University of California
Santa Barbara

Triple-heterojunction (3-HJ) TFETs Design and Fabrication for Low Power Logic

A dissertation submitted in partial satisfaction
of the requirements for the degree

Doctor of Philosophy
in
Electrical and Computer Engineering

by

Hsin-Ying Tseng

Committee in charge:

Professor Mark J.W. Rodwell, Chair
Professor Umesh K. Mishra
Professor Chris Palmstrøm
Professor Johanthan Klamkin

August 2021

The Dissertation of Hsin-Ying Tseng is approved.

Professor Umesh K. Mishra

Professor Chris Palmstrøm

Professor Johanthan Klamkin

Professor Mark J.W. Rodwell, Committee Chair

August 2021

Triple-heterojunction (3-HJ) TFETs Design and Fabrication for Low Power Logic

Copyright © 2021

by

Hsin-Ying Tseng

Curriculum Vitæ

Hsin-Ying Tseng

Education

2021	Ph.D. in Electrical and Computer Engineering, University of California Santa Barbara, Santa Barbara, CA
2016	M.S. in Electronics Engineering, National Tsing Hua University, Hsinchu, Taiwan
2014	B.S. in Electrical and Computer Engineering, National Tsing Hua University, Hsinchu, Taiwan

Publications

1. **H.Y. Tseng**, Y. Fang, W.J. Mitchell, A.A. Talor, M.J.W. Rodwell, "Atomic Layer Deposition of TiN/Ru Gate in InP MOSFETs," submitted to Applied Physics Letter.
2. C.Y. Chen, **H.Y. Tseng**, H. Ilatikhameneh, T.A. Ameen, G. Klimeck, M.J.W. Rodwell, M. Povolotskyi, "Doping Profile Engineered Triple Heterojunction TFETs With 12-nm Body Thickness," IEEE Transactions on Electron Devices **68**, 3104-3111 (2021)
3. A. Goswami, S.T. Brunelli, B. Markman, A.A. Taylor, **H.Y. Tseng**, K. Mukherjee, M.J.W. Rodwell, J. Klamkin, C.J. Palmstrøm, "Controlling facets and defects of InP nanostructures in confined epitaxial lateral overgrowth," Physical Review Materials **4**, 123403 (2020)
4. A. Goswami, S.T. Brunelli, B. Markman, D. Pennachio, **H.Y. Tseng**, S. Choi, A.A. Taylor, J. Klamkin, M.J.W. Rodwell, C. Palmstrom, "Effects of growth parameters on faceting and defects in confined epitaxial lateral overgrowth," Bulletin of the American Physical Society 65 (2020)
5. **H.Y. Tseng**, Y. Fang, S. Zhong, M.J.W. Rodwell, "InP MOSFETs Exhibiting Record 70 mV/dec Subthreshold Swing," 2019 Device Research Conference (DRC), 183-184 (2019).
6. Y. Fang, **H.Y. Tseng**, M.J.W. Rodwell, "$W_E = 100$nm InP/InGaAs DHBT with Self-aligned MOCVD Regrown p-GaAs Extrinsic Base Exhibiting 1 $\Omega - \mu m^2$ Base Contact Resistivity." 2019 Device Research Conference (DRC), 179-180 (2019).
7. S.T. Brunelli, A. Goswami, B. Markman, **H.Y. Tseng**, M.J.W. Rodwell, C. Palmstrøm, J. Klamkin, "Horizontal Heterojunction Integration via Template-Assisted Selective Epitaxy," Crystal Growth and Design **19**, 7030-7035 (2019)

