reduce effective-oxide-thickness (EOT) for better scalability. After DE, the sample is dipped in diluted H2SO4 in DI water as the final surface preparation step and transferred to the ALD chamber where Al2O3 is deposited, followed by 40 nm W sputtering as the gate metal. HfO2, another industry-standard High-κ material that can potentially deliver better electrical performance because of its higher dielectric constant compared to Al2O3, is not used here because HfO2 cannot be selectively removed against FOX before the contact metal is put down. Forming gas anneal (FGA) is then applied to the sample at 350 ºC for 30 min to improve the oxide/semiconductor interface.

Since the gate metal is now covering the whole wafer surface and the sidewall of the nanowire as a result of the ALD and sputtering processes, a gate metal patterning step is needed to remove the unwanted W. The W patterning involves two steps. A PMMA lift-off process with EBL is first done to evaporate a 5/20 nm Ti/Au layer onto a small area where the gate via is be opened. This thin Au layer serves as an etch stopper that protects the W gate from the bottom contact via etch, as explained later together with via opening process. Then a 300 nm
thick negative tone EBL resist, maN 2403 from *micro resist technology gmbh*, thick enough to cover the nanowires and resistive to W dry etch, is exposed and developed by EBL, leaving the desired gate metal pattern and nanowires protected. Unwanted W is dry-etched away in SF₆ chemistry selectively to the gate dielectric protecting the bottom semiconductor plane, completing the front-end processing (Step 4 in Figure 2-3).

Backend process to connect device terminals to pads begins with removing the gate metal from the top of the nanowire where the top contact needs to be formed. The challenge is to control the alignment of the gate metal to a certain position in the semiconductor heterostructure (depending on the device design) as precise as possible, 10 nm misalignment at most. As this alignment is important for both MOSFETs and TFETs, Chapter 3 and 4 address this issue in more detail. In this work, the removal of gate metal and subsequent isolation between the gate metal and top contact metal deposited later are done by two successive planarization and etch back processes. Flowable oxide, (FOX), a class of the spin-on glass (SOG) material of 450 nm (essentially a thicker version of HSQ) is deposited by spin-coating as the planarization agent and baked to drive away the solvents. Etch back proceeds in the CF₄ chemistry which etches W at a slow rate so W can still protect the top of the nanowire. After the etch back is stopped by timing, a low power SF₆ chemistry is used to remove W completely from the top of the nanowire while minimizing the dry etch damage (Step 5 in Figure 2-3). The details of the etch back process, including the control of timing is described in section 2.4.1. Next, another planarization and etch back is performed with HSQ to leave another 10-20 nm material to isolate the gate metal with contact metal deposited later. The gate dielectric Al₂O₃ is still present since the etch rate is very slow in CF₄.
Contact via is opened next with dry etch in CF$_4$ chemistry through FOX layers (Step 6 in Figure 2-3) to expose the gate metal (W) and semiconductor at the bottom plane, with ZEP520A$^2$, a positive EBL resist from Zeon Chemicals as the mask. Since the W metal surface is 40 nm higher than the semiconductor plane, in order to etch the FOX completely from the bottom semiconductor surface, W gate can be attacked by CF$_4$ dry etch and even damaged or removed completely. To protect W gate metal, Ti/Au is lifted-off on top of W, as mentioned above. The sample is then annealed in N$_2$ ambient for 1 hr at 350 °C to cure the FOX, preventing it from being attacked by a subsequent 25% TMAH dip to remove the gate dielectric Al$_2$O$_3$ that still covers the top of the nanowire. HfO$_2$ cannot be used as the gate dielectric because it cannot be removed selectively to FOX. After TMAH dip, the sample is immersed in 1:3 HCl:water solution to remove the native oxide and transferred immediately to a vacuum chamber where 40 nm of Mo is sputtered as the contact metal to both top and bottom semiconductor materials. 20/250 nm Ti/Au is then evaporated by a PMMA lift-off process, followed by Mo dry etch in SF$_6$/O$_2$ chemistry, completing the device fabrication (Step 7 in Figure 2-3). The process modules including nanowire formation, gate stack and backend process are discussed next following the order in the process flow. The detailed process flow and recipes are documented in the Appendix.

$^2$ Supply of resist Zep520A was discontinued. Resist gL2000 from MicroChem is used as the replacement.
2.2 NanowireFormation

2.2.1 Electron beam lithography

The electron beam lithography technique used in this thesis has been treated in detail previously [32] and a short summary is given here. HSQ used in this work was a flowable oxide (FOX) diluted with Methyl-IsoButyl-Ketone (MIBK), available under trade name of XR-1541™ from Dow Corning. As shown in [32], HSQ does not adhere well to the surface of III-Vs, especially for small features. It is well known that HSQ exhibits good adhesion on Si due to the formation of Si-Si bonds [33], so a thin layer (2-3 nm) of Si₃N₄ was deposited by CVD, serving as the adhesion layer between InGaAs and HSQ. After spin coating, the patterns are exposed in an Elionix ELS-F125 system, with which all the patterning steps in the process are done. It can achieve high precision alignment down to a few tens of nanometers. A tilted SEM image of HSQ mask with 30 nm diameter is shown in Figure 2-4.

Figure 2-4. Tilted SEM image of HSQ mask with 30 nm diameter pattern.
2.2.2 Reactive ion etching

The most important process module for top-down vertical nanowire transistor technology is a dry etch that can deliver high-aspect ratio and nanometer-scale NWs with vertical and smooth sidewalls in a variety of III-V materials containing In, Ga, As, Sb and P elements. In [32], vertical etching of GaAs is achieved with BCl3/N2 chemistry at 40 °C, which can also be extended to GaSb. Exemplary GaAs and GaSb nanowires fabricated by BCl3/N2 chemistry are shown in Figure 2-5. The first antimonide FinFETs [34] were fabricated using this recipe. However, this recipe does not work well on InGaAs, evidenced by the comparison of GaSb and InGaAs NWs etched at the same time in Figure 2-6. The etch depth decreases from 210 nm to 85 nm for InGaAs and the InGaAs NW sidewall is much more tapered.

Figure 2-5. Tilted SEM image of GaAs (left) and GaSb (right) NWs fabricated with BCl3/N2 recipe at 40 °C.

GaSb nanowires were etched by Wenjie Lu at MIT.
Dry etch of indium-containing III-V materials, mostly InP/InGaAsP, has been studied extensively for optical device applications [35-37]. High verticality and high aspect ratio have been demonstrated in a variety of chemistries, including those based on CH₄/H₂ [38], Cl₂ [35, 37, 39], BCl₃ [40, 41], SiCl₄ [42], and HBr [36]. Sub-20 nm features in In-based heterostructures with vertical sidewalls fabricated by RIE was demonstrated for the first time in this thesis by improving and further optimizing the BCl₃/N₂ chemistry used in [32] for GaAs.

Successful dry etch is developed [43] using BCl₃/SiCl₄/Ar chemistry in a SAMCO RIE-200iP ICP (inductively coupled plasma) system with a heated chuck and backside cooling. Samples cut into small pieces are loaded on a 6-inch ceramic carrier wafer without thermal grease. To avoid loading effects, samples with a total area of 1 cm × 1 cm are placed around the test pieces. Further details are included in the Appendix. A systematic study was performed with the MOSFET heterostructure shown in Figure 2-2 (In₀.₅₃Ga₀.₄₇As) by varying substrate temperature, gas flows, chamber pressure and RF platen power.
Figure 2-7 shows a NW with a diameter of 15 nm (D = 15 nm) and aspect ratio greater than 10 fabricated by optimized RIE conditions: 20 W ICP power, 160 W RF platen power (resulting in 280 V substrate bias voltage), 0.2 Pa chamber pressure, 250 °C substrate temperature and gas flows of 7/0.55/7 SCCM for BCl₃, SiCl₄ and Ar, respectively. The etching rate is about 1.8 nm/s (total etching time is 135 s) and the selectivity to the HSQ mask is approximately 8:1. The etched NW features very smooth sidewalls and nearly vertical profiles.

Figure 2-7. 15 nm diameter InGaAs NW defined by optimized RIE technique with an aspect ratio greater than 15.

A slight footing behavior towards the bottom of the nanowire and some degree of trenching is present [37]. Although the roughness on the sidewall is below the detection limit of SEM in the presence of imaging noise, quantitative information cannot be obtained at this moment. Stoichiometry on the sidewall is also unclear at this moment also because of the difficulty to
perform chemical analysis on the sidewall of these small NWs. However, the sidewall stoichiometry might be altered from that of the bulk according to [44], which shows that the sidewall of the InGaAs fin can be As-rich after un-optimized dry etch. Further process development on the aspect of dry etch and metrology is needed to address these issues.

The results shown in Figure 2-7 are obtained after systematic optimization of RIE conditions. For the purpose of illustrating the role of each etch parameter, we compare InGaAs pillars etched with optimal conditions in Figure 2-8 (a) and non-optimal conditions in Figure 2-8 (b) through (e). All samples are etched for 135 s. When pressure is increased from 0.2 to 0.8 Pa in Figure 2-8 (b), we see a clear undercut. This is partially due to the less directional ion bombardment that results from enhanced ion collisions in the plasma under higher pressure [41]. Enhanced chemical etching also contributes to the undercut since more reactive species are produced at higher pressure [35]. A similar undercut appears when the RF platen power is decreased from 160 to 50 W (substrate bias voltage reduced from 280 to 120 V) as in Figure 2-8 (c) as a consequence of less directional ions. Trenching around the sidewall is greatly reduced in (b) and (c), indicating that this is related to bombarding ions bouncing off the sidewall at an angle.

The BCl₃/Ar plasma alone is highly directional and causes little lateral etching [41]. Introducing SiCl₄ into the plasma offers more lateral etching due to its low anisotropy at low ion density [42]. This changes the profile from tapered to vertical. Without SiCl₄ we only obtain tapered sidewalls. This can be seen in Figure 2-8 (d) that features a reduced flow of SiCl₄.

The balance between BCl₃ and Ar affects the surface roughness and etching profile. With a higher Ar to BCl₃ ratio as in the case of Figure 2-8 (e), the surface roughness is decreased because of more balanced removal rates of difference elements in InGaAs due to
enhanced physical sputtering. However, this enhancement leads to lower selectivity to HSQ mask (4.5:1). The etching profile also develops a slight undercut.

Figure 2-8. InGaAs pillars etched under different RIE conditions: (a) optimum parameters (see text); (b) through (e) one parameter different from (a). (b) Chamber pressure = 0.8 Pa (vs. 0.2 Pa), (c) RF platen power = 50 W (vs. 160 W, with substrate bias voltage of 120 V vs. 280 V), (d) SiCl₄ flow rate = 0.25 SCCM (vs. 0.55 SCCM), (e) BCl₃/Ar flow rate = 2/12 SCCM (vs. 7/7 SCCM).
We show the impact of substrate temperature in Figure 2-9. As the temperature is increased from 100 °C to 250 °C (other conditions unchanged), the surface roughness is greatly reduced and the sidewall becomes vertical. The rough surface and positive sloped sidewall at low temperature is caused by the residual InCl₃ that cannot be removed as fast as GaCl₃ and AsCl₃ due to its low volatility [39]. Increasing substrate temperature results in more balanced removal rate of In with respect to Ga and As. The overall InGaAs etch rate rises at higher substrate temperatures.

![Figure 2-9](image)

Figure 2-9. Impact of substrate temperature on etching rate, profile, and surface roughness of InGaAs NWs: (a) 100 °C, (b) 175 °C, (c) 250 °C.

Another concern for the NW dry etch is material selectivity among different III-Vs. As argued in Chapter 1, band structure engineering by heterostructure growth is an important advantage of III-V VNW technology, which requires a uniform and smooth NW sidewall in the presence of complex heterostructures, implying minimal material selectivity between compounds containing elements including In, Ga, Al, As, P and Sb. Using the optimized RIE recipe, 7 nm fin with aspect ratio greater than 5 is demonstrated as shown in Figure 2-10 (a).
and the corresponding heterostructure is captured in Figure 2-10 (b), containing InP and InAlAs layers. The etched fin structure features very smooth sidewalls and no notching is observed, confirming the applicability of the dry etch technique to III-Vs containing In, Ga, Al, As and P.

Figure 2-10. (a) 25 nm fins containing InP and InAlAs, etched with optimized BCl₃/Ar/SiCl₄ recipe. (b) Starting heterostructure used for etching.

State-of-the-art InGaAs FinFETs [45-48] are enabled by the optimized dry etch recipe. Antimonides can also be etched with this recipe, evidenced by a 10 nm fin with smooth sidewall in Figure 2-11. The starting heterostructure is comprised of the 20 nm of In₀.₃₈Ga₀.₆₂Sb on 200 nm Al₀.₆₅Ga₀.₃₅Sb buffer. The absence of material selectivity is due to the strong physical etching component in the optimized recipe, which also leads to the superior directionality.

---

4 Antimonide fins were etched by Wenjie Lu.
2.2.3 Digital etch

RIE damage is well known to Si [49] and III-V materials [50-52] due to the high energy ions bombarding the semiconductor. For vertical nanowire transistor technology, the most critical interface is formed between the semiconductor sidewall that receives RIE damage and the gate dielectric in the channel. Since both MOSFETs and TFETs are surface-channel devices where the current flows in close proximity (a few nanometers) to this interface, it is of paramount importance to improve the quality of the sidewall for better oxide-semiconductor interface. In Si nanowire transistors, high-temperature annealing (> 850 °C) has been shown as a powerful technique to improve the sidewall quality [53]. However, high-temperature annealing does not readily apply to III-V materials [54] as III-Vs are compounds formed with different elements and group V elements tend to escape before effective annealing happens, causing damage rather than improvement. While wet etch is effective in removing RIE damage for III-
Vs [51], the application on ultra-scaled nanowire transistors is unfeasible because of the difficulty of achieving the required nanometer etching accuracy.

In this regard, digital etch (DE) techniques [55], with the characteristics of conventional wet etch but with a greater degree of control, is promising to remove RIE damage. In this approach, the two elemental components of etching, oxidation and oxide removal, are applied separately. One cycle of DE comprises of oxidation. This is followed by oxide removal in acid. For III-Vs, a typical oxidation agent is an H2O2 solution while the native oxide is stripped by acids [55, 56]. Due to the self-limiting nature of the H2O2 oxidation, DE can be very accurately controlled. However, this scheme (wet oxidation + wet oxide removal) can result in chemical cross contamination that leads to poor process repeatability. Our approach consists of native III-V oxide formation through exposure to low power oxygen plasma in a commercial Branson barrel asher system for 3 min, followed by oxide removal in diluted H2SO4 (H2SO4:H2O=1:1) for 1 min at room temperature followed by deionized water rinse. By repeating these two steps, the semiconductor can be removed in a controlled manner.

Figure 2-12, reproduced from [57] shows measured InP (on a planar wafer) digital etch rate per cycle as a function of oxygen exposure time. The InP etch rate increases for the first 150 s of oxidation time but it then saturates to a rate of about 0.9 nm/cycle. Since the oxide removal step completely removes the grown oxide, the etch rate is directly proportional to the oxide thickness. Therefore, the saturation behavior in Figure 2-12 reflects the self-limiting nature of the oxidation process in which once the oxide thickness is thick enough, the diffusion of oxidant species through it is sharply curtailed. For much longer oxidation time, the sample heats up and the oxidation rate increases somehow. The data between 50 to 500 s follows Lukeš’ rate law for GaAs oxidation: \( d = A + B \times \ln(t + t_0) \), with fitting parameters \( A \sim 0, B = \)
0.19 \text{nm/dec}, \text{ and } t_0 = -37 \text{ s} [58]. The negative $t_0$ value reflects the incubation time of the O$_2$ plasma oxidation process.

Figure 2-12. Etch rate per cycle of digital etch as a function of oxidation time under O$_2$ plasma, reproduced from [57] and fitting with Lukeš’ model [58].

Figure 2-13 shows an InGaAs NW with $D = 28$ nm defined by the RIE technology on the left and an identical NW with $D = 18$ nm on the right after 5 cycles of DE, giving $\sim 2$ nm etch rate per cycle. It can be seen that the DE preserves the overall shape and the roughness is still below the SEM detection limit. Due to the lack of proper metrology techniques (high resolution TEM is challenging to do on scaled NWs), quantitative information on the sidewall roughness and stoichiometry is currently unavailable. The sidewall stoichiometry is unlikely to be disturbed significantly, otherwise a porous or very rough sidewall is expected after several cycles of digital etch.
To demonstrate the impact of DE process on the device electrical performance, Figure 2-14 [43] shows the subthreshold characteristics and transconductance ($g_m$) (inset) of a VNW InGaAs MOSFET with and without DE. The device features 30 nm diameter 4.5 nm Al$_2$O$_3$ (EOT=2.2 nm) and 80 nm channel length. Drain current ($I_d$) and $g_m$ are normalized by the NW circumference ($\pi D$). The NW diameter after RIE is different (50 vs. 30 nm) so that after 10 cycles of DE on the thicker one, their final diameters are, within experimental uncertainty, identical and equal to 30±3 nm, as measured by SEM on pieces cleaved from the device samples. Except for DE, both samples went through the same process simultaneously, including a diluted H$_2$SO$_4$ dip prior to ALD. The digital etch is seen to improve the interface quality as evidenced by a significant reduction in subthreshold swing (S) (at $V_{ds} = 0.05$ V) from 190 to

Figure 2-13. (a) D= 28 nm InGaAs NW fabricated by RIE. (b) Same NW as in (a) after 5 subsequent cycles of digital etch.
150 mV/dec. While both devices have a similar ON resistance ($R_{on}$) of about 760 $\Omega \cdot \mu m$, the peak $g_m$ at $V_{ds} = 0.5$ V increases from 155 to 280 $\mu S/\mu m$ indicating a significant reduction in sidewall damage by DE. ON current ($I_{on}$, extracted at $V_{ds} = 0.5$ V and gate overdrive of 1 V) with and without DE is 205 and 130 $\mu A/\mu m$ respectively. Measurements on 10 working devices of each kind show consistent improvement in average $S$ (155 vs. 185 mV/dec at $V_{ds} =$50 mV) and $g_m$ (150 vs. 255 $\mu S/\mu m$ at $V_{ds} = 0.5$ V). These improvements can be attributed to the removal of damaged layers where crystal structure and stoichiometry is disturbed.