8. S.T. Brunelli, B. Markman, A. Goswami, **H.Y. Tseng**, S. Choi, C. Palmstrøm, M.J.W. Rodwell, J. Klamkin, "Selective and confined epitaxial growth development for novel nano-scale electronic and photonic device structures," Journal of Applied Physics **126**, 015703 (2019)
9. Y. Fang, **H.Y. Tseng**, H.W. Chiang, B. Markman, M.J.W. Rodwell, "90nm Emitter Width InGaAs/InP DHBTs with DC Current Gain $\beta > 40$," Lester Eastman Conference (2018)
10. J. Wu, Y. Fang, B. Markman, **H.Y. Tseng**, M.J.W. Rodwell, "$L_g = 30$ nm InAs Channel MOSFETs Exhibiting $f_{max} = 410GHz$ and $f_t = 357GHz$," IEEE Electron Device Letters **39**, 472-475 (2018)
11. **H.Y. Tseng**, W.C. Yang, P.Y. Lee, C.W. Lin, K.Y. Cheng, K.C. Hsieh, K.Y. Cheng, C-H Hsu, "GaN Schottky Diodes with Single-Crystal Aluminum Barriers Grown by Plasma-assisted Molecular Beam Epitaxy," Applied Physics Letters **109**, 082102 (2016)
12. W.C. Yang, P.Y. Lee, **H.Y. Tseng**, C.W. Lin, Y.T. Tseng, and K.Y. Cheng, "Mg Incorporation in GaN Grown by Plasma Assisted Molecular Beam Epitaxy at High Temperatures," Journal of Crystal Growth **439**. 87-92 (2016)
13. **H.Y. Tseng**, W.C. Yang, P.Y. Lee, C.W. Lin, K.C. Hsieh, K.Y. Cheng, "Schottky Barrier Diodes with In-situ Grown Single Crystal Aluminum on GaN by Plasma assisted Molecular Beam Epitaxy," PCSI 43 (2016)

Abstract

Triple-heterojunction (3-HJ) TFETs Design and Fabrication for Low Power Logic

by

Hsin-Ying Tseng

Power/heat density becomes excessive in VLSI. Supply voltage in the current 7/10 nm process has been scaled to 0.7 V for logic devices to reduce dynamic power loss ($P_{dynamic} \propto C_{wire}{V_{dd}}^2$). The low supply voltage (V_{dd}), on the other hand, poses a challenge to static power loss ($P_{static} = I_{off}V_{dd}$), because off-current ($I_{off}$) is limited by the Boltzmann tail of the Fermi-Dirac distribution, corresponding to a minimum 60 mV/dec subthreshold swing SS at room temperature. To scale V_{dd} to 0.3 V while maintaining a 6 decades on/off ratio, transistors need to have a subthreshold slope (SS) $<$ 60 mV/dec. The tunneling field-effect transistor (TFET) is one technology that holds such promise as electron tunneling is not constrained by thermal statistics. Current state-of-the-art TFETs, however, have low on current (I_{on}) $\approx$ 10 $\mu A/\mu m$ at V_{dd} = 0.3 V that precludes their applications in actual circuits. In this work, an InP-based triple heterojunction (3-HJ) TFET design is proposed to significantly increase I_{on} by exploiting a minimized tunneling distance, as well as resonant states in the 3-HJ. PNPN doping profile is employed in 3-HJ TFETs to eliminate the need for an ultra-thin body to ease processing. The simulated I_{on} reaches $\approx$ 300 $\mu A/\mu m$ at $V_{dd} = 0.3$ V.

To realize 3-HJ TFETs, three generations of device structures are tried. The main challenges come from the incompatibility between process modules within a small/limited process window. Among all device structures, the qualitative high-k/channel interface and gate metallization is important for high TFET performances. ALD post-metal H_2 annealing is developed, and different gate metallization processes, including physical

vapor deposition (PVD) of W, Ni, Al, Ti, and atomic layer deposition (ALD) of Ru, TiN, and TiN/Ru, are investigated. With the high-quality ALD TiN/Ru gate as well as the high-k/InP interface, a record low averaged SS of 68 mV/dec for long gate length devices, and a high peak transconductance of 0.75 $mS/\mu m$ at $V_{DS} = 0.6$ V in an 80 nm gate length InP channel planar MOSFET are demonstrated.

Vertical InP channel MOSFETs and 3-HJ InGaAs/GaAsSb/InAs/InP TFETs with conformal TiN/Ru gate are fabricated using 3rd generation top-down planarization process. A high peak transconductance of 0.42 $mS/\mu m$ in a 50 nm gate length vertical MOSFET at $V_{DS} = 0.6$ V, which is comparable with planar MOSFETs at the same gate length, proves the process. The first demonstrated 3-HJ TFETs (having a gate length of 30 nm) show a strong short channel effect, and exhibiting SS of >80 mV/dec at $V_{DS} = 0.1$ V, and I_{on} of ≈ 2 $\mu A/\mu m$ at $V_{DS} = 0.3$ V. The high SS and low I_{on} comparing to simulation could be resulted from short channel effect, gate misalignment, reduced electric field at tunneling junction due to fringing field, and rough heterojunction interfaces given by a problematic epi. Negative differential resistance at $V_{DS} = 0.9$ V, and high drain current under reverse drain bias are observed. A heterojunction interface trap-assisted tunneling model is used to explain those behaviors in 3-HJ TFETs.

Contents

Curriculum Vitae **iv**

Abstract **vi**

List of Figures **x**

List of Tables **xviii**

1 Introduction **1**
1.1 Background 1
1.2 Power Density Issue in VLSI 3
1.3 Steep Slope Transistors 6
1.4 Tunneling Field Effect Transistors 8
1.5 Ferroelectric field effect transistors(this section needs to be modified to specify the importance of high on current in tfet 9
1.6 Outline 10

2 Design of TFETs **12**
2.1 Principles of Operation–Zener Tunneling 12
2.2 Tunneling Probability Maximization 17
2.3 Electrostatics of TFETs 23
2.4 Challenges and Potential Issues of TFETs 27

3 High On-Current 3-HJ TFETs with PNPN Doping Profile **38**
3.1 Doping Profile Engineering 39
3.2 Heterojunction Engineering 41
3.3 3-HJ TFETs with Resonant Enhancement 42
3.4 Thick Body– Doping Profile Engineering on InGaAs/GaAsSb/InAs/InP TFET 54
3.5 Design Finalization and Potential Issues 62

4 Development of Conformal ALD TiN/Ru Gate Metallization **70**
4.1 Overview 70
4.2 MOSCAPs with Different Gate Metallization 71
4.3 InP Channel Planar MOSFETs with ALD TiN/Ru Gate 90
4.4 Low Temperature Conformal ALD Ru Growth 104

5 Process Development and First Demonstration of 3-HJ TFETs **107**
5.1 Structure Overview– From Planar to Vertical TFETs 108
5.2 Process Development Towards 3-HJ TFETs 114
5.3 Fabrication of Vertical Planarization FETs 132
5.4 Vertical MOSFETs and 3-HJ InGaAs/GaAsSb/InAs/InP TFETs 144

6 Conclusion **159**
6.1 Summary 159
6.2 Future work 161

A Planar and Vertical FETs Process Flow **162**
A.1 Overview 162
A.2 Planar MOSFET Process Flow 172
A.3 Vertical FET Process Flow 173

Bibliography **175**

List of Figures

1.1 (a) Cross-sectional TEM of Intel's 32 nm node P-channel MOSFETs [1, 2]. (b) Schematic finfet structure of Intel's 22 nm node [3]. (c) Cross-sectional TEM of TSMC's 7 nm node. 2
1.2 Cross-sectional TEM of IBM Research Alliance's 5nm GAA transistors [4]. 3
1.3 Schematic diagram showing the operation of CMOS inverter, from 1 to 0 (left) and 0 to 1 (right). 5
1.4 (a) Supply voltage vs. technology node by Microwind [5] and (b) Microprocessor clock frequency vs. time created by Michael L Rieger [6]. . . . 5
1.5 Schematic I_D-V_G characteristics of MOSFETs (with minimum SS = 60 mV/dec) and steep slope transistors (with SS < 60mV/dec). 6
1.6 Schematic band diagram of N-channel MOSFETs (a) and TFETs (b) at on and off states. 8
1.7 Si SOI MOSFET using ferroelectric HfZrO2 in gate dielectric by UC Berkeley research group [10]. An improved short channel behavior is seen comparing to HfO_2 control sample. 10