As in the case of the RIE, minimal material selectivity is obtained in the aqueous DE technique. Figure 2-15 shows the fin structure comprising of InGaAs, InAlAs and InP in Figure 2-10 after 8 cycles of DE, displaying very smooth sidewalls. However, aqueous DE does not
work on antimonides, an important class of materials with potential for p-channel MOSFETs and tunnel-FETs. Antimonides are highly reactive and can be etched by water, negating the tight etching control that is sought in DE. Another problem with the aqueous DE technique is the inability to deliver sub-10 nm vertical NWs with high yield that is required for sub-10 nm technology nodes. This is a result of nanowire breakage as a consequence of the strong mechanical stress exerted by the water-based acids used in the oxide removal step of DE.

Figure 2-15. Fin structure in Figure 2-10 after 8 subsequent cycles of digital etch, demonstrating the uniformity of DE technique for a heterostructure containing In, Ga, Sb, Al and P.

To overcome these challenges, a novel non-aqueous DE technique is introduced5. The new approach uses acids dissolved in alcohol which has much less surface tension (about a third) than water to enhance the survivability of very thin nanowires. 10% H₂SO₄ in methanol, and 10% HCl in isopropanol (IPA) purchased from Sigma-Aldrich are tested and compared with the previous solution, 1:1 H₂SO₄ in DI water. In all cases, the oxidation step is unchanged, i. e.

5 In collaboration with Wenjie Lu at MIT.
oxygen plasma for 3 min. Figure 2-16 (a) shows an arsenide VNW after 7 cycles of conventional DE. A final NW diameter of 8 nm is expected but all structures are destroyed. In contrast, Figure 2-16 (b) shows an identical sample processed side-by-side in HCl:IPA. 8 nm diameter VNWs are demonstrated with over 97% yield. Table 2-1 summarizes the yield of arsenide VNWs after 7 DE cycles for samples with different initial diameters for water as well as alcohol-based acids. The yield of 10 nm NW is improved from 5% to 100%, demonstrating the superiority of the new technique. Figure 2-17 shows the narrowest arsenide VNW obtained with a 5 nm diameter and 230 nm height (aspect ratio=46) after 10 cycles of DE in 10% H₂SO₄ in methanol, demonstrating the potential of solvent DE to obtain sub-10 nm features.

![Broken NW](image1.png)  
![8 nm](image2.png)

Figure 2-16. Arsenide VNW array after 7 DE cycles in (a) 10% HCl in DI water, and (b) 10% HCl in IPA. The insets show close ups of the etched structures. The nanowire yield in (a) is 0% while the yield in (b) is 97%.

<table>
<thead>
<tr>
<th>Initial NW D (nm)</th>
<th>Final NW D (nm)</th>
<th>NW Yield HCl:DI</th>
<th>NW Yield HCl:IPA</th>
</tr>
</thead>
<tbody>
<tr>
<td>28</td>
<td>14</td>
<td>100%</td>
<td>100%</td>
</tr>
<tr>
<td>26</td>
<td>12</td>
<td>31%</td>
<td>100%</td>
</tr>
<tr>
<td>24</td>
<td>10</td>
<td>5%</td>
<td>100%</td>
</tr>
<tr>
<td>22</td>
<td>8</td>
<td>0%</td>
<td>97%</td>
</tr>
<tr>
<td>20</td>
<td>6</td>
<td>0%</td>
<td>5%</td>
</tr>
</tbody>
</table>

Table 2-1. Yield of InGaAs VNWs after 7 cycles of DE.
Figure 2-17. InGaAs VNW with 5 nm diameter, 230 nm height obtained after 10 cycles of digital etch in H₂SO₄:methanol.

2.3 Gate Stack

Gate stack engineering constitutes an important component in MOSFET and TFET research. High-κ metal gate, HKMG, has been widely used in the recent generation of CMOS devices. In this thesis, HKMG gate stack on InGaAs nanowires has been investigated, as discussed below.

After DE, the samples are treated in diluted H₂SO₄ in DI water and transferred immediately into the chamber of Cambridge Nanotech ALD system for Al₂O₃ deposition at 250 °C. To confirm the uniformity of the ALD oxide coverage on the high aspect ratio NW sidewall, 80 cycles of Al₂O₃ is deposited on the InGaAs NW in Figure 2-18 (a) and the same NW is captured in Figure 2-18 (b) after ALD deposition. It can be seen that the Al₂O₃ covers the sidewall uniformly and there is no indication of increasing roughness. For device demonstrations a thinner oxide is used.
AJA ATC sputter system is employed next for the deposition of 40 nm W as the gate metal. W is chosen in this process because it can be removed with a high selectivity to FOX in the etch back process. Figure 2-19 presents an InGaAs NW covered with sputtered 40 nm W and reveals that the NW is completely covered in metal. While W is 40 nm thick on the floor, the thickness reduces to 20 nm on the sidewall as the nanowire diameter increases from 40 to 80 nm after metal sputtering, which means that the deposition rate is half of that on a planar surface. The top of the nanowire is no longer flat, which means that the metal thickness is not uniform. The gate stack module completes with a 30 min FGA (8% H$_2$ in N$_2$ ambient) at 350 °C in a furnace.
Although the electrical quality of the oxide/semiconductor interface becomes apparent when analyzing the transistor current-voltage (IV) characteristics, it is important to have a method to evaluate this interface with quick turnaround for optimizing the gate dielectric deposition and monitor the ALD chamber condition. For this purpose, MOS capacitor (MOSCap) is an efficient way of characterizing oxide/semiconductor interface by analyzing the capacitance-voltage (CV) characteristics. The MOSCap fabrication is very simple and begins with lightly doped In$_{0.53}$Ga$_{0.47}$As grown on top of heavily doped InP substrate. The process flow involves a backside contact metal sputtering, front side surface preparation, ALD deposition and metal contact evaporation, which can be finished within several hours. Figure 2-20 displays the CV characteristics of an InGaAs – 2.5 nm Al$_2$O$_3$ MOSCap before and after FGA with identical
surface preparation, gate dielectric deposition and FGA condition to the transistor demonstrations. The FGA clearly reduces the frequency dispersion below threshold significantly, indicating a large drop in the interface trap density ($D_{it}$) in the midgap enabled by FGA.

![Figure 2-20. CV characteristics of InGaAs MOSCAP with 3 nm Al$_2$O$_3$ dielectric: (left) before and (right) after FGA.]

An effective way of improving device drive current is increasing the oxide capacitance by using dielectrics with higher $\kappa$ [59]. As explained earlier, industry-standard HfO$_2$ with higher $\kappa$ than Al$_2$O$_3$ cannot be employed due to process constraints, demanding a high $\kappa$ dielectric that can be selectively removed to FOX in a isotropic way (non-HF). LaAlO$_3$ with dielectric constant as high as 16 has been shown to form superior interfaces on InGaAs [60] and can be removed selectively to FOX in HCl. Lanthanum tris($N,N'$-diisopropylformamidinate), trimethylaluminum, and H$_2$O serve as the precursors for LaAlO$_3$ deposition at 300 °C. This novel oxide is currently being tested and integrated into the process flow$^6$. For gate metal, WN

---

$^6$ In collaboration with Xiabing Lou from Prof. Gordon’s research group at Harvard.

63
can be deposited right after dielectric deposition in the same ALD chamber at 375 °C using \( \text{bis(tert-butylimido)bis(dimethylamido)tungsten(VI)} \) vapor and ammonia gas as the precursors. Alternatively, W can be sputtered in a different chamber in AJA after ALD oxide deposition. MOSCaps with 0.5 nm Al₂O₃/3 nm LaAlO₃ are fabricated with both WN and W as the gate metal, finishing with FGA at 350 °C for 30 min. The CV characteristics are compared in Figure 2-21. The frequency dispersion for W below threshold is much higher than WN, signaling the necessity of in-situ WN deposition right after LaAlO₃.

![Figure 2-21. CV of InGaAs MOSCap with 0.5/3 nm Al₂O₃/LaAlO₃ dielectric after FGA with W (left) and WN (right) gate metal.](image)

### 2.4 Backend Process

#### 2.4.1 Planarization and etch back

With gate stack finished, as shown in Figure 2-19, the gate metal covering the top of the NW needs to be removed and replaced with contact metal. Alignment of the gate edge relative to the channel material with nanometer scale accuracy is critical for both MOSFETs [61] and TFETs [62]. To achieve this goal, a planarization and etch back process module is developed.
using FOX®-16 from Dow Corning. FOX is a class of SOG material that can be deposited by spin-coating and has structural and chemical properties similar to Si oxide upon high temperature annealing or EBL. Two steps of planarization and etch back are done and a detailed process flow is shown in Figure 2-22. In short, FOX is first etched back leaving only the top part to be etched while protecting the gate metal on the sidewall below the top contact. The complicated scheme stems from the limited reproducibility of the etcher and FOX spin coating. Alignment accuracy below 10 nm cannot be achieved with a single timed etch back step. Multiple short-time etching and frequent SEM monitoring in between are needed, as described below.
1st planarization and etch back begins with gate stack covering InGaAs NW in Figure 2-22 (a). $t_w$ is the metal thickness on the top surface and the red line represents the target metal edge, $t_s$ below the semiconductor NW top, determined by the semiconductor heterostructure. FOX is first coated on the wafer to obtain a film of 450 nm thick that is almost twice the NW height for achieving flat surface (Figure 2-22 (b)). Subsequent baking at 200 and 250 °C for 1 min each is performed to reflow the FOX surface for better planarization and drive off the residue solvents.

Next, the FOX is etched back in a single step roughly until the tip of the NW is exposed in (c), with CF$_4$ chemistry in an electron cyclotron resonance (ECR) RIE system Plasmaquest. The total etch time is estimated based on the ratio of $d_1$ and the calibrated etch rate. An SEM image at this stage is shown in Figure 2-23 (a). This step usually can be done in a single etch since accuracy is not critical here. Further FOX etch back for a thickness of $d_2$ in Figure 2-22 (d) is needed to position the FOX layer accurately. $d_2$ is determined by $t_s + t_w - 2*t_u$, where $t_u$ is the undercut shown in Figure 2-22 (c). This undercut is produced when W is removed with a SF$_6$ etching recipe in the same etcher. The metal etch is usually 45 s, leaving an undercut $t_u$ on the order of 10 nm, although with a fairly large uncertainty. As $t_w$ is also not very reproducible, the etch step to remove a thickness of highly uncertain $d_2$ is usually done in multiple short etching steps with each step removing about 20 nm. SEM imaging is taken after each step to measure the etching depth. Figure 2-23 (b) displays an SEM image corresponding to Figure 2-22 (d). W is attacked by the CF$_4$ chemistry during the FOX etch back, but not completely removed. As mentioned above, a short metal etch in SF$_6$ leads to Figure 2-22 (e). Usually at this step the 1st
planarization is not complete yet, because the uncertainty in $t_w$ usually leads to some error. Short FOX and metal etch are added if necessary to adjust the final metal edge. Accounting for the uncertainty in $t_w$, etch rate and FOX spinning, a variation of 10-20 nm of the final FOX surface is usually observed across different devices on a die. Figure 2-22 (f) and Figure 2-23 (c) show the schematic and SEM image at the end of 1st planarization and etch back, respectively. It
is worth noting that the gate dielectric $\text{Al}_2\text{O}_3$ has 1:100 etch selectivity to the $\text{CF}_4$ and $\text{SF}_6$ dry etch compared to FOX and W, so it is expected to cover and protect the semiconductor from direct RIE damage. Figure 2-24 shows a SEM image of a device at this stage in the process (SOG removed).

![30 nm](image)

Figure 2-24. SEM image of a D=30 nm NW MOSFET after 1st planarization and etch back (SOG and dielectric removed in BOE). The intrinsic region and the bottom n+ region of the nanowire are wrapped around by W. The top n+ region is exposed for ohmic contact formation. The roughness on the ground plane is caused by the erosion of W in BOE.

To isolate the gate metal from the contact metal that is sputtered later, the 2nd etch back process starts with another planarization using spin-coated HSQ of 90 nm thickness, shown in Figure 2-22 (g). Since the exposed semiconductor is usually less than 60 nm tall, HSQ is used instead of the much thicker FOX. After planarization, the etch back takes place in multiple short-time FOX dry etch and SEM imaging cycles, until a thin layer of FOX is left in Figure 2-22 (h) and Figure 2-23 (d). The FOX thickness above the gate metal edge, $t_o$, should be minimized to leave as much semiconductor as possible to minimize the contact resistance but large enough to ensure enough isolation. 10 – 15 nm is generally used. In preparation for contact process, the sample goes through a FOX curing step, essentially annealing in a furnace.
at 350 °C in N₂ ambient for 1 hr. The purpose is to increase the etch resistance of FOX to TMAH, as explained in the following section.

### 2.4.2 Contact

Following FOX curing, via process begins with resist pattern with EBL. The FOX etch back recipe with CF₄ is used for via opening. The standard oxide etch recipe involves both CF₄ and H₂ and was tested initially for FOX etching. Although FOX can be etched with CF₄/H₂, residue is left behind that cannot be removed, unless with uncontrollable HF. Figure 2-25 shows a comparison of FOX etched with and without H₂. The FOX residue, seen as white roughness is clearly visible using H₂.

![SEM images of the via bottom after FOX dry etch without (left) and with (right) H₂ recipe. Residue FOX is clearly seen on the right.](image)

After via dry etch, gate dielectric Al₂O₃ is removed with 25% TMAH in water. The FOX curing process mentioned earlier is done to increase the etch resistance of FOX so that it can withstand the TMAH dip. Following dipping in 1:3 HCl:H₂O for 1 min to remove the native oxide, the sample is transferred immediately into the metal deposition chamber. The process considerations for the choice of contact metal are several folds. Firstly, sputtered metal is
needed to cover the sidewall of the top portion of the nanowire. Secondly, since sputtering is not very compatible with lift-off process, metals that can be selectively dry etched with regard to FOX are preferred. Thirdly, a single metal contacting both n and p InGaAs is ideal, as both n and p contacts are needed in TFETs. From a standpoint of device performance, the contact resistivity and metal film resistivity have to be minimized [11] while taking account for process considerations. In this thesis, sputtered Mo is chosen because it satisfies the process considerations mentioned above. Furthermore, as the contact resistance between metal and n+ NW top is revealed to be a bottleneck for VNW MOSFET performance, Mo stands out due to the record contact resistivity below 1 Ω·µm² to n+ InGaAs [63]. After 40 nm Mo is sputtered on the sample, a final lift-off of Ti/Au pads is performed, followed by a Mo etch with Ti/Au as the hard mask, which completes the process.

2.5 Chapter Summary

In this chapter, detailed description of the design and fabrication technology of III-V VNW field-effect transistors are introduced. A novel ICP-RIE technique based on a BCl3/SiCl4/Ar chemistry for fabricating sub-20 nm diameter InGaAs nanowires with smooth, vertical sidewall and high aspect ratio (> 10) is developed for the first time. To mitigate dry-etch damage, RIE is followed by a digital etch method comprised of multiple cycles of self-limiting low power O2 plasma oxidation and diluted H2SO4 rinse. Digital etch improves both the subthreshold swing and peak transconductance of InGaAs MOSFETs, indicating enhanced sidewall interfacial quality, and is capable of controllably thinning nanowire diameter while preserving its general shape. A novel alcohol-based digital etch technique for III-V 3D structures is then introduced. The new technique addresses the limitations of the water-based
approach in enabling structures with sub-10 nm 3D features. Sub-10 nm fins and nanowires with a high yield and mechanical stability have been achieved. InGaAs nanowires with diameter of 5 nm and an aspect ratio > 40 have been demonstrated.

ALD high-κ dielectrics Al₂O₃ and sputted metal is shown to form gate-all-around structures. MOSCaps process is developed to evaluate the oxide/semiconductor interface and confirms the importance of FGA annealing in reducing the interface trap density. A novel oxide LaAlO₃ with higher dielectric constant than Al₂O₃ is being tested and the in situ deposition of gate metal is proved to be required for better interfacial quality.

To isolate the gate metal with the top contact metal, the planarization and etch back based on SOG material is developed. Due to the poor reproducibility of etching tool and SOG spin-coating, a complicated scheme to achieve nanometer accuracy for positioning gate metal edge is established, featuring multiple short-period etching and SEM monitoring in between. The variation of FOX surface is typically within 10-20 nm of the target. Sputtered Mo contact is used to contact both n and p InGaAs.