2.1 Schematic energy band structure of (a) a zener diode at on-state. (b) a N-TFET at off-state, and (c) a N-TFET at on-state. 13
2.2 Schematic band structure of a P^+/N^+ zener diode under reverse bias (a) and the equivalent triangular energy barrier for electrons tunneling from the valence band into conduction band at a certain energy (b). 14
2.3 The calculated zener tunneling current density per 1 V of reverse bias using eq. 2.7 for homogeneous Si and InGaAs PN junctions. 17
2.4 Schematic diagram of thin body double-gated TFETs in different device orientations (a) and their energy k-space contours (b). Courtesy of Purdue team. 19
2.5 Band structures for conduction band of InAs and valence band of GaSb in (001) (a) and $(1\bar{1}0)$ (b) confinement, and energy band diagram (c) and transmission vs. energy (d) for a InAs/GaSb TFET at the design orientation. [23] 20

2.6 Schematic energy band structure of a GaSb/InAs tunneling junction. . . 21
2.7 Energy band diagrams of InGaAs homojunction TFET and triple-heterojunction TFET at t_{body} of 3 nm (a) and their tunneling probability vs. energy at V_{GS} = 0.3 V. Courtesy of Chin-Yi Chen. 22
2.8 Schematic double-gated TFET structure (a) and the calculated channel potential vs. distance with different cavity thickness D (b). 24
2.9 (a) Double-side gated P-GaSb/InAs TFET structure and its (b) band diagram, (c) tunneling probability vs. energy at V_g of -0.7 V and -1 V, and (d) the I_D - V_G characteristic. Courtesy of Chin-Yi Chen. 29
2.10 Summary of 5 ways to improve ambipolar leakage. (a) A normal P-i-N TFET. (b) Hetero-gate TFET. (c) Gate/drain underlap TFET. (d) Asymmetric doping TFET. (e) Gate/drain overlap TFET. (f) Heterojunction TFET. 30
2.11 Energy band diagram of (a) a normal TFET and (b) a gate/drain overlap TFET. L_{ov} is the overlap distance, which is 30 nm in this case. [28] . . . 31
2.12 Energy band diagram comparing different solutions to ambipolar leakage problem. HG is hetero-gate and GDU is gate/drain underlap. [31] 32
2.13 Gate misalignment analysis for GaSb/InGaAsSb/InAs nanowire TFETs done by the research team in Lund University [33]. (a) The structure of nanowire TFETs, and (b) I_D - V_G simulation fitting to the experimental gate/source overlap condition. (c) The simulated transfer characteristics of TFETs having aligned gate and 5 nm gate/source overlap and underlap gates. ©2018, ISTE® (International Society for Technology in Education), iste.org. All rights reserved. 34
2.14 Schematic band structures showing TAT given by (a) heterojunction interface traps and (b) oxide/semiconductor interface traps in TFETs. Note that TAT happens two-way (source to channel and channel to source), dependent on the bias. 36

3.1 (a) Schematic structure of a PNPN doping TFET, and its band diagram at on and off states. (c) I_D - V_G characteristics of a PNPN TFET and conventional P-i-N TFET. (d) Band diagram of PNPN TFETs with pocket doping of $4 \times 10^{19} cm^{-3}$ and pocket thickness of 3 nm to 13 nm. [28] . . . 40
3.2 Schematic structures of 1-HJ TFET and 3-HJ TFET (a), their energy band diagram at tunneling junction (b), and the energy vs. tunneling probability plots (c). [45] 42
3.3 Schematic band structure of a 3-HJ TFET that has layers specified from left: source barrier, source well, channel well, channel barrier. By incorporating three heterojunctions, the tunneling interface is split into junctions 1, 2, and 3. The tunneling energy window (marked grey) is in between the conduction band edge of the channel well and the valence band edge of the source well. 43