The processing technology described in this chapter forms the basis for the device demonstrations in Chapter 3 and Chapter 4. Utilizing the fabrication techniques, three generations of InGaAs vertical nanowire MOSFETs are demonstrated and record performance is achieved, as discussed in Chapter 3.
Chapter 3 InGaAs Vertical Nanowire MOSFETs

3.1 Introduction

As discussed in Chapter 1, to continue Moore’s law of MOSFET scaling, it is imperative to find effective ways to enhance carrier transport in scaled dimensions. III-V materials clearly stand out in this regard. The ultimate scalable MOSFET design is the gate-all-around (GAA) nanowire architecture. III-V nanowire MOSFETs come in two different geometries, horizontal and vertical. Horizontal nanowire MOSFETs are essentially FinFETs in which the channel has been suspended through selective etching and a gate is wrapped around its entire periphery [64, 65]. This provides enhanced charge control and an ability to scale to smaller dimensions. Lateral InGaAs nanowires are more frequently prepared through etching, though lateral growth of GaAs and InAs nanowires has been demonstrated by the vapor-liquid-solid technique (VLS) [66]. When the direction of the current flow is aligned vertically, the resulting vertical nanowire (VNW) MOSFET concept is a particularly attractive design because footprint scaling and gate length scaling become uncoupled. This promises high transistor density, stemming from a very small footprint, yet acceptable short-channel effects due to the flexibility in gate length, $L_g$, design. In addition, there is also greater freedom in the selection of contact length and spacer length since the contacts do not contribute to the footprint. This should translate into lower contact resistance and higher performance.

An intriguing aspect of VNW-MOSFETs is that they offer a plausible path for integration on a Si substrate. Bottom-up growth of InGaAs nanowires on Si substrates is relatively well established through vapor-liquid-solid (VLS) epitaxy, selective-area epitaxy
(SAE) and templated-assisted growth techniques [66-68]. Schematics and SEM images of III-V NWs by these three methods are shown in Figure 3-1, reproduced from [68]. Numerous impressive VNW-MOSFET demonstrations have been published with NWs grown using either of these techniques. SEM images of a few examples [69-71] are put together in Figure 3-2.

Figure 3-1. Schematics (left column) and SEM images (right column) of III-V VNWs grown on Si substrate by VLS, SAE and template-assisted techniques [68].
InGaAs VNW MOSFETs were demonstrated through a top-down approach for the first time as part of this thesis work. In this chapter, the DC characteristics of three generations of devices are reviewed first, followed by performance benchmarking against other published III-V VNW results. The device characteristics are then analyzed from the perspective of both electrostatics and ON-state performance, revealing the importance of interface states and top contact resistance. Source/drain asymmetry inherent to VNW MOSFETs is experimentally studied.

### 3.2 Device Electrical Characteristics

#### 3.2.1 First generation

Figure 3-3 shows the design parameters of the first generation (G1) InGaAs VNW MOSFETs [72] and the heterostructure, which is identical to the MOSFET structure in Figure 2-2. Device fabrication followed the process flow described in Chapter 2. 4.5 nm ALD Al₂O₃ (EOT=2.2 nm) was deposited on the sidewalls. All devices have a single NW and a channel length of 80 nm, set by the undoped InGaAs layer thickness. 10 cycles of digital etch were employed to trim the nanowire diameter (D) by 20 nm and improve the quality of the
semiconductor/oxide interface. InGaAs NW used in this work is shown in Figure 3-4, which has much more tapered sidewall compared to the NW in Figure 2-7. This is because the RIE recipe was not optimized at the time of the first device demonstration.

Starting heterostructure:  Design parameters:

- n+ InGaAs, 70 nm  - D= 30, 40, 50 nm
- i InGaAs, 80 nm  - L_{ch}= 80 nm
- n+ InGaAs, 300 nm  - EOT= 2.2 nm
- No. of wires = 1

Figure 3-3. Starting heterostructure and first generation device design parameters.

Figure 3-4. D = 30 nm InGaAs NW after 10 cycles of digital etch used for first generation device demonstration.

Figure 3-5 and 3-6 show the electrical characteristics of a single D=30 nm NW MOSFET. As is common, drain current (I_d) and transconductance (g_m) are normalized by the nanowire circumference (\pi D). The output characteristics demonstrate excellent saturation at low V_{ds} with ON resistance (R_{on}) = 759 \Omega \cdot m (8062 \Omega) at V_{gs}=1 V. A peak g_m (g_{m,pk}) of 280 \mu S/\mu m is extracted at V_{ds}=0.5 V. Subthreshold swing of 145 mV/dec at 0.05 V and 200 mV/dec at 0.5 V.
V are obtained. DIBL is 195 mV/V. The gate leakage current is below $10^{-9}$ A throughout the measurement range. Fluctuations in the drain current measurement are observed, which can be attributed to charge trapping in the very few oxide and interface defects that are present in these extremely small single NW devices [73].

![Figure 3-5](image_url)

Figure 3-5. Output characteristics of a D = 30 nm InGaAs single NW MOSFET.

![Figure 3-6](image_url)

Figure 3-6. Subthreshold characteristics (left), transfer and $g_m$ characteristics at $V_{ds} = 0.5$ V (right) of the device shown in Figure 3-5.

Electrical properties of a single nanowire device with D = 50 nm are shown in Figure 3-7 and 3-8. The electrostatics are worse than D=30 nm device, as evidenced by increased S (210
mV/dec at $V_{ds} = 0.05$ V) and DIBL (360 mV/V). However, $R_{on}$ is reduced by a factor of 4 and a high $g_{m, pk}$ of 730 $\mu$S/\mu m at $V_{ds} = 0.5$ V is achieved. $R_{on}$ is believed to be dominated by contact resistance of the top n` region due to the small contact area. The level of current fluctuation is also reduced as the nanowire diameter widens.

![Figure 3-7. Output characteristics of a D = 50 nm InGaAs single NW MOSFET.](image)

The impact of digital etch can be appreciated in Figure 3-9 that shows the subthreshold characteristics of $D = 40$ nm devices with and without digital etch. The starting NW diameter is different so that after digital etch of one of them, both final diameters are identical. The digital
etch improves the interface quality as evidenced by a reduction in $S$ from 220 mV/dec to 180 mV/dec. Interestingly, DIBL is not affected. However, the threshold voltage $V_t$ (defined at 1 $\mu$A/$\mu$m at $V_{ds}$=0.05 V) shifts positive by about 0.1 V as a result of digital etch. This could be due to the elimination of B$^+$ ions incorporated at the sidewalls during the RIE process [41].

Digital etch also improves transport, as evidenced by a 25% increase in $g_{m,pk}$ (from 402 to 498 $\mu$S/$\mu$m at $V_{ds}$=0.5 V). Figure 3-10 shows that gate leakage also significantly decreases by digital etch. Output characteristics of $D = 40$ nm without and with digital etch are compared in Figure 3-11.

Figure 3-9. Impact of digital etch on the subthreshold characteristics of $D = 40$ nm devices. This is the final device diameter.

Figure 3-12 shows the evolution of key figures of merit as a function of final NW diameter for devices fabricated with and without digital etch. S and DIBL greatly improve as D is reduced, as expected. Peak normalized $g_m$ is reduced as D is decreased. This is probably due to the strong dependence of series resistance on D that is evident in Figure 3-12 (d). Thus, a trade-off between transport (best indicated by $g_{m,pk}$) and electrostatics (measured by S) is
observed in vertical InGaAs NW MOSFETs of different diameters. Digital etch improves $S$ and $g_m$ but does not affect DIBL and $R_{on}$.

Figure 3-10. Impact of digital etch on gate leakage current of $D = 40$ nm single NW MOSFETs.

Figure 3-11. Output characteristics of $D = 40$ nm InGaAs single NW MOSFETs fabricated without (left) and with (right) digital etch.
Figure 3-12. Key device figures of merit vs. final nanowire diameter for devices fabricated with and without digital etch (over 10 devices for each nanowire diameter are measured): (a) Linear ($V_{ds} = 0.05 \text{ V}$) and saturated ($V_{ds} = 0.5 \text{ V}$) subthreshold swing, (b) DIBL, (c) Normalized $g_m$ ($V_{ds} = 0.5 \text{ V}$), and (d) Normalized $R_{on}$.

3.2.2 Second generation

The starting heterostructure of the second generation (G2), shown in Figure 3-13 was designed to closely resemble the previous generation, albeit with a critical difference. The wafer essentially consists of an 80 nm undoped In$_{0.53}$Ga$_{0.47}$As channel sandwiched between two $n^+$ contact regions, same with the first generation. The major difference is the design of the contact region, now featuring the insertion of $n^+$ 7 nm In$_{0.53}$Ga$_{0.47}$As/ 2 nm InAs/ 6 nm In$_{0.7}$Ga$_{0.3}$As on top of 55 nm In$_{0.53}$Ga$_{0.47}$As, as opposed to 70 nm $n^+$ In$_{0.53}$Ga$_{0.47}$As in G1. The insertion of the
InAs/In$_{0.7}$Ga$_{0.3}$As layer is to reduce the contact resistivity as InGaAs is reported to form better ohmic contacts with higher InAs concentration [74].

**Starting heterostructure:**

- n+ InGaAs: 7 nm
- n+ InAs: 2 nm
- n+ In$_{0.7}$Ga$_{0.3}$As: 6 nm
- n+ InGaAs, 55 nm
- i InGaAs, 80 nm
- n+ InGaAs, 300 nm
- i InP substrate

**Design parameters:**

- D= 20, 30, 40 nm
- L$_{ch}$= 80 nm
- EOT= 1.5 nm
- No. of wires = 1/100

Figure 3-13. Starting heterostructure of 2\textsuperscript{nd} generation devices and design parameters.

Device fabrication follows the process flow in Chapter 2 but it incorporated a number of innovations compared to the first generation. Nanowire etching was done through optimized InGaAs dry etch technology that yields nearly vertical and smooth sidewalls and a very uniform NW diameter, as shown in Figure 2-7. This was obtained by increasing the substrate temperature during etch and optimizing the etching conditions (gas flow ratio, etc) after the demonstration of the first generation device. In addition, this sample featured a scaled gate oxide consisting of 3 nm of Al$_2$O$_3$ (~1.5 nm EOT) and greater attention to ALD conditioning to improve the quality of the oxide semiconductor interface. DE were employed to trim the
nanowire diameter by 20 nm. The final device has the same channel length \( L_{ch} \) of 80 nm and diameters of 40 / 30 / 20 nm.

The output, subthreshold and \( g_m \) characteristics of an exemplary single NW device with 40 nm diameter are shown in Figure 3-14 and 3-15. The linear \( (V_{ds} = 0.05 \, \text{V}) \) and saturated \( (V_{ds} = 0.5 \, \text{V}) \) subthreshold swing \( (S_{\text{linear}} \, \text{and} \, S_{\text{sat}}) \) are 98 and 114 mV/dec, while DIBL is 177 mV/V. The electrostatics are greatly improved compared to G1, clearly shown in the benchmark figure in the next section, attributing to the improved oxide/semiconductor interface by better ALD chamber conditioning. A peak transconductance \( (g_{m,pk}) \) of 620 \( \mu \text{S}/\mu \text{m} \) is achieved with an ON resistance of 895 \( \Omega \cdot \mu \text{m} \). \( g_{m,pk} \) here is comparable to G1, despite a scaled EOT (1.5 vs. 2.2 nm). This is mainly due to the increased resistance of the top contact in the present devices, unexpected from an effort to improve the top contact by new heterostructure design. The cause of this is hypothesized to be a lower doping of the \( n^+ \) contact layers than designed caused by reduced cracking of the \( \text{Si}_2\text{H}_6 \) precursor at the reduced epitaxial growth temperature of 450 °C.

\[
\begin{align*}
V_{gs} = & -0.2 \, \text{V to 0.7 \, V in 0.1 \, V step} \\
V_{ds} (\text{V}) & \quad 0.0 \quad 0.1 \quad 0.2 \quad 0.3 \quad 0.4 \quad 0.5 \\
I_d (\mu \text{A}/\mu \text{m}) & \quad 350 \quad 300 \quad 250 \quad 200 \quad 150 \quad 100 \quad 50 \quad 0 \\
\end{align*}
\]

Figure 3-14. Output characteristics of a D = 40 nm InGaAs single NW MOSFET from G2.
Figure 3-15. Subthreshold and $g_m$ (inset, $V_{ds} = 0.5$ V) characteristics of the device shown in Figure 3-14.

The smallest diameter found in devices with switching characteristics is 20 nm in this sample. The subthreshold characteristics of a representative device are shown on the right in Figure 3-16. However, the ON current is very low at both drain biases and the difference in ON current is about 2 orders of magnitude, signaling the behavior of a Schottky contact. This is verified by the super-linear turn-on behavior in the device output characteristics shown on the right of Figure 3-16. The bottom metal/semiconductor contact is ohmic because the IV characteristics between the bottom pads in two adjacent devices are linear within 0.3 V bias range, which leads to the hypothesis that the top semiconductor/metal contact is Schottky. This would happen because the tiny NW tip with diameter of 20 nm might be fully depleted by the Schottky barrier height around the NW as a result of the large width of the depletion region caused by accidental low doping during growth. To alleviate this problem, higher doping level,
thicker NW diameter in the contact region and/or contact metal with lower Schottky barrier are needed.

Figure 3-16. Subthreshold (left) and output (right) characteristics of a D = 20 nm InGaAs single NW MOSFET from G2.

Figure 3-17 compares the statistics of major figure of merits (FOMs) related to device electrostatics for D = 40 and 30 nm devices, including $S_{lin}$ and $S_{sat}$, DIBL. It is revealed that there is no statistical difference in any of these FOMs between 30 and 40 nm devices. The similar subthreshold swing and DIBL are expected based on standard theory of electrostatics, which will be discussed in section 3.4 However, Figure 3-17 also points out that 30 nm devices have smaller $g_{m, pk}$, likely to be caused by the larger top contact resistance due to smaller contact area and thinner NW diameter, evidenced by the larger $R_{on}$, also shown in Figure 3-17.

Due to the improved yield in G2, devices containing 100 NWs (array NW) placed in a square pattern (10 row $\times$ 10 column) with 500 nm spacing are found to be working. The output, subthreshold and $g_m$ characteristics of an exemplary array NW device with 40 nm diameter are shown in Figure 3-18 and 3-19. However, the array NW (ANW) devices have worse electrostatics and transport in comparison to single NW (SNW) devices, as shown in Figure 3-
where the statistics of $S_{\text{lin}}$, $S_{\text{sat}}$, and $g_{m,pk}$ are compared. The degradation of electrostatics can be explained by the smearing due to the variability among individual NWs in parameters such as diameters, planarization height, etc [75]. The reduction in peak transconductance stems from the increased effect of the extrinsic resistance, including contact and part of the access.
resistance from the bottom contact [76]. For SNW and ANW devices, this extrinsic resistance on the bottom is roughly unchanged while the channel resistance scales inversely proportional to the number of NWs. As a result, the voltage drop on the extrinsic resistance is larger in ANW devices, reducing the intrinsic gate overdrive voltage.

Figure 3-18. Output characteristics of a D = 40 nm InGaAs array NW MOSFET.

Figure 3-19. Subthreshold characteristics (left) and $g_m$ characteristics at $V_{ds} = 0.5$ V (right) of the device shown in Figure 3-18.
Figure 3-20. Statistical comparison of single and array nanowire devices with 40 nm diameter in key FOMs including $S_{lin}$, $S_{sat}$, and $g_{m,pk}$.

Another interesting observation can be made by comparing the level of fluctuation in the IV characteristics of single and array NW devices, e.g., Figure 3-18 vs. 3-14. It is clear that the family of curves for ANW is much smoother than SNW, meaning a smaller noise. To gain quantitative information on the noise, 10 min sampling measurements were performed on both SNW and ANW devices at identical bias, $V_{ds} = 0.25 \, \text{V}$, $V_{gs} = 0.3 \, \text{V}$ (ON state), as shown in Figure 3-21. Figure 3-22 presents the noise spectral density (PSD) normalized by the square of current as a function of frequency for both SNW and ANW devices, a common measure of noise [76] extracted using sampling measurement results via Welch’s PSD estimate
implemented in Matlab\textsuperscript{7} with the correction for current drift. The red dotted line serves as a visual aid with 1/f frequency dependence. The normalized PSD in both ANW and SNW follows 1/f noise. ANW has roughly 100 times less noise power compared to SNW, corresponding to the number of NWs. The reduction of normalized noise by the number of NWs can be explained with simply statistical theory as follows. Suppose the current of a single NW is comprised of a noise source $\varepsilon$ and a static current $I_0$:

$$I_s = I_0 + \varepsilon$$ \hspace{1cm} (3-1)

The variance of the current is a measure of the noise power:

$$Var(I_s) = \varepsilon^2$$ \hspace{1cm} (3-2)

The current of an ANW device is the sum of the currents from $n$ individual NWs and its variance is given by:

$$I_A = \sum_{i=1}^{n} I_{si}, Var(I_A) = n\varepsilon^2$$ \hspace{1cm} (3-3)

The variance is only $n\varepsilon^2$ instead of $n^2\varepsilon^2$, because the noise components in different NWs are uncorrelated so the covariance is zero. As a result, the normalized power of ANW device is $n$ times smaller than that of SNW:

$$\frac{Var(I_A)}{I_A^2} = \frac{n\varepsilon^2}{n^2I_0^2} = \frac{\varepsilon^2}{nI_0^2} = \frac{1}{n} \times \frac{Var(I_s)}{I_s^2}$$ \hspace{1cm} (3-4)

This result also means that devices with larger area have less noise.

\textsuperscript{7} https://www.mathworks.com/help/signal/ref/pwelch.html
Figure 3-21. 600 s sampling measurements in drain current as a function of time with 50 ms resolution, in both SNW (left) and ANW (right) devices.

Figure 3-22. Normalized power spectral density as a function of frequency for ANW and SNW devices. The red dotted line serves as a reference with 1/f dependence on the frequency.