3.4 Band alignment (bandgap heaven) of widely used III-V materials. [48] . . 45
3.5 Energy band diagrams and energy vs. tunneling probability of 3-HJ TFETs with (a) aligned bound states in channel and source well, and (b) misaligned quantum state energies. © 2021 IEEE 49
3.6 (a) Band diagrams of InAlAs/GaAsSb/InAs/InP 3-HJ TFETs with source well (GaAsSb) thickness ranging from 1.8 nm (case 1), 2.4 nm (case 2), to 4 nm (case 3). (b) and (c) the I_D - V_G characteristics and energy vs. tunneling probability for three cases. Courtesy of Chin-Yi Chen. 50
3.7 Energy band diagrams of InAlAs/GaAsSb/InAs/InP 3-HJ TFET (case 3) at V_G of 0 V, -0.3 V, -0.6 V (a), and the density of state distribution showing bound state energy alignment at V_G of 0.3 V, 0 V, -0.3 V. Courtesy of Chin-Yi Chen. 51
3.8 (a) Full band diagram of proposed high on-current InGaAs/GaAsSb/InAs/InP 3-HJ TFET design and (b) the structure design in the simulation set up. Channel thicknesses are set to 4 nm, and high-k thickness is 3.2 nm with a dielectric constant of 15. 53
3.9 Band diagram of proposed 3-HJ TFET at on state (V_G = 0.3 V). The resonant state in the InAs channel well aligns to the bound states in the GaAsSb source well and lies in the tunneling window (marked in between dash lines). © 2021 IEEE 54
3.10 Electrostatics of wave functions in thin body and thick body wells. Electrons are highly confined in a thin well and spread over the whole body in a thick well. A P body doping supplements the loss of confinement, thus improves electrostatics. 55
3.11 Band structures and resonant state alignment of InGaAs/GaAsSb/InAs/InP TFETs with (a) P-i-N at t_{body} = 4 nm, (b) PNPN at t_{body} = 4 nm, (c) P-i-N at t_{body} = 12 nm, (d) PNPN at t_{body} = 4 nm. ©2021 IEEE 57
3.12 Electric field vs. distance for InGaAs/GaAsSb/InAs/InP TFETs with different body thickness and doping profiles. ©2021 IEEE 58
3.13 Energy band diagrams of InGaAs/GaAsSb/InAs/InP TFETs with InP channel doping of (a) U.I.D, (b) P-1× 10^{16}, (c) P-5× 10^{18}, (d) P-2× 10^{19} cm^{-3} at $V_{GS}-V_{TH}$ = 0.3 V. ©2021 IEEE 59
3.14 (a) Equal potential contours of 3-HJ TFETs with InP doping density of (a) U.I.D, (b) P-1× 10^{16}, (c) P-5× 10^{18}, (d) P-2× 10^{19} cm^{-3}. (b) I_D - V_G curves and (c) SS vs. I_D plots for 4 cases. ©2021 IEEE 60
3.15 Energy band diagram of InGaAs/GaAsSb/InAs/InP TFETs at (a) body edge (near to channel/high-k interface) and (b) body center with different InP channel doping. ©2021 IEEE 62
3.16 Band diagrams of InGaAs/GaAsSb/InAs/InP PNPN TFETs in bulk condition at different V_{DG} with InP channel of (a) 30 nm P-5× 10^{18} cm^{-3}, and (b) 15 nm U.I.D. + 15 nm P-5× 10^{18} cm^{-3}. (c) Ambipolar leakage current vs. V_{GD} for 2 channel designs. 64

3.17 Schematic TFET structures for gate alignment analysis in 3-HJ TFETs. 67
3.18 Energy band diagrams at on-state comparing gate aligned case with (a)overlap/underlap by 2 nm cases, (b) underlap by 4 nm cases, and (c) overlap by 4 nm cases. Courtesy of Purdue team. 68
3.19 I_D - V_G characteristics of gate aligned, gate/source overlap by 2/4 nm (+2/+4 nm), gate/source underlap by 2/4 nm (-2/-4 nm) 3-HJ TFETs. Courtesy of Purdue team. 68

4.1 C-V characteristics of InGaAs channel MOSCAPs using thermal evaporated Ni/Au gate with (a) as-deposited, (b) 3 minutes of ALD post gate H_2 annealing, (c) 9 minutes of ALD post gate H_2 annealing at 300°C. . . 74
4.2 C-V characteristics of InGaAs channel MOSCAPs using thermal evaporated Ni/Au gate with (a) 15 minutes of ALD post gate H_2 annealing at 300°C, (b) 15 minutes of ALD post gate H_2 annealing at 325°C, (c) 30 minutes of ALD post gate H_2 annealing at 350°C. 76
4.3 C-V characteristics of InGaAs channel MOSCAP using sputter W gate. Sputter W: power: 100 W, pressure: 3 mtorr. ICP W etch: SF_6/Ar: 20/5 sccm, power: 600/200 W, pressure: 1 Pa. 77
4.4 C-V characteristics of InGaAs channel MOSCAP using sputter W gate. (a) W stacks: 20/20nm with pressure: 15/3 mtorr. ICP W etch: SF_6/Ar: 20/5 sccm, power: 600/50 W, pressure: 1 Pa. (b)W stacks: 40/20nm with pressure: 15/3 mtorr. ICP W etch: SF_6/Ar: 20/5 sccm, power: 600/50 W, pressure: 5 Pa. 77
4.5 Cross-sectional SEM images of the test structures that have blanket W sputtered at a pressure of (a) 3 mtorr and (b) 15 mtorr. Red arrows indicate sputtered W layers. 79
4.6 C-V characteristics of InGaAs channel MOSCAP using 25 nm TiN gates with (a) as-grown and (b) post-metal H_2 annealing at 300°C for 30 minutes. 81
4.7 C-V characteristics of InGaAs channel MOSCAP using 30 nm Ru gate. Blue and red curves are measured at 1k and 1M Hz, respectively. 82
4.8 Threshold voltages of different gates vs. metal work functions. 83
4.9 C-V characteristics of InP channel MOSCAPs measured at 1k - 1M Hz using (a) thermally evaporated Ni/Au gate and (b) Ru gate. 85
4.10 C-V characteristics of InP channel MOSCAPs using (a) TiN (1 nm)/Ru gate and (b) TiN (2 nm)/Ru gate. 88
4.11 (a) First sweep and (b) second sweep of C-V characteristics of InP channel MOSCAPs using TiN/Ru gate. 89
4.12 Planar MOSFET fabrication process before gate metallization 94
4.13 Planar MOSFET fabrication process after gate metallization 95
4.14 (a) The as-grown schematic diagram showing TiN/Ru gate converage on mesa edges. (b) The surrounding TiN/Ru gate metal (over mesa edges) remains after gate patterning. 96

4.15 (a) SEM images of an L_g=30 nm planar MOSFET using thermal Ni/Au gate and © 2019 IEEE (b) schematic top view MOSFET layout. 97
4.16 Transfer charateristics of 40 nm, 100 nm, and 1 μm-L_g MOSFETs using Ni/Au gates. © 2019 IEEE 98
4.17 Output charateristics of 40 nm, 100 nm, and 1 μm-L_g MOSFETs using Ni/Au gates. © 2019 IEEE 98
4.18 (a) SS vs. L_g at V_{DS} =0.1 and 0.6 V; (b)peak g_m vs. L_g at V_{DS} =0.6 V; (c) DIBL vs. L_g at V_{TH} =1 $\mu A/\mu m$ of MOSFETs using Ni/Au gates. © 2019 IEEE 99
4.19 The STEM image of an L_g = 30 nm MOSFET with TiN/Ru gate. The inner plot in (b) is the HAADF-STEM image highlighting the structure at InP channel. Layer 1, 2, and 3 represent 1 nm AlO_xN_y, ≈2.5 nm ZrO_2 and ≈2 nm TiN, respectively. 100
4.20 Transfer charateristics of 30 nm, 80 nm, and 200 nm-L_g MOSFETs using TiN/Ru gates. 102
4.21 Output charateristics of 30 nm, 80 nm, and 200 nm-L_g MOSFETs using TiN/Ru gates. 102
4.22 (a) Peak g_m at V_{DS} =0.6 V vs. L_g and (b) SS vs. L_g at V_{DS} =0.1 V (b) of MOSFETs using thermal evaporated Ni/Au gates, ALD TiN/Ru gates and ALD Ru gates. 103
4.23 Schematic diagram explaining the process of Ru growth using (a) traditionally AB-type growth (Ru precursor-O_2) with a RuO_x starting surface, and (b) the ABC-type growth (Ru precursor-O_2-H_2) with a Ru starting surface, and (c) the sequence of ABC-type reactants exposure. (a) and (b) Reprinted with permission from [92]. © 2015 American Chemical Society. 105
4.24 TEM images of vertical MOSFETs using TiN/Ru gates in which Ru layer is grown using (a) AB-type growth at 300°C and (b) ABC-type growth at 250°C. 106