In order to improve the contacts, rapid thermal anneal (RTA) after the completion of device fabrication was applied to G2, in N₂ ambient at 350 °C for 3 min. Figure 3-23 and 3-24 shows the effect of RTA on the subthreshold, $g_m$ and output characteristics of an exemplary D = 40 nm device. The $g_{m,pk}$ is improved from 455 to 502 $\mu$S/µm while $R_{on}$ stays roughly the same at
1030 Ω·μm, which means that the $g_m$ increase comes from an improvement in the gate stack, not contacts. The huge improvement in $S_{\text{linear}}$ from 105 to 65 mV/dec confirms the better quality in semiconductor/oxide interface. DIBL is also reduced from 155 to 88 mV/V. A large positive threshold voltage shift of 0.33 V is observed. Electrical characteristics of one of the best device (D = 40 nm) are shown in Figure 3-25. $S_{\text{linear}} = 70$ mV/dec, $S_{\text{sat}} = 80$ mV/dec and $g_{m,pk} = 720$ μS/μm at $V_{ds} = 0.5$ V are obtained.

Figure 3-23. Impact of RTA on subthreshold (left) and transconductance (right) characteristics of an exemplary D = 40 nm device.

Figure 3-24. Output characteristics of an exemplary D = 40 nm device before (black lines) and after (red lines) RTA.
3.2.3 Third generation

The heterostructure of G3, shown in Figure 3-26 was designed to be essentially the same with G2 but correcting the growth conditions and increasing the doping in $\text{In}_{0.7}\text{Ga}_{0.3}\text{As/InAs}$ from $3 \times 10^{19}$ with Si to $7 \times 10^{19}$ with Te to lower the top contact resistance. Another minor change was to increase the thickness of the top-most layer ($\text{In}_{0.53}\text{Ga}_{0.47}\text{As}$) from 7 nm to 11 nm. The purpose was to leave InAs layer to directly contact the metal because the 10 nm will be removed by DE during processing. The major process change was the shift from aqueous DE to
solvent based DE (10% H2SO4 in methanol), in order to demonstrate devices with sub-20 nm diameter. The nanowire diameters were trimmed by 20 nm with DE. 2 nm Al2O3 was deposited and final RTA in N2 ambient at 350 °C for 3 min was performed.

**Starting heterostructure:**

- n+ InGaAs: 11 nm
- n+ InAs: 2 nm
- n+ In0.7Ga0.3As: 6 nm
- n+ InGaAs, 55 nm
- i InGaAs, 80 nm
- n+ InGaAs, 300 nm
- i InP substrate

**Design parameters:**

- D= 40, 30, 20, 17, 14 nm
- Lch = 80 nm
- EOT= 1 nm
- No. of wires = 1/100

Figure 3-26. Starting heterostructure of 3rd generation devices and the design parameters.

The electrical characteristics of one of the best devices (SWN, D = 40 nm) are shown in Figure 3-27. Unfortunately, the output characteristics reveal a Schottky contact even in the 40 nm devices, possibly due to processing related issues such as RIE damage to the nanowire top or incompletely removal of gate dielectric. Despite the Schottky contact, $S_{\text{linear}} = 82 \text{ mV/dec}$, $S_{\text{sat}} = 80 \text{ mV/dec}$ and $g_{m,pk} = 775 \mu \text{S/μm}$ at $V_{ds} = 0.5 \text{ V}$ are obtained. The minimum subthreshold swing of 70 mV/dec at $V_{ds} = 50 \text{ mV}$ is observed in one device in Figure 3-28 (a), confirming that the solvent DE technique delivers sidewall quality comparable to the aqueous approach.
Switching characteristics are observed in devices with 14 nm diameter, shown in Figure 3-28 (b). However, the device performance is severely degraded by the Schottky top contact.

Figure 3-27. Subthreshold, $g_m$ and output characteristics of one of the best devices in G3 (D = 40 nm) after RTA.

Figure 3-28. (a) Subthreshold characteristics of a D = 40 nm with $S_{\text{linear}} = 70 \text{ mV/dec}$ (b) Subthreshold characteristics of a D = 14 nm showing switching behavior at $V_{ds} = 0.5 \text{ V}$.
3.2.4 Benchmark

Figure 3-29 benchmarks the transport and electrostatics in recently published III-V VNW MOSFETs [69-71, 77-80] by plotting $g_{m,pk}$ vs. minimum $S_{sat}$, both at $V_{ds} = 0.5$ V.

![Figure 3-29. Benchmark of $g_{m,pk}$ vs. minimum $S$ at $V_{ds} = 0.5$ V for recently published InGaAs and InAs VNW MOSFETs. A trade-off is evident in terms of electrostatics and transport.](image)

A trade-off between transport and short-channel effects in all VNW MOSFETs is observed clearly. G2 and G3 devices after RTA demonstrate a significant improvement in the electrostatics compared to G1, mainly due to the scaled gate oxide and the improved oxide/semiconductor interface. To the best knowledge, G3 achieves the highest quality factor $Q$ (defined as $g_{m,pk}/S_{sat}$ at $V_{ds} = 0.5$ V) of 9.7 in any III-V VNW MOSFETs by either bottom-up or top down approach, despite the Schottky contact. $I_{on}$ of 224 $\mu$A/$\mu$m is obtained at $I_{off} = 100$
nA/μm with $V_{dd} = 0.5$ V, which also sets the record of III-V VNW MOSFETs. These results highlight the potential of top-down III-V VNW transistor technology developed in this thesis.

### 3.3 Electrostatics

The major advantage of NW geometry is the excellent scalability. To gain a perspective of the electrostatics of the III-V VNW MOSFETs, the experimental data from published results are compiled and compared with the standard theory of the electrostatics of NW devices [25, 81]. In essence, when the materials (channel, gate oxide), material properties (channel material and gate oxide dielectric constants, $\varepsilon_s$ and $\varepsilon_{ox}$) and geometry (gate length $L_g$, diameter $D$ and oxide thickness $t_{ox}$) of a NW MOSFET is given, the ideal $S_{linear}$ and DIBL can be estimated based on the ratio of $L_g$ and the natural length $\lambda$. This ratio is a measure of how far into the channel the drain impacts the surface potential, and the SCEs are negligible if $L_g/\lambda$ is larger than 6. For the Gate-All-Around NW geometry, $\lambda$ is given by [81]:

$$\lambda = \sqrt{\frac{\pi \varepsilon_s D^2}{4C_{ox}}} + \frac{D^2}{16}$$  \hspace{1cm} (3-5)

where $C_{ox}$ is the total capacitance of the oxide and barrier semiconductor layers:

$$\frac{1}{C_{ox}} = \frac{1}{C_{ox,1}} + \frac{1}{C_{ox,2}} + \ldots = \frac{\ln\left(1 + \frac{2t_{ox,1}}{D}\right)}{2\pi\varepsilon_{ox,1}} + \frac{\ln\left(1 + \frac{2t_{ox,1}}{D + 2t_{ox,1}}\right)}{2\pi\varepsilon_{ox,2}} + \ldots$$ \hspace{1cm} (3-6)

Figure 3-30 (a) compiles the $S_{linear}$ or $S_{sat}$ (if $S_{lin}$ is not available) of experimental III-V VNW MOSFETs [69-71, 77, 79, 80, 82-87] as a function of $L_g/\lambda$, where $\lambda$ was calculated according to Equation 3-5 and 3-6. Relative dielectric constants of 8, 18 and 16 were used for
ALD Al$_2$O$_3$, HfO$_2$\textsuperscript{8}, and LaAlO$_3$ [88]. InAs, In$_{0.7}$Ga$_{0.3}$As, In$_{0.53}$Ga$_3$As, and InP have relative dielectric constants of 15.2, 14.3, 13.9 and 12.5\textsuperscript{9}. The possible interfacial oxide layer [89] was not taken into account due to lack of specific information. The red line is the ideal relationship obtained by electrostatic simulation [81]. Although a few experimental devices including G2 and G3 after RTA yield values close to the theoretical expectations, most of the results, e. g. G1, have $S_{\text{linear}}$ far larger than the simulated values. The dominating factor to explain the discrepancy is the poor gate efficiency $\eta$, which measures the effectiveness of the gate voltage in modulating the channel potential $\Phi_x$. The change of charge in the depletion layers, channel and interface states as a function the channel potential are differential capacitors ($C_n$, $C_d$ and $C_{it}$) that form a voltage divider with the gate oxide capacitor $C_{ox}$, shown schematically in the inset of Figure 3-30 (a). Therefore, the change in the channel potential is only a fraction of the change in the gate voltage and the gate efficiency is always less than 1. The most significant factor in the subthreshold region is $C_{it}$, which is proportional to the interface trap density $D_{it}$.

When taking into consideration of gate efficiency, $S_{\text{linear}}$ becomes:

$$S_{\text{linear}} = S_0 \times \eta = S_0 \times \left(1 + \frac{C_n + C_d + C_{it}}{C_{ox}}\right) = S_0 \times (1 + \frac{q\pi d D_{it}}{C_{ox}}) \quad (3-7)$$

$S_0$ is the simulated subthreshold swing accounting for 2D effects. As the channel is undoped, $C_d$ is very small. $C_n$ is also very small in the subthreshold region because the charge concentration in the channel is negligible. According to Equation 3-7, the $D_{it}$ levels for G1 to G3 were estimated to be $1.6 \times 10^{13}, 1.4 \times 10^{12}, 3.9 \times 10^{12} \text{ eV}^{-1} \text{ cm}^{-2}$, respectively. Although CV

\textsuperscript{8} Dielectric constants of ALD Al$_2$O$_3$ and HfO$_2$ are determined from CV characteristics of silicon MOSCap.

\textsuperscript{9} Dielectric constants of semiconductors are obtained from http://www.ioffe.ru/SVA/NSM/Semicond/

96
Figure 3-30. (a) $S_{\text{linear}}$ plotted versus $L_g/\lambda$ in recently published III-V VNW MOSFETs (symbols). The red line is the simulation [81]. The inset shows the degradation of gate efficiency $\eta$ due to the capacitance from channel charge, depletion region and interface traps ($C_n$, $C_d$ and $C_{it}$). (b) DIBL as a function of $L_g/\lambda$. The symbols are consistent with (a). The red line is the simulation from [81].
characteristics are more commonly used for the more precise extraction of $D_{it}$ as a function of trap state energy, the single NW devices present a significant challenge for CV due to the extremely small gate area and large parasitic capacitances. Compared to G1, G2 reduces the $D_{it}$ by one order of magnitude, due to improved ALD chamber condition. However, significant efforts to further improve the interfacial quality needs to be devoted to achieve the $10^{11} \text{eV}^{-1}\text{cm}^{-2}$ $D_{it}$ level demonstrated in planar InGaAs/ ALD Al$_2$O$_3$ structures [90]. Figure 3-30 (b) plots the DIBL as a function of $L_G/\lambda$. Although $S_{\text{linear}}$ in G2 and G3 is close to ideal, DIBL is still significantly worse than the theoretical prediction, which could also be caused by interface states [23].

3.4 Source/Drain Asymmetry

Two generations of samples (G1 and G2) that yield many working devices with relatively well behaved electrical characteristics have allowed us to study the asymmetry of VNW InGaAs MOSFETs, an issue of great importance to circuit designers. This is a topic that has been investigated theoretically [61] but has not been studied experimentally in the InGaAs system. VNW transistors are intrinsically asymmetric to source-drain (S/D) swapping, in contrast with intrinsically symmetric horizontal devices. We carried out a study of device asymmetry and its impact on electrical FOMs by performing device electrical measurements with source and drain swapped. The characterization procedure was benign and the device characteristics remained stable after repeated measurements. The electrical characteristics for G2 devices used in this section are obtained before final RTA step.

A typical example of the impact of S/D swap for a G2 device is shown in Figure 3-31 and 3-32, where measurements with bottom electrode as the source (BES, black) and top
electrode as the source (TES, red) are plotted. We see that the subthreshold behavior barely changes while the ON current at $V_{ds}=0.5$ V is greatly affected. The asymmetry in the ON regime can be quantified by the change in peak transconductance which is 620 $\mu$S/\(\mu\)m for BES vs. 200 $\mu$S/\(\mu\)m for TES (inset of Figure 3-31). This strong asymmetry can be attributed to the significant difference in series resistance at the bottom and at the top of the nanowire. In our NW devices, the contact to the top electrode only covers the pillar tip, a very small area. The contact to the bottom electrode, on the other hand, takes place on the bottom conductive plane and has a much larger contact area. As a result, the top contact presents much more resistance than the bottom contact. In consequence, the TES configuration offers a small drain resistance $R_d$ and a large source resistance $R_s$, while BES gives a small $R_s$ and large $R_d$. It is well established that $R_s$ has a much larger effect on extrinsic $g_m$ than $R_d$ [91], and this results in significantly higher $g_m$ for BES with respect to TES.

Figure 3-31. Subthreshold characteristics of an exemplary device in G2 measured with bottom electrode at the source (BES, black) vs. top electrode as the source (TES, red) (inset shows transconductance characteristics).
Figure 3-33 compares the output characteristics of the same device of Figure 3-31 in both configurations. As expected, $R_{on}$ is roughly the same in both configurations (895/897 Ω·μm for BES/TES). However, in addition to the difference in maximum current already discussed, TES enables better saturation than BES. This is a consequence of the smaller $R_d$ under TES.

In an FET, the intrinsic drain-to-gate voltage $V_{dg,i}$ establishes the saturation of the device output characteristics. For example, in a classic long channel MOSFET, drain current saturation happens around $V_{dg,i} = -V_t$. In the presence of series drain and source resistance, $V_{dg,i}$ differs from the extrinsic drain-to-gate voltage $V_{dg}$ and is given by:

$$V_{dg,i} = V_{dg} - I_d \times R_d$$ (3-8)

This means that, to the first order, the drain resistance contributes to the reduction of $V_{dg,i}$ but the source resistance does not. In consequence, high $R_d$ hampers device saturation while high $R_s$ does not. In our devices in the TES configuration, $R_s \gg R_d$ while in the BES configuration, $R_d \gg R_s$. As a result, the TES configuration exhibits better current saturation than the BES configuration.
Figure 3-33 (a) compares the behavior of ON regime figures of merit, transconductance and $R_{on}$, in the BES and TES configurations for G1 and G2 devices. As is reasonable to expect,

![Figure 3-33](image)

there is no asymmetry in $R_{on}$. We can see, however, that there is strong asymmetry in $g_{m, pk}$ in this work. Also, as mentioned before, $g_{m, pk}$ for G2 is comparable to G1 in the BES configuration, despite a scaled EOT (1.5 vs. 2.2 nm). This is mainly due to the increased resistance of the top
contact in the present devices which also contributes to the greater $g_m$ asymmetry observed in this work and the sharp increase in $R_{on}$.

In addition to asymmetry in the ON regime, there is also asymmetry in the subthreshold characteristics to consider. This was very prominent in G1 but has been eliminated in G2 devices. This is shown in Figure 3-33 (b). This graph plots key FOM values pertaining to subthreshold characteristics for G1 and G2 devices. For DIBL and $S_{sat}$, G1 devices displayed significant asymmetry. As discussed theoretically in [61], DIBL and $S_{sat}$ are primarily determined by the nanowire diameter at the source which is about 15 nm wider in the BES configuration with respect to the TES configuration in G1 devices, determined from the shape of the NW shown on the right in Figure 3-34. The improved dry etch technique in G2 leads to a much more uniform NW cross-section, as shown on the left in Figure 3-34. This results in the suppression of DIBL and $S_{sat}$ asymmetry. Also, consistent with simulations in [61], $S_{linear}$ does not reveal significant asymmetry in either chip.

Figure 3-34. Improved NW profile in G2 (left) compared to G1 (right).
3.5 Roadblock: Top Contact

Future generations of MOSFET technologies will require a series resistance below 100 Ω·μm [11]. For NW MOSFETs to maintain acceptable SCEs, sub-10 nm NWs are needed [26]. In contrast, G1 VNW MOSFETs have \( R_{on} \) of 400 Ω·μm in D = 50 nm devices and this number rises up to 900 Ω·μm in devices with diameters scaled to 30 nm. In G2, 20 nm devices even show Schottky contact (probably due to the large depletion width underneath the contact as a result of the accidental low doping level). Even with doping as high as \( 3 \times 10^{19} \text{ cm}^{-3} \), the depletion width underneath the Mo contact in InGaAs still approaches 5 nm [92], which means that NWs with sub-10 nm diameter could be fully-depleted under the contacts. Figure 3-35 plots the \( g_{m, pk} \) at \( V_{ds} = 0.5 \text{ V} \) as a function of \( 1/R_{on} \) for both G1 and G2 devices, revealing a clear negative relationship between the transconductance and on resistance. The transconductance is significantly degraded by the series resistance. All these difficulties stem from the tiny area available on top of the ultra-thin NW for contacts (can be seen in Figure 3-1), a major roadblock for VNW transistors that necessitates substantial efforts to be invested to push forward this technology.

Another challenge for III-V VNW transistor technology is the lack of convenient methods to quantify the series resistances. Unlike planar technology where series resistances are symmetric and can be determined from transfer length measurement (TLM) methods by varying the gate length, the III-V VNW transistors have asymmetric source/drain series resistances and fixed gate length. In order to calculate the source and drain series resistances from a single device, reciprocal transconductance method (RTM) [93] and drain current conductance method (DCCM) [94] were employed. In essence, RTM method leverages the asymmetric source/drain
terminals discussed in the previous section and the difference in the source and drain series resistance can be calculated as:

\[ R_T - R_B = \frac{1}{g_{m,T}} - \frac{1}{g_{m,B}} \]  

(3-9)

\( R_T \) and \( R_B \) are the source resistance in the TES and BES configurations, respectively. \( g_{m,T} \) and \( g_{m,B} \) are the peak transconductances measured with TES and BES configurations. When the \( g_{m,pk} \) at 0.5 V is used to analyze the G2 and G1 devices, the values of \( R_T - R_B \) for all devices are consistently larger than \( R_{on} \), which is supposed to be bigger than \( R_T + R_B \). This means that the simple model in Equation 3-9 does not apply to these devices, probably due to the nonlinearity in access resistances [20].