5.1 Schematic gated mesa edge vertical TFET structures (a) with pyramid-shape, and (b) with channel perpendicular to tunneling junction. 108
5.2 Schematic vertical TFET structures using top down approach with (a) drain contact first and (b) drain contact last. 109
5.3 Schematic vertical nanowire TFET structure (a) and multi-coreshell nanowire with source at the bottom (b). 111
5.4 Schematic lateral TFET structure using direct wafer bonding and source/drain selective regrowth. 112
5.5 Schematic lateral TFET structures using TASE technique with (a) all layers grown in confined box, and (b) only U.I.D. channel grown in confined box and source/drain selective regrown. 113
5.6 116

5.7 InP CELO grown at 600°C on (001) InP substrate with template sitting along [110]. (a) and (b) top view TEM images at two ends of the crystal. (c) and (d) cross-sectional TEM images cutting from red dash lines in (a) and (b). (e) Cross-sectional TEM images focusing on defects. (f) Schematic structure showing the formed growth planes in the box. [100] Copyright © 2020 by American Physical Society. All rights reserved. . . 117
5.8 (a) Cavity bow vs. cavity length, and (b) cavity morphology before and after growth. Process A and R stand for template fabrication using thin-film deposited and resist-based sacrificial and top dielectric layers, respectively. Reproduced from [101], with the permission of AIP Publishing. 118
5.9 Top view (a) SEM and (b) OM images of parasitic nucleation on CELO template after MOCVD growth. Courtesy of Brian Markman, Simone Šuran Brunelli, Aranya Goswami. 119
5.10 Top view TEM images of InP CELO growth on (100) InP substrate at (a) 630°C, (b) 600°C, (c) 570°C, (d) 550°C. Template sits along $[0\overline{1}\overline{1}]$. [100] Copyright © 2020 by American Physical Society. All rights reserved. . . 120
5.11 (a) Top view schematic anchor-shape vertical structure and (b) 3 different cuts in cross-sections used in explaining the process flow for anchor-shape vertical FETs (that will be shown later). 121
5.12 Front end process flow of anchor-shape vertical FETs. 123
5.13 Isolation process of anchor-shape vertical FETs. 124
5.14 Back end wiring process flow of anchor-shape vertical FETs. 124
5.15 SEM image of the test structure after isolation, having fin width of 80 nm, fin length of 1 μm, and L_{sp} of 100 nm. L_{sp} is the spacing between pads and the edge of mesa. 125
5.16 SEM images of different test structures with (a) fin length of 2, 3, and 5 μm at $L_s p$ of 100, 150, 500 nm, respectively, and (b) the multi-finger design at a fin length of 2 μm and $L_s p$ of 200 nm. 126
5.17 Cross-sectional TEM images of 8th batch anchor-shape vertical MOSFETs along A cut (according to fig. 5.11 (b)). The inner plot shows the EDX analysis of elemental Pt. 127
5.18 Cross-sectional TEM images of (a) 8th, (b) 9th, and (c) 10th batch anchor-shaped vertical MOSFETs along C cut, A cut and C cut, respectively (acoording to fig. 5.11 (b)). 128
5.19 Top view schematic structure of vertical planarization FETs. The yellow area indicates measurement pads. 129
5.20 Front-end process flow for vertical top-down planarization FETs. 131
5.21 Back-end wiring process flow for vertical top-down planarization FETs. . 132
5.22 30°tilted SEM images of TiW/Mo stack after dry etch and hard mask lifting-off. (a) 60 nm fin width design at the center die, and (b) 80nm fin width design at the edge die of a 1.5”×1.5” sample. 134