![Figure 3-35. \( g_{m,pk} \) plotted as a function of 1/\( R_{on} \) for G1 (black) and G2 (blue) devices](image)

The DCCM method utilizes the asymmetric drain current and drain conductance in linear region and extract the series source and drain resistances at high gate voltages. Again, \( R_T - R_B \) are extracted to be larger than \( R_{on} \). Novel methodology needs to be developed.
quantify and breakdown the components of the series resistances in a device to gain insight on how to reduce the parasitic resistances below the target.

3.6 Chapter Summary

In this chapter, the first demonstration of III-V VNW MOSFETs by top-down approach is presented. As the diameter of NW devices are scaled from 50 to 30 nm, the electrostatics are improved drastically at the cost of a significant drop in transconductance, caused by increasing series resistance. Digital etch improves both the subthreshold characteristics and the transport properties. An enhancement of 20% in peak $g_m$ is observed.

The electrostatics are greatly improved in second generation devices by better ALD chamber conditioning and oxide scaling. Due to the improved yield, array NW devices are found to be working, although with worse electrostatics and transport compared to single NW devices. The degradation in electrostatics is attributed to device variability in diameters, planarization height, etc. The increased effect of the extrinsic resistance explains the reduction in the peak transconductance. Sampling measurements and subsequent Fourier transform reveals the lower noise power density in ANW devices due to the averaging effect.

With the introduction of solvent based digital etch, third generation devices present excellent subthreshold characteristics, signaling great oxide/semiconductor interface. Switching behavior is observed in devices with diameter as low as 14 nm, enabled by the improved structural integrity due to the low surface tension of solvents. G3 devices demonstrate the highest quality factor and $I_{on}$ in any III-V VNW MOSFETs, highlighting the potential of the top-down approach.
Electrostatics of three device generations are benchmarked with standard scaling theory based on the ratio of gate and natural length. G2 and G3 achieve near ideal subthreshold characteristics and a tenfold reduction in the interface trap density, owing to the drastically improved oxide/semiconductor interface.

Two generations of high-yield devices enable the first experimental study of source/drain asymmetry in III-V VNW MOSFETs. The transconductance differs significantly when swapping source and drain due to inherently different top and bottom contacts, which also result in distinct asymmetry in the saturation behavior of the output characteristics. On the other hand, diameter uniformity along the nanowire length is responsible for asymmetry in the subthreshold characteristics. A uniform nanowire cross-section, enabled by the improved InGaAs dry etch technology G2, eliminates the asymmetry of the electrostatics that was observed in G1.

The clear inverse dependence of the device performance on the series resistance reveals the metal contact to the NW tip as the major roadblock to III-V VNW MOSFET technology. Sub-10 nm NWs needed for future technology nodes can potentially be fully depleted under metal contacts. Conventional series resistance extraction methods using single device do not readily apply to the VNW MOSFETs demonstrated in this work, necessitating innovative approaches.

Following InGaAs VNW MOSFET development, III-V VNW TFETs are investigated in the next chapter as potential device technology for ultra-low power systems. Two generations of TFETs are demonstrated, delivering record performance. The analysis on the interface trap density enabled by VNW MOSFETs is critical in identifying oxide-semiconductor interface trap-assisted tunneling as a major leakage mechanism in III-V TFETs.
Chapter 4 InGaAs/InAs Heterojunction Vertical Nanowire TFETs

4.1 Introduction

As discussed in Chapter 1, TFET is considered as one of the most promising device technologies to realize ultra-low power systems by reducing the supply voltage. Figure 4-1 (a) sketches the simplistic device structures for an n-MOSFET and a n-TFET. While n+/i/n+ doping profile is used in MOSFETs, p+/i/n+ gives the tunneling operation in TFETs. The physical mechanisms of current conduction in MOSFETs and TFETs are schematically shown in the band diagrams in Figure 4-1 (b).

Figure 4-1. (a). Schematic cross-section of n-type MOSFET and TFET (b). Schematic band diagram of n-type MOSFET and TFET. The solid red lines refer to the on state while the dashed blue lines represent the off state. The grey arrow highlights the current path when the device is turned on. In TFETs, the energy window for tunneling is $\Delta \Phi$. 
MOSFETs rely on thermal injection of carriers over a barrier, and the subthreshold slope is fundamentally limited by the Boltzmann factor (60 mV/dec at room temperature) as the carriers obey Boltzmann distribution in the subthreshold regime. In contrast, the primary injection mechanism in a TFET is band-to-band tunneling (BTBT) [95], whereby charge carriers transfer from one energy band into another at a p-n junction. In the off state, no empty states are available in the channel for tunneling from the source, so the off current is low. Increasing gate bias moves the conduction band energy ($E_c$) of the channel below the valence band energy ($E_v$) of the source so that interband tunneling can occur. This switches the device to the on state, in which electrons in the energy window, $\Delta \Phi$ (green shading), can tunnel from the source valence band into the channel conduction band. Electrons in the tail of the Fermi distribution cannot tunnel because no empty states are available in the channel at their energy, so a slope of less than 60 mV/dec can be achieved. In other words, the band edge effectively filters out carriers with high energies.

Although the physics of TFETs promises sub-60 mV/dec S at room temperature (RT), which is also confirmed with numerous simulation studies, the experiments have not yet shown results superior to the well-established Si CMOS technology. Ref. [21] summarizes the simulation and experimental TFET results, which are reproduced in Figure 4-2 and Figure 4-3. A tremendous gap exists between the simulation studies predicting superior performance compared to Si MOSFETs and the experimental observations of much poorer devices. Sub-thermal subthreshold characteristics are rarely observed in these results, especially for the non-Si material systems; even for those devices that beat the thermal limit, this was only observed at very low currents, at least a few orders of magnitude lower than the typical OFF-current of MOSFETs. Moreover, the gate voltage span for the transition from the sub-60 mV/dec regime
to ON-state with sufficient drive current in these devices is large, which is disadvantageous for low Vdd application. Lastly, the ON-currents of the demonstrated devices are far below the ON-current, sometimes even lower than the OFF-current, of CMOS technology. Although efforts have been made to improve the ON-currents of the TFETs, S in those instances became extremely poor due to the leaky tunneling junction and possible parasitic leakage paths in the device [96, 97]. Such obvious discrepancy between the simulation and experimental studies of TFETs are mostly due to one or more intrinsic or technological non-idealities in the device.

Figure 4-2. Comparison of simulated n-type (right) and p-type (left) TFET transfer characteristics, reproduced from [21]. The numbers on the curves label the applied Vds. The curves are shifted so that the gate voltage where the steepest sub-threshold slope occurs is at the origin. The majority of the results predict steep SS and high on-current for a variety of materials and device architectures.
Non-BTBT transport mechanisms and band tails are the most prominent intrinsic non-idealities. Various thermal leakage currents, including substrate leakage current [98], direct generation-recombination, Shockley-Reed-Hall (SRH) [99] and Auger [100] generation-recombination currents, are the most common examples of incorrect device physics. These currents flow in parallel to the tunnel currents and can easily overshadow the true steepness of the BTBT process, resulting in poor subthreshold characteristics and temperature-dependent subthreshold characteristics. According to Equation 1-4, to the first order there is no temperature dependence expected for BTBT process.
TFETs rely on the band edge to cut off carrier transport as discussed earlier. Before the valence band in the source and the conduction band in the channel overlap in an n-TFET, ideally there is no current at all. However the band edge is not perfectly sharp and there are finite number of states in the bandgap. Before band overlapping, current can still flow by electrons tunneling through states inside the bandgap. The energy band tail of intrinsic semiconductors [101] has an exponentially decaying density of states (DOS) (also known as the Urbach tail), potentially due to the spatial and temporal fluctuation of the conduction and valence band edges due to phonons [102, 103]. The optical absorption coefficients measured in previous studies reveal a steepness of 27 mV/dec for the exponential decay of these Urbach tails in both Si and InGaAs [102, 104]. TFET performance can be significantly degraded by the Urbach tail as the subthreshold swing of a transistor is limited by the steepness of the Urbach tail [105]. Furthermore, the dopants in the semiconductor introduce impurity states in the bandgap which become delocalized at high doping levels (especially when the doping exceeds 10^{18} cm^{-3}) [104], resulting in much larger non-abrupt decay of conducting DOS at the band-edges. Therefore, although heavy doping reduces the depletion width and hence the tunneling distance [95], the tunneling steepness may be degraded due to the heavy band tail.

Technological factors that degrade TFET performance include the gate efficiency, spatial inhomogeneity, etc. The gate efficiency measures the effectiveness of the gate in modulating the channel potential, roughly equal to the potential across the tunnel junction. As discussed in section 3.3, gate efficiency can be degraded due to a high concentration of interface traps and poor electrostatic design [106]. Even if the tunneling junction has 30 mV/dec steepness, the subthreshold swing of a TFET can degrade to 60 mV/dec given a gate efficiency
of 0.5, which is not uncommon in III-V TFETs [106]. A transistor technology with low interface trap density and tight electrostatic control is needed to boost the gate efficiency.

Spatial inhomogeneity refers to the variation in parameters such as doping, thickness, material composition and strain at different spatial positions in real devices. The degradation in array NW MOSFETs compared to single NW devices discussed in section 3.2.2 illustrates the detrimental effect of spatial inhomogeneity. When a transistor with significant spatial inhomogeneity is being turned on, regions with different parameters behave as individual transistors with different threshold voltages that are being turned on together. The total current is comprised of the sum of the currents from all the individual transistors and the subthreshold characteristics are smeared out due to the variability in different transistors, resulting in large subthreshold slope. As shown in [107], electrically measured joint density of states in tunnel diodes have generally indicated a steepness > 90 mV/decade, unlike the optical Urbach measurements that are < 30 mV/decade in good semiconductors [104]. This was attributed to the smearing due to spatial inhomogeneity [107]. In order to improve TFET performance, spatial inhomogeneity in a device needs to be minimized by transistor design.

Although there has been considerable research efforts in TFETs worldwide due to its huge potential, experimental devices that can challenge Si MOSFETs are yet to be demonstrated because of the intrinsic and technological non-idealities described above. To compete with Si CMOS, it is of paramount importance to develop TFETs with steep subthreshold slope in a large current range and high current drive. $I_{60}$, defined as the highest current level where the subthreshold characteristics exhibit a transition from sub to super-60 mV/dec behavior [108], captures the requirement for high current drive.
To demonstrate TFETs with steep subthreshold characteristics and high $I_{0}$, innovations on both materials and device architectures are needed to combat the intrinsic and technological non-idealities. Equation 1-4 demonstrates the advantage of III-V materials compared to Si or Ge to obtain high drive currents due to the direct bandgap, small carrier effective masses and flexibility in band structure engineering. In this thesis, InGaAs/InAs heterojunction is used as a vehicle to study physics and optimize performance of TFETs.

As for the device architectures, TFETs are divided into two categories, transverse and inline devices, based on the directions of tunneling versus the gate electric field. The device structures are shown schematically in Figure 4-4. The transverse TFETs has a tunneling direction perpendicular to the gate electric field while in the inline TFETs the tunneling direction aligns with the gate electric field. The inline TFETs feature 1D electrostatics, uniform tunneling current and great flexibility in heterostructure design, which are excellent platform to study tunnel physics [98, 106, 109-113]. The tunneling current in the inline TFETs is proportional to the gate area, which may be a disadvantage in terms of scaling. In comparison, transverse TFETs, resembling MOSFETs, have better scalability (gate efficiency) and manufacturability, which can be used to develop a deployable technology [114-117].

![Figure 4-4](image.png)

Figure 4-4. Schematic cross-section of n-type transverse (left) and inline (right) TFET. The arrows highlight electron flow.
As shown in Chapter 1, maintaining the transistor scalability requires a 3D channel design such as NWs [115]. Using a nanowire channel in a transverse TFET not only increases the gate efficiency but also provides a way to combat spatial inhomogeneity, another important technological non-ideality, i.e., when the NW diameter is reduced to sub-10 nm, which is actually required for future technology nodes, single quantum channel operation can be realized, minimizing the device spatial inhomogeneity. Aligning the transverse NW TFET in the vertical direction leads to VNW TFET architecture. In addition to the scalability (gate efficiency) and minimized inhomogeneity of the transverse NW TFET, VNW TFET also leverages the critical advantage of inline TFETs for flexible heterostructure design. As a result, III-V VNW TFET is an ideal candidate for a deployable TFET technology in future ultra-low power systems due to the scalability (high gate efficiency), minimized inhomogeneity and flexible heterostructure design. As in the case of III-V VNW MOSFETs, bottom-up techniques have yielded excellent III-V VNW TFET device demonstrations [82, 118-123]. III-V VNW TFET technologies via a top-down approach has been pioneered in this thesis, built with the process technologies introduced in Chapter 2. The device performance and physics are discussed in this chapter.

4.2 First Generation

The starting heterostructure for the first generation VNW TFETs [124], grown by MBE on an InP wafer, is shown in Figure 2-2. The tunneling junction consists of a p⁺-i In₀.₅₃Ga₀.₄₇As heterostructure in which a 2 nm i-InAs/8 nm i-In₀.₇Ga₀.₃As “notch” has been inserted to reduce the tunnel barrier height and yield steeper subthreshold characteristics and high ON current [114]. The p⁺ source and n⁺ drain have a nominal 10²⁰ cm⁻³ C and 6×10¹⁹ cm⁻³ Si doping, respectively. In Figure 4-5, the energy band diagram along the nanowire was simulated with
Nextnano3 by solving Schrodinger-Poisson self-consistently in a double-gate geometry. The band diagram for the ON state \((V_{ds} = 0 \text{ V}, V_{gs} = 0.5 \text{ V})\) and OFF state \((V_{ds} = 0 \text{ V}, V_{gs} = -0.2 \text{ V})\) are shown on the left and right, respectively. The efficacy of the InAs/In_{0.7}Ga_{0.3}As notch to lower the tunnel barrier is evident [114].

![Energy band diagram](image_url)

**Figure 4-5.** Simulated energy band diagram along the NW for ON state (right) and OFF state (left). The self-consistent Poisson-Schrodinger solver implemented in Nextnano3 is used. The simulation was done with a 2D double-gate geometry to approximate the 3D NW channel, in order to speed up the simulation. The current model or interface states model are not included.

Device fabrication followed the process flow described in Chapter 2, with the optimized RIE recipe [43] to ensure a uniform NW profile. The major change compared to the generic process flow in Chapter 2 was an introduction of a planarization and etch back step right after NW formation. This formed a 50 nm SOG layer that covers the bottom p^+ source on which the gate stack was deposited, captured by the SEM image shown in Figure 4-6. In this way, the source-to-gate leakage current that affected the first [72] and second generation VNW MOSFETs was reduced, which enabled the scaling of the gate dielectric thickness down to 2.5 nm ALD Al2O3 (EOT = 1.2 nm). All devices have a final diameter in the intrinsic region of 50 nm and a channel length of 60 nm given by the undoped InAs/InGaAs composite layer.
thickness. Figure 4-7 shows a schematic view of the transistor fabricated in this work and the heterostructure used.

Figure 4-6. SEM image of a 50 nm SOG layer covering the p+ source on which the gate stack was deposited. The purpose of this SOG layer was to reduce gate leakage and enable further EOT scaling.
Figure 4-7. Schematic of InGaAs/InAs heterojunction VNW TFET cross-section and heterostructure.

Figure 4-8 shows subthreshold and transfer characteristics of one of the best performing single-NW TFETs. A subthreshold swing of 75 mV/dec averaged over $I_d$ from $10^{-9}$ to $10^{-7}$ A/μm is obtained at $V_{ds} = 0.3$ V. ON current of 0.27 μA/μm is extracted with $I_{off} = 100$ pA/μm and $V_{dd} = 0.3$ V ($V_{ds}=0.3$ V, $ΔV_{gs}=0.3$ V). Interestingly, $S$ slightly improves at higher $V_{ds}$ [4]. The ON/OFF current ratio exceeds $10^5$ in this device. The gate leakage current is below $10^{-12}$ A/μm in the subthreshold regime. The drain current fluctuations are attributed to the single NW nature of the device [4]. As a result, average values are reported for the figures of merit. Hysteretical behavior is observed in the subthreshold characteristics. Measurements in a narrower $V_{gs}$ range from -0.2 to 0 V yield $S=79$ mV/dec when averaging both sweeping directions. A peak $g_m$ value of 8 μS/μm is obtained.

![Graph showing subthreshold and transfer characteristics](image)

Figure 4-8. Subthreshold characteristics (left), transfer and $g_m$ characteristics at $V_{ds} = 0.3$ V (right) of one of the best performing devices.

The output characteristics of the same device are shown in Figure 4-9 (left). Triode-like characteristics are observed [118] with a typical low $V_{ds}$ super-linear behavior characteristics of
TFETs [125]. Figure 4-9 (right) shows the output characteristics in a semilog scale including the reverse regime. Clear negative differential resistance (NDR) is observed for $V_{ds}<0$ and high $V_{gs}$, confirming the tunneling nature of the device operation in the ON regime. At $V_{gs}=0.8$ V, a peak-to-valley ratio in $I_d$ of 6.2 is observed. This is the highest value reported at room temperature in III-V NW TFETs.

![Figure 4-9](image)

**Figure 4-9.** Output characteristics in linear scale (left) and in semilog scale for positive and negative $V_{ds}$ (right) of the device shown in Figure 4-8.

Across the sample a spread of device characteristics is observed. Figure 4-10 presents subthreshold characteristics of three different single-NW TFETs including the one shown in Figure 4-8 (black curve) at $V_{ds} = 0.3$ V. The device with the most positive $V_t$ shows the steepest subthreshold regime. Output characteristics of these devices in an identical scale are presented in Figure 4-11. Devices with positive $V_t$ show triode-like characteristics but devices with negative $V_t$ exhibit saturating behavior and significantly more current. While low-$V_{ds}$ super-linear behavior is observed at all $V_{gs}$ values in devices with the most positive $V_t$, devices with more negative $V_t$ only show super-linear onset at high $V_{gs}$. At low $V_{gs}$, MOSFET-like turn-on with $V_{ds}$ is observed.
Figure 4-10. Subthreshold characteristics of three single-NW TFETs at $V_{ds} = 0.3$ V illustrating device-to-device variability.

Figure 4-11. Output characteristics of the three devices shown in Figure 4-10 for positive and negative $V_{ds}$. Clear NDR is observed in all 3 devices.
To better characterize the variability in these small NW devices, minimal subthreshold swing at $V_{ds} = 0.3$ V ($S_{sat}$), DIBL and $g_{m, pk}$ at $V_{ds} = 0.3$ V for tens of devices are plotted against $V_t$ at $V_{ds} = 0.3$ V in Figure 4-12. It is clear that devices with more negative $V_t$ show worse subthreshold characteristics, i.e. larger $S_{sat}$ and DIBL. The reasons for the wide distribution of device characteristics are not clear at this moment, but the high sensitivity of TFET characteristics to small geometrical variations is likely a contributing factor. In particular, the dependence of $S_{sat}$ and DIBL on $V_t$ can be explained by the variability of the NW diameters. The peak transconductance does not have a clear dependence on $V_t$. Figure 4-13 plot DIBL and $g_{m, pk}$ at $V_{ds} = 0.3$ V against minimal subthreshold swing at $V_{ds} = 0.3$ V ($S_{sat}$). DIBL shows a reasonable correlation with $S_{sat}$ but $g_{m, pk}$ has no clear dependence on $S_{sat}$.

![Graphs showing subthreshold swing, DIBL, and peak transconductance vs. threshold voltage](image)

Figure 4-12. Minimum subthreshold swing at $V_{ds} = 0.3$ V ($S_{sat}$), DIBL and $g_{m, pk}$ at $V_{ds} = 0.3$ V for several devices plotted against $V_t$ at $V_{ds} = 0.3$ V.
Figure 4-13. DIBL and $g_{m,pk}$ at $V_{ds} = 0.3$ V plotted against minimum subthreshold swing at $V_{ds} = 0.3$ V ($S_{sat}$).

1st generation InGaAs/InAs heterojunction VNW TFETs are benchmarked with 1st generation InGaAs VNW MOSFETs in Figure 4-14. While TFETs offer much better OFF state characteristics (subthreshold swing), MOSFETs significantly outperform TFETs in terms of ON characteristics ($g_{m,pk}$).

Figure 4-14. Benchmark of $g_{m,pk}$ vs. minimum $S$ at $V_{ds} = 0.5$ V for first generation InGaAs VNW MOSFETs and InGaAs/InAs VNW TFETs.
4.3 Temperature Dependence

To further understand the physics of device operation, temperature (T) dependent measurements on another device (with relatively negative Vt) were performed from liquid N2 T (77 K) up to room T (RT), with standard three terminal analysis (Id - Vds and Id – Vgs) and diode-like analysis (Id - Vgd) methods, as discussed below.

4.3.1 Output and subthreshold characteristics

Figure 4-15 and 4-16 show the output characteristics with positive and negative Vds at varying T in linear and semilog scales, respectively. To compare the ON characteristics at different T, Id-Vds characteristics at Vgs = 0.5 V from different T are plotted together in Figure 4-17. It is helpful to establish the notion that negative/positive Vds for the TFET corresponds to the forward/reverse bias situation for a tunnel diode. Three distinct regions are clearly visible, separated by two vertical dashed lines. In the middle region (Esaki), NDR is clearly seen at all T and a certain steepness is observed. While the pre-NDR and NDR regions show very little temperature dependence, confirming the BTBT current conduction, the valley current is sensitive to the temperature, increasing with rise in temperature and suppressing the peak-to-valley current ratio. The origin of valley current is attributed to trap-assisted tunneling at the oxide/semiconductor interface [126]. A detailed discussion on the interpretation of the steepness is given in the next section.

When Vds is further decreased (region on the left to Esaki), the current eventually increases with temperature as the diode is turned on in the forward bias regime and this current is contributed by thermionic emission. In the positive Vds region, Id increases slightly with T. Figure 4-18 plots the drain current (log scale) in the linear region (Vgs = 0.5 V, Vds = 0.05 V) as
Figure 4-15. Output characteristics in linear scale of a representative device at different temperatures.

Figure 4-16. Output characteristics in Figure 4-15 plotted in semilog scale.
Figure 4-17. $I_d-V_{ds}$ characteristics at $V_{gs} = 0.5$ V in semilog scale at different temperatures from 77 to 240 K. Negative/positive $V_{ds}$ region of a TFET corresponds to forward/reverse bias of a tunnel diode. Three regions, thermionic emission, Esaki and BTBT are observed.

Figure 4-18. Drain current at $V_{ds} = 0.05$ V and $V_{gs} = 0.5$ V plotted as a function of the 1.5 power of the InAs bandgap at measured temperatures from 77 K to 240 K.
a function of 1.5 to the power of the InAs bandgap at measured temperatures. The linear relation confirms BTBT nature of the current because the observed $e^{-E_g^{1.5}}$ dependence matches Equation 1-4 for BTBT.

The $I_d$-$V_{ds}$ characteristics at varying $T$ with the device turned off ($V_{gs} = -0.1$ V) are shown in Figure 4-19. The NDR (Esaki) and diode turn-on (thermionic emission) regions are also observed as when the device is turned on. The critical difference is the appearance of a thermally activated regime ($V_{ds} > -0.2$ V) with huge temperature dependence.

![Diagram of $I_d$-$V_{ds}$ characteristics](image)

Figure 4-19. $I_d$-$V_{ds}$ characteristics at $V_{gs} = -0.1$ V in semilog scale at different temperatures from 77 to 240 K. Compared to $V_{gs} = 0.5$ V characteristics in Figure 4-17, $V_{gs} = -0.1$ V also shows the thermionic emission and Esaki region. The critical difference is the presence of a thermally activated region ($V_{ds} > -0.2$ V).
To better characterize the thermally activated region, subthreshold characteristics at different T and $V_{ds} = 0.05$ V are shown in Figure 4-20. A T-independent leakage current floor below pA/um range is reached, a unique feature of the NW geometry not seen in planar TFETs [99]. Another observation in Figure 4-20 is the sharp saturation of $I_d$ at high $V_{gs}$, which could be due to high interface states ($D_{it}$) density inside InGaAs conduction band or the depletion of p+ source in the overlapping region with the gate.

Figure 4-20. Subthreshold characteristics at $V_{ds}=0.05$ V and T = 77 to 300 K.

The subthreshold current is very sensitive to T, which is not expected from a pure BTBT conduction mechanism. This suggests that a thermal process is involved. Figure 4-21 (a) plots $S$ as a function of drain current at different T, which clearly shows the strong temperature dependence. At low temperatures (190, 140 and 77 K), $S$ remains very flat over 4 orders of magnitudes of current, which is a typical feature in MOSFETs. The average $S$ at $V_{ds} = 0.05$ V is
shown as a function of T in Figure 4-21 (b), indicating that the thermal limit is never reached. An Arrhenius plot of the drain current \([\ln (I_d/T^{3/2}) \text{ vs. } (1/kT)]\) at several \(V_{gs}\) in the subthreshold regime is shown in Figure 4-22 (a). The linear relations confirm the thermal activated behavior. As observed in Figure 4-22 (b), the extracted thermal barrier height \(q\Phi_B\) from the Arrhenius plot is linearly dependent on \(V_{gs}\), with a slope of 0.89.

Figure 4-21. (a) Subthreshold swing at \(V_{ds} = 0.05\) V as a function of drain current, extracted at different temperatures. (b) Temperature dependence of average S at \(V_{ds} = 0.05\) V. Red dashed line is the thermal limit of the subthreshold swing in a MOSFET.

Figure 4-22. Arrhenius plot of the subthreshold current at various \(V_{gs}\) values versus inverse thermal energy. The slope of the linear relationship is the thermal barrier height. (b) Thermal barrier height extracted from the slopes in (a) as a function of \(V_{gs}\).
Temperature-dependent subthreshold characteristics in TFETs are not that uncommon, especially in III-V TFETs. Figure 4-23 compiles the subthreshold characteristics at different temperatures in recent published III-V TFETs from various research groups [96, 99, 118, 121, 127-130], revealing a common problem in the field. In section 4.5, the physical mechanism behind the unexpected T dependence is discussed in detail.

4.3.2 Conductance-voltage characteristics

As mentioned earlier, the steepness of the band edge states (Urbach tail) is an intrinsic non-ideality that limits the transistor subthreshold swing. Urbach tail can be significantly degraded by heavy doping or spatial inhomogeneity and the nanowire channel design is promising in reducing spatial inhomogeneity. Therefore, it is of great interest to extract the Urbach tail steepness in NW TFETs but the subthreshold characteristics presented previously...
fail to reveal the true tunneling steepness because of overshadowing by a thermal leakage current and the limited gate efficiency.

Recently, conductance-voltage (GV) characteristics have emerged as a promising method to estimate the band tails in tunnel diodes and TFETs while minimizing the impact of the gate efficiency [107]. In this section, GV characteristics of G1 TFETs are analyzed in an effort to estimate the band edge sharpness in the VNW TFETs.

The basic idea of GV method to measure the Urbach tail steepness in tunnel diodes is introduced according to [107]. In a tunnel diode the BTBT current can be expressed as

$$I \propto \int (f_c - f_v) \times T(V) \times D_j(V, E) \times dE$$

(4-1)

$E$ is energy, $V$ is bias, $f_c, f_v$ are the Fermi occupation probabilities on the n and p sides, $T(V)$ is the tunnel probability (function of $V$ only to the first order, as carriers at different energy have the same energy barrier), $D_j(V, E)$ is the joint density of states (depends on both $V$ and $E$). $\frac{d \int T(V) \times D_j(V, E) \times dE}{dV}$ is the quantity of interest (tunnel steepness), but the current is complicated by the Fermi probability $(f_c - f_v)$. The trick is to obtain the absolute conductance $G$:

$$G = \frac{I}{V} = \int \frac{(f_c - f_v)}{V} \times T(V) \times D_j(V, E) \times dE$$

(4-2)

Because $\frac{(f_c - f_v)}{V}$ is a linear function of $V$ and $T(V) \times D_j(V, E)$ varies exponentially with $V$, the slope of the conductance, $S = \left(\frac{d \log G}{dV}\right)^{-1}$ measures mostly the tunnel steepness $(T(V) \times D_j(V, E))$ [107]. It is worth noting that this method yields an estimation, not the exact value. In addition, the conductance slope captures the steepness of the Urbach tails and tunnel probability combined, not just Urbach tails. $S$ should be sharper than the steepness of Urbach tails or the tunnel probability.
A diode is a two-terminal device while a TFET has three terminals. To obtain similar G-V characteristics in three-terminal n-type TFETs, \( V_s \) is grounded and \( V_g \) and \( V_d \) are swept at the same time with fixed \( V_g - V_d \) during the measurements. The conductance is the ratio of the drain current to the drain bias \( V_d \). In this way, the potential change across the tunnel junction is maximized, partially overcoming the problem of limited gate efficiency [107].

Figure 4-24 (a) plots the standard subthreshold characteristics of a representative device at various temperatures from 77 K to 423 K and Figure 4-24 (b) – (d) shows the conductance-voltage characteristics of the same device at three different temperatures (a) (b) (c) (d). Conductance voltage characteristics of the same device with various \( V_{gd} \) biases at 77, 290 and 423 K.
(77, 290 and 423 K), with several fixed $V_{gd}$ values at each temperature. Compared to the standard output characteristics ($I_d$ vs. $V_{ds}$ at fixed $V_{gs}$) in Figure 4-16, the G-V curves show smooth transition around $V_d = 0$ V rather than a singular behavior, allowing the extraction of steepness. Two distinct regimes are clearly visible, the Esaki diode regime and the backward diode regime, as labeled in Figure 4-24 (b). The device shows NDR behavior and very small temperature dependence in the Esaki diode region. On the contrary, the backward diode regime is strongly temperature dependent.

To quantify the temperature dependence, two types of steepness, $S_E$ and $S_B$ corresponding to these two regions can be extracted, which are plotted as a function with temperature in Figure 4-25, together with the steepness calculated from the subthreshold characteristics, $S_s$ (the subthreshold swing). $S_E$ in the NDR region has very small positive temperature dependence throughout the wide $T$ range, signaling a pure tunneling current. $S_s$ exhibits strong thermally activated behavior up to 423 K. $S_B$ in the backward diode region, on the other hand, has a very strong temperature dependence and closely follows the subthreshold swing from low temperatures up until RT where $S_B$ crosses and tracks $S_E$. The difference between Esaki and backward diode regimes is that in Esaki diode region, $V_{gd}$ is positive enough that an inversion charge layer under the channel (electrons are supplied by the drain) is formed. In the Backward diode regime, as there is no inversion charge layer in the channel, the drain has only weak impact on the channel potential through short-channel effects, which explains the similar values observed in $S_B$ and $S_s$. The implication is that $S_B$ should have the same interpretation as the subthreshold swing $S_s$ below 300 K, which is discussed in section 4.5. Above 300 K, $S_s$ stops following $S_s$ and begins to track $S_E$, which can be understood with the G-V$_d$ characteristics at 423 K in Figure 4-24. The G-V curve at negative $V_{gd}$ behaves similarly
to positive $V_{gd}$, because at high temperatures, considerable inversion charges are present in the channel even at very negative $V_{gd}$ and the drain is able to grab the channel potential.

![Graph showing minimal $S_E$, $S_B$, extracted from NDR and backward diode regimes from GV characteristics and $S_S$ measured from subthreshold characteristics plotted as a function of temperature. The green dashed line represents the thermal limit.](image)

Figure 4-25. Minimal $S_E$, $S_B$, extracted from NDR and backward diode regimes from GV characteristics and $S_S$ measured from subthreshold characteristics plotted as a function of temperature. The green dashed line represents the thermal limit.

In the Esaki regime where there exists inversion charge in the channel, both the gate and drain can affect the potential in the channel. The gate is expected to be rather inefficient because the charges in the channel are able to screen the gate potential. The drain, on the other hand, can grab the channel potential effectively through the inversion layer. Figure 4-26 shows the conventional output characteristics ($I_d$ vs $V_{ds}$ at fixed $V_{gs}$) and the corresponding conductance of the same device in semilog scale. At the same bias condition where $S_E$ is calculated at 290 K in Figure 4-24 (c), a steepness denoted $S_O$ of 160 mV/dec can be obtained from Figure 4-26 (b). Note that this number is softer than $S_E = 108$ V/dec in Figure 4-24 (c) because $V_{gs}$ also changes when evaluating $S_E$. The effect of $V_{gs}$ can be appreciated in Figure 4-27 which shows $I_d$ vs $V_{gs}$ at $V_{ds} = -0.25$ V. At $V_{gs} = 0$ V, which is roughly the bias point where $S_E$ is evaluated, the
steepness there is very soft, around 500 mV/dec. This means that the gate is not very effective at modulating the potential in the channel due to the presence of the inversion layer.

Figure 4-26. (a) Output characteristics at 290 K in semilog scale. (b) Corresponding conductance versus V_{ds} in semilog scale. A minimum slope of 160 mV/dec can be extracted from the NDR region.

Figure 4-27. I_{d} – V_{gs} characteristics at V_{ds} = -0.25 V and T = 298 K. The minimum slope is around 500 mV/dec.

The purpose of the G-V analysis is to estimate the tunneling or band tail steepness for VNW TFETs. In the three types of steepness in Figure 4-25, only $S_0$ extracted from NDR region is temperature independent, and is therefore the only potential candidate to reveal the
steepness of the band tails. The interpretation of $S_E$ in the NDR region depends on the band alignment of the valence band in the source and the conduction band in the channel. $S_E$ of 108 mV/dec is observed around $V_{gs} = -0.2$ V and $V_{ds} = -0.2$ V in the G-V characteristics of $V_{gd} = 0$ V at 300 K shown in Figure 4-24 (c). The band alignment around this bias point can be inferred by examining the subthreshold characteristics shown in Figure 4-24 (a). Due to the significant thermal signature, it is clear that in the subthreshold region the bands do not overlap, otherwise BTBT would dominate and there should be no temperature dependence. Therefore, at the bias

![Figure 4-28. Distribution of minimum subthreshold swing ($S_S$) at $V_{ds} = 0.3$ V ($S_{sat}$), minimum $S_E$ at $V_{gd} = 0$ V versus $V_t$ at $V_{ds} = 0.3$ V. $S_E$ has much less variability compared to $S_S$.]

point of $V_{gs} = -0.2$ V and $V_{ds} = 0.05$ V at 300 K, the bands do not overlap. When $V_{ds}$ is decreased from 0.05 V to -0.2 V where $S_E$ is extracted, the bands still do not overlap because the conduction band in the channel is further lifted. The implication is that $S_E$ of 108 mV/dec reflects the combined steepness of the band edge and tunnel probability, representing a lower limit to the band edge steepness. However, it is unclear at this point if the band edge states
originates from Urbach tails or oxide/semiconductor interface trap states. One observation in favor of the Urbach tails is that while SS has a wide distribution in the devices, SE features much less variability, shown in Figure 4-28 where the distributions of SS and SE are plotted versus Vt. The small variability of SE suggests that this steepness is more likely due to intrinsic property – Urbach tails. Further evidence supporting the role of the Urbach tails is presented in Section 4.5.

4.4 Second Generation

A second device fabrication run followed the first generation on the same heterostructure, but with greater attention to ALD conditioning to improve the quality of the oxide/semiconductor interface. After RIE etching, aqueous digital etch was performed in order to trim the nanowire diameter by 20 nm and improve the quality of the interface. A 15 nm drain underlap at the top of the nanowire was introduced to improve the subthreshold characteristics. A drain underlap reduces the electric field at the drain-gate junction that promotes ambipolar tunnel leakage, leading to sharper subthreshold characteristics [115]. An RTA step at 350 °C for 3 min in N2 environment was performed at the very end of the process to further improve the oxide/semiconductor interface quality, as in G2 MOSFETs. The final device features a single nanowire with channel length of 60 nm, D = 40 nm and 3 nm Al2O3 (EOT ~ 1.5 nm). It is worth pointing out that this sample was fabricated side-by-side with G2 MOSFETs detailed in Chapter 3. Dry etch, digital etch, ALD processes and RTA were performed in the same tools at the same time.

Figure 4-29 shows the output characteristics of a representative device in a linear scale on the left and in semilog scale including the negative Vds regime on the right. The linear output characteristics exhibit good saturation. Clear negative differential resistance was observed for
$V_{ds} < 0$ and high $V_{gs}$, confirming the tunneling nature of device operation in the ON regime. At $V_{gs} = 0.6 \, V$, a peak-to-valley ratio of 3.4 is observed.

![Figure 4-29. Room temperature output characteristics of an exemplary VNW TFET in linear scale (left) and in semilog scale (right).](image)

The subthreshold and $g_m$ characteristics of the device in Figure 4-29 are shown in Figure 4-30 (top). A peak $g_m$ ($g_{m, pk}$) of 17 $\mu$S/$\mu$m was obtained at $V_{ds} = 0.3 \, V$ and DIBL is 102 mV/$V$. This is to be compared with $g_{m, pk} = 8 \, \mu$S/$\mu$m and DIBL = 320 mV/$V$ in G1 devices, despite a thicker oxide thickness here.

To take a closer look at the subthreshold characteristics, the subthreshold swing vs. drain current is plotted in Figure 4-30 (b). A minimum $S$ of 55 mV/dec was observed at $V_{ds} = 0.05 \, V$ and was further reduced to 53 mV/dec at $V_{ds} = 0.3 \, V$. The slight improvement of $S$ with drain bias might be due to an increase in gate efficiency as a result of the enhanced depletion of channel charge by the higher drain bias. At $V_{ds} = 0.3 \, V$, the subthreshold swing remains sub-thermal for over two orders of magnitude of current, delivering an $I_{60}$ as high as $4.3 \, nA/\mu$m, which set the record of published experimental TFETs for $V_{ds} < 1 \, V$ at the time of device demonstration. The improvement in the subthreshold characteristics may be partially explained...
by the introduction of the 15 nm drain underlap region which reduces channel-drain electric field and consequently decreases ambipolar tunnel leakage current [115]. Another reason for this improvement is revealed by the temperature dependent measurements discussed next.

Figure 4-30. Subthreshold and $g_m$ characteristics (top) and subthreshold swing vs. $I_d$ characteristics (bottom) of InGaAs VNW TFET.
In Figure 4-31 (top), subthreshold characteristics at different T and \( V_{ds} = 0.05 \) V are shown from liquid nitrogen temperature to room temperature. Apart from a positive threshold voltage shift with decreasing temperature, the subthreshold swing shows minimal change as a function of temperature. This is evidenced in Figure 4-31 (bottom) (the green stars) where the minimum S at \( V_{ds} = 0.05 \) V is plotted against T. Figure 4-31 (bottom) also includes data for G1 TFETs (50 nm diameter) and G2 MOSFETs with 40 nm diameter. While S in G1 TFET suffered from a significant temperature dependence, changing from 145 to 30 mV/dec (room to liquid N\(_2\) T), it only varies from 55 to 37 mV/dec over the same T range in G2 devices.

**Figure 4-31.** Temperature dependence of subthreshold characteristics at \( V_{ds} = 0.05 \) V (top) and minimum S at \( V_{ds} = 0.05 \) V as a function of temperature (bottom) for 2 generations of TFETs and G2 MOSFETs.

### 4.5 Trap Assisted Tunneling and Urbach Tail Steepness

In this section, the significant temperature dependence observed in G1 TFETs and many others shown in Figure 3-23 is explored. Many mechanisms can possibly explain the thermal signature in the subthreshold regime, including temperature-dependent gate efficiency due to
large interface state density ($D_{it}$), phonon-assisted tunneling, and trap-assisted tunneling via oxide-semiconductor interface traps or band edge states.

The gate efficiency due to $D_{it}$ is unlikely to be the root cause as a different temperature signature due to $D_{it}$ is observed on MOSFETs on similarly etched InGaAs surfaces [48]. Phonon-assisted tunneling in III-V TFETs has been shown via simulation to increase the leakage current without impacting significantly on the subthreshold swing [131]. Trap-assisted tunneling (TAT) via oxide-semiconductor interface trap states [99, 105, 132-136] has been found theoretically to severely degrade the subthreshold swing and lead to significant temperature dependence. To our best knowledge, the comparison of two generations of VNW TFETs serve as the first experimental evidence of this hypothesis because the only significantly difference between the first and second generation TFETs is the drastic reduction in the interface trap density and the temperature dependence is significantly suppressed consequently. As estimated in Chapter 3, a ten-fold decrease in $D_{it}$ ($1.6 \times 10^{13}$ vs. $1.4 \times 10^{12}$ cm$^{-3}$) was achieved in G2 devices due to better ALD chamber conditioning and post-deposition RTA. This result establishes the notion that the significant temperature dependence of the subthreshold characteristics in G1 is due to extrinsic property – the interface trap density, rather than intrinsic properties, such as Urbach band tails.

In essence, oxide-semiconductor TAT happens because an electron can reach the conduction band from the valence band via a combination of tunneling and thermal emission, as shown in Figure 4-32. This undesired tunneling process is electric field dependent in the same way as the BTBT current. The electric field enhanced generation rate is much higher than the classical Shockley-Read-Hall (SRH) formalism that neglects tunneling. The generation rate is further enhanced by the Poole-Frankel effect, the lowering of the electron barrier due to the
Coulomb interaction of the trap with the lattice [105]. While in MOSFETs $D_{it}$ only affects the gate efficiency, in TFETs interface traps participate directly in the current conduction and a much stricter requirement (low-$10^{11}$ cm$^{-3}$ range) in TFETs is needed to unveil their true potential.

Figure 4-32. Schematic of band structure of the VNW TFETs illustrating the trap-assisted tunneling process. An electron can reach the conduction band from the valence band via a combination of tunneling and thermal emission.

The residual weak T dependence in G2 TFETs might be due to an Auger generation process [100] or Urbach tails-assisted tunneling. It is worth pointing out that the minimal S in Figure 4 (b) for both generations of TFETs keeps improving with decreasing T into the liquid N$_2$ T range, highlighting the potential of obtaining a true tunneling steepness sharper than 37 mV/dec.
Following the discussion on the implication of the GV characteristics at the end of section 4.3.2, SE from G2 devices are extracted to be in the range of 140-160 mV/dec, compared to the 100-120 mV/dec in G1 TFETs. The even worse slope despite a reduction in the Dit and improvement in the subthreshold characteristics supports the hypothesis that Urbach tails are responsible for SE, not interface traps. Therefore, 100-120 mV/dec extracted from G1 devices appears to be the limit of the Urbach tail steepness. The worse slope in G2 devices could be due to the lower drain efficiency in controlling the channel potential due to the introduction of a drain underlap region. The large value despite a NW channel could be attributed to the significant adverse effect of heavy doping [104, 137] in the source (1×10²⁰ cm⁻³). This result again highlights the strict demand for higher perfection in TFETs as compared with MOSFETs [135].

4.6 Benchmark

Following [21], G2 TFETs are benchmarked against experimental TFETs with sub-thermal S at RT and V_ds < 1V¹⁰ in Figure 4-33, at the time of device demonstration. Top-down VNW TFETs offer the highest I₆₀ in devices with V_ds < 1 V and show sub-thermal subthreshold characteristics over two orders of magnitude of current in a clean, smooth measurement with ample data points. This is testimony to the increased flexibility and precision heterostructure growth that is afforded by a top-down fabrication approach.

Another benchmark methodology uses the I_on-versus-I_off metric to capture the tradeoff between the dynamic switching speed and the standby power, commonly seen in mature Si

¹⁰ Private communication with P. Paletti and A. Seabaugh.
CMOS technology [138]. Figure 4-34 benchmarks $I_{on}$ vs. $I_{off}$ among published vertical NW TFETs based on III-V materials at $V_{dd} = 0.3$ V ($V_{ds} = 0.3$ V, $\Delta V_{gs} = 0.3$ V). Several points in this figure have other $V_{ds}$ values as marked due to data availability. Compared to other vertical III-V NW TFETs, the top-down approach demonstrated in this thesis exhibit an excellent combination of steep slope and ON current, delivering high $I_{on}$ at low $I_{off}$.

![Subthreshold Swing vs Drain Current](image)

Figure 4-33. $S$ vs $I_d$ in experimental TFETs with sub-thermal $S$ at RT and $V_{ds} < 1$ V (Adapted from [21]).
In this chapter, the physical mechanism (BTBT) and non-idealities including band tails, gate efficiency and spatial inhomogeneity for TFET technologies are first reviewed. Leveraging the flexible heterostructure design capability of III-V materials and the ultimate scalability of NW channel while minimizing the spatial inhomogeneity and gate inefficiency, top-down III-V VNW TFET is an ideal transistor technology to realize the ultra-low power electronics system.

4.7 Chapter Summary

Figure 4-34. $I_{on}$ vs. $I_{off}$ at $V_{dd} = 0.3 \text{ V}$ ($V_{ds}=0.3 \text{ V}, \Delta V_{gs}=0.3 \text{ V}$ with exceptions marked next to data points) for recently published vertical NW TFETs containing III-V materials. All devices but the present ones are fabricated through bottom-up techniques.
The first experimental demonstration of III-V VNW TFETs with InGaAs/InAs heterojunction by top-down approach is then introduced, starting from a discussion of the excellent room temperature electrical performance. Device variability is documented in detail, which is likely caused by geometrical variations in nanowire diameter and heights of the planarization layers. BTBT proves to dominate the ON current while the subthreshold characteristics suffer from unexpected temperature dependence. Detailed analysis confirms the strongly thermal activated behavior in the subthreshold regime. Conductance-voltage characteristics are carefully studied, revealing a steepness of 100 – 120 mV/dec for band edge states in VNW TFETs, due to either Urbach tails or oxide/semiconductor interface traps.

Second generation devices demonstrate sub-thermal subthreshold characteristics over two orders of magnitude of current. A minimal subthreshold swing of 53 mV/dec at $V_{ds} = 0.3$ V has been obtained at room temperature. An $I_{60}$ of 4.3 nA/μm has been achieved at $V_{ds} = 0.3$ V, setting the record of any experimental TFETs for $V_{ds} < 1$ V. This result testifies the increased flexibility and precision heterostructure growth that is afforded by a top-down fabrication approach.

The comparison of two generations of the devices confirms oxide/semiconductor interface trap-assisted tunneling as the source of the significant temperature dependence, because the T-dependence is greatly suppressed in the second generation devices where the only significant change compared to the first generation is a ten-fold reduction in the interface trap density. The consistent slopes observed in the NDR region of the conductance-voltage characteristics endorses the 100-120 mV/dec steepness of Urbach tails in these VNW TFETs, degraded by heavy doping in the source.
Chapter 5 Summary and Future Work

5.1 Summary

This work has pioneered III-V vertical nanowire transistor via top-down approach for logic applications in future ultra-low power systems. Vertical nanowire transistor technology has the potential to carry Moore’s law further than any other device architecture due to its ultimate scalability. The vertical channel design fully unleashes the potential of III-V materials by enabling band structure engineering along the transport direction. With the insertion of III-Vs as the channel material, a single technology platform can enable both optimal III-V TFETs and III-V MOSFETs through heterostructure engineering, which can potentially support a wide spectrum of applications form high-performance computing to ultra-low power systems. The combination of MOSFETs and TFETs enables tremendous circuit and system design flexibility, which can significantly boost the circuit performance.

In this thesis, an entire process flow and the critical fabrication modules for VNW transistor fabrication were developed from scratch. A novel ICP-RIE technique based on BCl₃/SiCl₄/Ar chemistry for fabricating sub-20 nm III-V nanostructures with a smooth, vertical sidewall and high aspect ratio (> 10) was developed for the first time. No undesirable material selectivity is observed in heterostructures containing In, Ga, Al, As, P and Sb. To mitigate dry etch damage in a controllable way, digital etch, comprised of multiple cycles of self-limiting low power oxygen plasma oxidation and diluted sulfuric acid rinse was applied to 3D InGaAs nanostructures for the first time. DE is capable of controllably thin down the nanowire diameter while preserving the shape and sidewall roughness. Both the subthreshold and ON...
characteristics of VNW MOSFETs are improved by DE, indicating enhanced sidewall interfacial quality. To demonstrate sub-10 nm III-V nanowires needed for future technology nodes, a novel alcohol-based digital etch technique was introduced, which addresses the limitations of the water-based approach. Sub-10 nm fins and nanowires with a high yield were achieved. A planarization and etch back process based on SOG was developed to remove part of and isolate the gate metal from the contact metal to NW top.

InGaAs vertical nanowire MOSFETs via a top down approach were demonstrated for the first time in this thesis, followed by two generations of devices improvements. G3 devices demonstrate the highest quality factor and ON current of any III-V VNW MOSFETs by bottom-up or top-down techniques, despite a relatively resistive Schottky top contact. This highlights the potential of the top-down fabrication approach developed in this thesis. With the introduction of solvent-based DE, switching characteristics are observed in G3 devices with a diameter as small as 14 nm. This is enabled by improved structural integrity due to the low surface tension of solvents. Close to ideal subthreshold swing was achieved in the latest two generation devices due to a tenfold decrease in the interface trap density compared to the first generation. This was enabled by better ALD chamber conditioning and oxide scaling.

The first experimental study of source/drain asymmetry in III-V VNW MOSFETs was also carried out. The asymmetry in transconductance and the saturation behavior of the output characteristics is attributed to the inherent difference between top and bottom contacts. Diameter uniformity along the nanowire is responsible for asymmetry in the subthreshold characteristics. The suppression of diameter uniformity in G2 devices eliminated the asymmetry of the electrostatics observed in G1 devices. The metal contact to the NW tip was revealed as
the bottleneck to the performance of III-V VNW MOSFET technology and prevented the demonstration of sub-10 nm NW devices.

The first experimental demonstration of III-V VNW TFETs was introduced with InGaAs/InAs heterojunction. BTBT was proven to dominate the ON regime while detailed analysis revealed strong thermal activated behavior in the subthreshold regime. Second generation TFETs demonstrate sub-thermal subthreshold characteristics over two orders of magnitude of current and much reduced temperature dependence. A record high $I_{60}$ among any experimental TFETs for $V_{ds} < 1$ V was achieved at the time of device demonstration, highlighting the increased flexibility and precision heterostructure growth afforded by the top-down approach.

The comparison between the two generations of TFETs confirms oxide/semiconductor interface trap-assisted tunneling as the source of the temperature dependence. This is because $T$ dependence is greatly suppressed in the second generation devices where the only significant change compared to the first generation is a tenfold reduction in the interface trap density. Detailed analysis on the conductance-voltage characteristics on both generations of devices reveals a 100-120 mV/dec steepness of Urbach tails in the VNW TFETs, pointing out the detrimental effect of heavy doping.

In conclusion, core processing technologies for top-down III-V VNW transistors were developed in this thesis, enabling the demonstration of state-of-the-art III-V VNW MOSFETs and TFETs with record performance. Fabrication technologies for sub-10 nm III-V VNW transistors were developed, including novel RIE technology and digital etch technique. Relevant device physics for VNW MOSFETs, including source/drain asymmetry and the impact of the
number of NWs in parallel were investigated and the roadblock, top metal/semiconductor contact, was identified. Oxide/semiconductor interface trap assisted tunneling was revealed as a major leakage mechanism in TFETs, precluding steep-slope device demonstrations and Urbach tail steepness in a nanowire channel was extracted.

\[ S = 60 \times \left(1 + \frac{qD_{it}t_{ox}}{c_{ox}}\right) \]

### 5.2 Future Work

III-V VNW transistor technology is still in the infancy and numerous innovations in many areas are necessary before it becomes a competitive technology to Si CMOS.

On the front-end of the process, self-aligned contacts are needed to replace the planarization and etch back technique which has limited manufacturability as a result of the use of timed etch. Beginning with sub-10 nm nanowires, sub-10 nm transistors are still challenging to fabricate structurally because of the many steps following NW formation that involve wet processes. Metrology methods for sidewall characterization, including morphology and chemical analysis (roughness and stoichiometry) on the ultra-scaled NWs need to be developed. Detailed study of the effect of digital etch on the surface of planar, fin and NW structures in terms of roughness and stoichiometry is also required to further improve this technique.

For III-V VNW MOSFETs, the top metal/semiconductor contact remains the bottleneck to transistor performance. Even with sub-10 nm transistor structures, the devices might not function due to the difficulty of contacting such a thin NW. Sidewall quality is another major concern that needs significant effort to perfect in order to achieve competitive performance. Co-integration with p-type MOSFETs on a Si substrate is a key requirement. The introduction of mechanical stress is challenging in vertical devices, but definitely worth the effort. It is also of
great value to develop electrical characterization techniques to measure the source/drain contact properties separately and the capacitance-voltage characteristics. Assessment of process variations on device variability is needed and methods to reduce process variations are worth developing. Source/drain asymmetry needs to be minimized for circuit applications.

For III-V VNW TFETs, the roadblock is the presence of high oxide/semiconductor interface trap density. Methods to reduce $D_{it}$ to the range of low $10^{11}$ eV$^{-1}$cm$^{-2}$ are of paramount importance. As Urbach tail steepness is significantly degraded by heavy doping, heterostructure and device architecture innovations to eliminate the need of high source doping are worth exploring. To achieve high ON current with acceptable OFF leakage, heterojunction with smaller tunnel barrier height needs to be optimized.

To overcome the difficulty of obtaining both high ON current and steep subthreshold slope in TFETs, superlattice-source (SLS) transistor is an interesting device concept worth pursuing and uniquely supported by the top-down vertical nanowire transistor technology developed in this thesis. Compared to TFETs that rely on band edges to filter out high-energy carriers, SLS device inserts a superlattice (SL) region in the source to achieve steep subthreshold slope operation [139]. It is well known that the band structure of a superlattice is comprised of minibands (where the DOS is high) and minigaps (where the DOS is low) whose energy scale is on the order of several tens to hundreds of meV. By accurately selecting the constituent materials and adjusting the SL physical dimensions, a SL region inserted in the source can filter out the carriers in the subthreshold and off-state by minigaps, thus resulting in steep turn-on characteristics. The device can possibly deliver a drive current that is comparable to that of MOSFETs when the miniband is properly aligned with the channel in the on-state, while achieving sub-thermal subthreshold characteristics. III-V vertical nanowire transistor
architecture is ideal to implement the SLS transistor because it requires precision heterostructure engineering and lateral quantization for SL, uniquely supported by III-V VNW transistor technology. Top-down approach has an advantage for this device concept because III-V superlattice structures with atomically sharp interfaces and nanometer period is challenging for bottom-up growth techniques, but can easily be implemented with well-developed epitaxial growth techniques [140].
## Appendix:

### A.1 Sample Process Flow

<table>
<thead>
<tr>
<th>Module</th>
<th>Step</th>
<th>Tool</th>
<th>Recipe</th>
<th>Spec</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alignment</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Sample clean</td>
<td>Photo-Wet-r</td>
<td>A/M/I 10s</td>
<td></td>
</tr>
<tr>
<td></td>
<td>EBL</td>
<td>Elionix</td>
<td>PMMA A8 SOP</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Metal dep</td>
<td>eBeamFP</td>
<td>Zerorotation/Ti-Au</td>
<td>5nmTi/60nmAu</td>
</tr>
<tr>
<td></td>
<td>Lift-off</td>
<td>Photo-Wet-r</td>
<td>Acetone</td>
<td></td>
</tr>
<tr>
<td>Nanowire</td>
<td>Clean</td>
<td>Asher</td>
<td>SOP</td>
<td>800 W, 2 min</td>
</tr>
<tr>
<td></td>
<td>Adhesion</td>
<td>sts-CVD</td>
<td>HFSINXIN</td>
<td>21 s</td>
</tr>
<tr>
<td></td>
<td>EBL</td>
<td>Elionix</td>
<td>HSQ SOP</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Check HSQ</td>
<td>SEM</td>
<td>SOP</td>
<td></td>
</tr>
<tr>
<td></td>
<td>RIE</td>
<td>SAMCO</td>
<td>Recipe 53</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Strip HSQ</td>
<td>Acidhood</td>
<td>BOE</td>
<td>1 min</td>
</tr>
<tr>
<td></td>
<td>Check NW</td>
<td>SEM</td>
<td>SOP</td>
<td></td>
</tr>
<tr>
<td>Gate stack</td>
<td>Digital etch</td>
<td>Acidhood+Asher</td>
<td>SOP</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Gate dielectric</td>
<td>ICL-ALD</td>
<td>Al₂O₃ at 250 °C</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Gate metal</td>
<td>AJA-TRL</td>
<td>XinW40nm200W</td>
<td>40 nm</td>
</tr>
<tr>
<td></td>
<td>Annealing</td>
<td>tube-B1</td>
<td>FGA</td>
<td>30 min at 350 °C</td>
</tr>
<tr>
<td>Gate pattern</td>
<td>Gate protect</td>
<td>Exposure</td>
<td>Elionix PMMA A8 SOP</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Metal dep</td>
<td>eBeamFP</td>
<td>5+15 nm Ti/Au</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Lift-off</td>
<td>Photo-Wet-r</td>
<td>Acetone</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Clean</td>
<td>Asher</td>
<td>SOP</td>
<td>800 W, 2 min</td>
</tr>
<tr>
<td>Metal pattern</td>
<td>Bakeout</td>
<td>Hotplate2</td>
<td>SOP</td>
<td>2 min at 200 °C</td>
</tr>
<tr>
<td></td>
<td>Spin</td>
<td>Coater</td>
<td>maN</td>
<td>1 min at 2.5K</td>
</tr>
<tr>
<td></td>
<td>Bake</td>
<td>Hotplate2</td>
<td>SOP</td>
<td>2 min at 90 °C</td>
</tr>
<tr>
<td></td>
<td>Exposure</td>
<td>Elionix</td>
<td>maN SOP</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Develop</td>
<td>Photo-Wet-r</td>
<td>MF-CD-26</td>
<td>75s</td>
</tr>
<tr>
<td></td>
<td>Check</td>
<td>SEM</td>
<td>SOP</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Check</td>
<td>AFM</td>
<td>SOP</td>
<td></td>
</tr>
<tr>
<td>Process</td>
<td>Equipment</td>
<td>Material</td>
<td>Parameters</td>
<td></td>
</tr>
<tr>
<td>-------------------------</td>
<td>------------------------</td>
<td>----------</td>
<td>------------------------------------------------</td>
<td></td>
</tr>
<tr>
<td>Gate etch</td>
<td>Plasmaquest</td>
<td>XZW</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Resist strip</td>
<td>Photo-Wet-r</td>
<td>Water bath</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Resist strip</td>
<td>Asher</td>
<td>SOP</td>
<td>Until resist cleared</td>
<td></td>
</tr>
<tr>
<td><em>1 etch back</em></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FOX spin</td>
<td>Coater</td>
<td>SOP</td>
<td>1 min at 4K</td>
<td></td>
</tr>
<tr>
<td>Reflow</td>
<td>Hotplate1</td>
<td>SOP</td>
<td>2 min at 150 °C</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Hotplate2</td>
<td>SOP</td>
<td>2 min at 200 °C</td>
<td></td>
</tr>
<tr>
<td>Etch back</td>
<td>Plasmaquest</td>
<td>XZSIOVIA+XZW</td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>2 etch back</em></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HSQ spin</td>
<td>Coater</td>
<td>SOP</td>
<td>1 min at 4K</td>
<td></td>
</tr>
<tr>
<td>Reflow</td>
<td>Hotplate1</td>
<td>SOP</td>
<td>2 min at 150 °C</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Hotplate2</td>
<td>SOP</td>
<td>2 min at 200 °C</td>
<td></td>
</tr>
<tr>
<td>Etch back</td>
<td>Plasmaquest</td>
<td>XZSIOVIA+XZW</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Via</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>EBL</td>
<td>Elionix</td>
<td>ZEP SOP</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Develop</td>
<td>Photo-wet-r</td>
<td>Xylene SOP</td>
<td>1 min</td>
<td></td>
</tr>
<tr>
<td>Resist check</td>
<td>AFM</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Via etch</td>
<td>Plasmaquest</td>
<td>XZSIOVIA</td>
<td>Etch until clear</td>
<td></td>
</tr>
<tr>
<td>Check</td>
<td>SEM</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Via Etch</td>
<td>PQ</td>
<td>XZSIOVIA</td>
<td>1 min</td>
<td></td>
</tr>
<tr>
<td>Resist strip</td>
<td>Photo-Wet-r</td>
<td>Water bath</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Resist strip</td>
<td>Asher</td>
<td>SOP</td>
<td>Until resist cleared</td>
<td></td>
</tr>
<tr>
<td><strong>Contact</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contact metal</td>
<td>Al₂O₃ remove</td>
<td>Acidhood</td>
<td>25 % TMAH 5 min</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Surface clean</td>
<td>Acidhood</td>
<td>HCl: H₂O= 1:3 1 min</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Metal sputter</td>
<td>AJA-TRL</td>
<td>XinW40nm200W 40 nm</td>
<td></td>
</tr>
<tr>
<td>Pads</td>
<td>EBL</td>
<td>Elionix</td>
<td>PMMA A8 SOP</td>
<td></td>
</tr>
<tr>
<td>Metal dep</td>
<td>eBeamFP</td>
<td>Zerorotation/Ti-Au</td>
<td>20nmTi/220nmAu</td>
<td></td>
</tr>
<tr>
<td>Lift-off</td>
<td>Photo-Wet-r</td>
<td>Acetone</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Etch</td>
<td>Plasmaquest</td>
<td>JQW1</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
### A.2 Recipes in The Process

<table>
<thead>
<tr>
<th>Adhesion Layer</th>
<th>Nanowire RIE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tool: sts-CVD/TRL</td>
<td>Tool: SAMCO/TRL</td>
</tr>
<tr>
<td>Deposit 30 nm Si$_3$N$_4$ on Si dummy</td>
<td>Chamber conditioning: 4 cycles of Recipe 91</td>
</tr>
<tr>
<td>Calibrate recipe to get 2-3 nm Si$_3$N$_4$ reliably (thickness measured with Filmetrics/TRL before/after short deposition)</td>
<td>Chamber conditioning: 4 cycles of Recipe 53</td>
</tr>
</tbody>
</table>
| Recipe: HFSINXIN | Recipe 53:
| Gas flow: | Gas flow: BCl$_3$/Ar/SiCl$_4$ = 3/11/0.3 SCCM |
| Gas temperature: | Necessary adjustment of gas flows needed |
| Substrate temperature: | Temperature: 250 °C |
|                     | ICP/Bias power: 20/160 W |
|                     | Rate: non-linear, 200 min for 2 min |

<table>
<thead>
<tr>
<th>Digital Etch</th>
<th>W Gate Sputtering</th>
</tr>
</thead>
<tbody>
<tr>
<td>Oxidation:</td>
<td>Recipe: XinW40nm200W</td>
</tr>
<tr>
<td>Tool: Asher/TRL</td>
<td>Loadlock pressure &lt; 10$^{-5}$ Torr</td>
</tr>
<tr>
<td>Power: 800 W</td>
<td>Main chamber pressure &lt; 10$^{-6}$ Torr</td>
</tr>
<tr>
<td>Time: 180 s</td>
<td>Strike step first at 100 W/ 20 mTorr</td>
</tr>
<tr>
<td>Oxide removal:</td>
<td>Deposition pressure: 3 mTorr.</td>
</tr>
<tr>
<td>H$_2$SO$_4$:H$_2$O = 1:1 at RT</td>
<td>Rate: 1.5 Å/s (calibration needed)</td>
</tr>
<tr>
<td>Time: 30 s</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>W Dry Etch</th>
<th>FOX Dry Etch</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tool: Plamsaquest/TRL</td>
<td>Tool: Plamsaquest/TRL</td>
</tr>
<tr>
<td>Recipe: XZW</td>
<td>Recipe: XZSIOVIA</td>
</tr>
<tr>
<td>Gas stabilization and plasma strike</td>
<td>Gas stabilization and plasma strike</td>
</tr>
<tr>
<td>Temperature: RT</td>
<td>Temperature: RT</td>
</tr>
<tr>
<td>Pressure: 60 mTorr</td>
<td>Pressure: 20 mTorr</td>
</tr>
<tr>
<td>ECR/DC bias: 150 W/ 50 V</td>
<td>ECR/DC bias: 300 W/ 50 V</td>
</tr>
<tr>
<td>SF$_6$ flow: 90 SCCM</td>
<td>CF$_4$ flow: 27 SCCM</td>
</tr>
<tr>
<td>Rate: 30 nm /min (calibration needed)</td>
<td>Rate: 13 nm /min (calibration needed)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Mo Contact Sputtering</th>
<th>Resist Strip</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tool: AJA-TRL/TRL</td>
<td>Cycle of solvents and O$_2$ plasma until resist cleared.</td>
</tr>
<tr>
<td>Recipe: XinMo40nmStrike</td>
<td>Solvent clean:</td>
</tr>
<tr>
<td>Loadlock pressure &lt; 10$^{-5}$ Torr</td>
<td>Tool: Photo-wet-r/TRL</td>
</tr>
<tr>
<td>Deposition chamber pressure &lt; 10$^{-6}$ Torr</td>
<td>NMP in water bath, on hotplate at 300 °C.</td>
</tr>
<tr>
<td>Loadlock pressure &lt; 10$^{-5}$ Torr</td>
<td>Time: 45 min</td>
</tr>
<tr>
<td>Main chamber pressure &lt; 10$^{-6}$ Torr</td>
<td>O$_2$ plasma:</td>
</tr>
<tr>
<td>Strike step first at 100 W/ 20 mTorr</td>
<td>Tool: Asher/TRL</td>
</tr>
<tr>
<td>Deposition pressure: 2 mTorr.</td>
<td>Power: 800 W</td>
</tr>
<tr>
<td>Deposition power: 120 W, no substrate bias.</td>
<td>Time: 3 min then 3 min waiting</td>
</tr>
<tr>
<td>Rate: 1 Å/s (calibration needed)</td>
<td></td>
</tr>
</tbody>
</table>
A.3 Reactive Ion Etching

III-V RIE throughout this work is performed in a *SAMCO RIE-200IP ICP* system with a heated chuck and backside He cooling. Samples cut into 1 cm × 1 cm pieces are loaded on a 6-inch ceramic carrier wafer without thermal grease. To avoid loading effects on smaller samples, dummy pieces are placed around the samples so that the total III-V area is roughly 1 cm × 1 cm. The thermal contact between wafer carrier and chuck is mediated by He, in order to maintain a constant sample temperature. At the beginning of every recipe, a He leak test is performed and the cleanliness of the backside of the carrier and the position of the carrier in the robotic arm are found to be important in order to pass the leak test.

Chuck temperature is raised from idle value, 40 °C to the target value, 250 °C, which normally takes 40 min to complete. Aside from the main etching step discussed in section 2.2.2, the complete recipe has 3 preceding steps, each serving a distinct purpose. Details are shown in the following table. During the first step, sample sits in the processing chamber for 15 min to reach thermal equilibrium without gas or power. The second step involves flowing identical gases as used in the main step but with no power, in order to stabilize the gas flows and the pressure. Without this step, pressure and gas flow rates can change when etching happens after RF powers are turned on, causing uncontrolled etch rate and profile. Right before the main etching step, a short strike step (5 s) is introduced with higher ICP power to ignite the plasma.

<table>
<thead>
<tr>
<th></th>
<th>Bias power/W</th>
<th>ICP power/W</th>
<th>BCl3</th>
<th>SiCl4</th>
<th>Ar /SCCM</th>
<th>Pressure/Pa</th>
<th>Time/s</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>900</td>
<td>-</td>
<td>900</td>
</tr>
<tr>
<td>2</td>
<td>-</td>
<td>-</td>
<td>7/0.55/7</td>
<td></td>
<td>0.2</td>
<td>30</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>160</td>
<td>100</td>
<td>7/0.55/7</td>
<td></td>
<td>0.2</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>160</td>
<td>20</td>
<td>7/0.55/7</td>
<td></td>
<td>0.2</td>
<td>As needed</td>
<td></td>
</tr>
</tbody>
</table>
A.4 Publication List

References


beyond 7nm technologies." in *Device Research Conference (DRC), 2014 72nd Annual*, 2014, pp. 133-134.


[77] K. M. Persson, M. Berg, M. Borg, W. Jun, H. Sjoland, E. Lind, and L. E. Wernersson, "Vertical InAs nanowire MOSFETs with $I_{DS} = 1.34 \, \text{mA/μm}$ and $g_m = 1.19 \, \text{mS/μm}$ at $V_{DS} = 0.5 \, \text{V}$." in Device Research Conference (DRC), 2012 70th Annual, 2012, pp. 195-196.


[79] K. Tomioka, M. Yoshimura, and T. Fukui, "Vertical In$_0.7$Ga$_{0.3}$As nanowire surrounding-gate transistors with high-k gate dielectric on Si substrate." in Electron Devices Meeting (IEDM), 2011 IEEE International, 2011, pp. 33.3.1-33.3.4.


[96] R. Bijesh, H. Liu, H. Madan, D. Mohata, W. Li, N. V. Nguyen, D. Gundlach, C. A. Richter, J. Maier, K. Wang, T. Clarke, J. M. Fastenau, D. Loubychev, W. K. Liu, V. Narayanan, and S. Datta, "Demonstration of In$_{0.9}$Ga$_{0.1}$As/GaAs$_{0.18}$Sb$_{0.82}$ near broken-gap tunnel FET with I$_{ON} =$740uA/um, G$_{m}=$700uS/um and gigahertz switching performance at V$_{DS}=0.5$V." in Electron Devices Meeting (IEDM), 2013 IEEE International, 2013, pp. 28.2.1-28.2.4.


[98] Y. Tao, T. T. James, A. A. Dimitri, and L. H. Judy, "Effects of substrate leakage and drain-side thermal barriers in In$_{0.53}$Ga$_{0.47}$As/GaAs$_{0.5}$Sb$_{0.5}$ quantum-well tunneling field-effect transistors."Applied Physics Express, vol. 7, no. 9, pp. 094201, 2014.