5.23 (a) Schematic diagram of device (fin) orientation and the resulted facets and sidewall angles after HCl:H_3PO_4=1:4 InP wet etch along A cut ([01$\bar{1}$]) (b), and along B cut ([011]) 136
5.24 Cross-sectional SEM image of fins after 10 nm InGaAs etch following a 2 minutes of 30 nm InP low temperature etch using (a) 30 nm HSQ mask, and (c) Cr/SiO_2/SiN_x/TiW/Mo (30/80/40/130/20 nm) drain stack as hard masks. (b) shows the structures in (a) with a reduced dimension after 3 cycles of digital etch. 138
5.25 Tilted SEM image of vertical FET after finishing front end process 140
5.26 The 30°tilted SEM images of the device having the top of fin and gate/source posts expose after BCB ashing. (a) shows the panoramic view, while (b) focuses on the gate post and (c) zooms into TiW fin. 141
5.27 Cross-sectional SEM images showing micro-trenching effect in BCB (a) without, and (b) with 1 minutes of Cl_2/O_2 plasma exposure. 142
5.28 30°tilted SEM images of vertical FETs (a) after BCB planarization, (b) after Ru dry etch following TiN/high-k/SiN_x sidewall removal, and (c) after Ru wet etch following TiN/high-k/SiN_x sidewall removal. 143
5.29 Transfer (a) and output (b) characteristics of vertical MOSFETs with TiN/Ru gate. Low ratio between gate length and fin width (50 nm: 90 nm) results in a poor gate control and shows a strong short channel behavior. 145
5.30 The schematic (a) cross-sectional and (b) top view vertical MOSFET structure, and (c) the cross-sectional STEM image of the vertical MOSFET with TiN/Ru gate cutting along A and A', as indicated in (b). The P-InP channel length is 50 nm in this study, while the TiW/Mo drain metal stack is as thick as 520 nm to facilitate contacting the drain. The inner plot in (c) is the image in a large field. 146
5.31 (a)-(d) Schematic cross-section diagrams explaining what happens during the back end wet etches. (e) OM images of vertical MOSFETs before (left) and after (right) back end Ru wet etch. As indicated by the red arrow, the color of gate metal (closed to fin) fades after Ru wet etch, indicating Ru undercut is extended to the planar regime. 147
5.32 Cross-sectional STEM-EDX mapping of elemental Ru, Ti, Si, In, P, W, Ga, and As on working MOSFETs (that measures the data in fig. 5.29). 148
5.33 Cross-sectional STEM-EDX mapping of elemental Ru, Ti, Si, In, P, W, Ga, and As on opened gate MOSFETs with high drain resistance at the same time. 148
5.34 Transfer characteristics of 3-HJ TFETs having estimated body thickness of 100 nm to 170 nm at tunneling junction. 150
5.35 Output characteristics of 3-HJ TFETs having estimated body thickness of 100 nm to 170 nm at tunneling junction. Red dash lines are for the guidance of the degree of short channel effect. 151

5.36 SEM images in the fabrication of 3-HJ TFETs (a) after InP wet etching, and (b) after finishing cycles of digital etch prior to high-k deposition. The exposed surface is InAs and GaAsSb in (a) and (b), respectively. 153
5.37 Comparison of the transfer and output characteristics of InP channel (a) MOSFET and (b) 3-HJ TFET at similar bias condition. 154
5.38 Source current vs. V_{DS} with a floating gate (a), and the $I_D - V_D$ characteristics of 3-HJ TFETs under negative drain bias and V_{GS} of 0 - 0.3 V in (a) linear and (b) log scale. 155
5.39 Schematic band diagram of thick body 3-HJ TFETs near to high-k/semiconductor interface without (top) and with (bottom) heterojunction interface traps under different bias conditions. 156
5.40 Schematic band diagram of thick body 3-HJ TFETs in the body without (top) and with (bottom) heterojunction interface traps under different bias conditions. 157

List of Tables

2.1 Bandgap energy and effective mass for semiconductor materials. InGaAs listed has composition of 53% of In and 47% of Ga. [17][18][19][20][21][22] 18

2.2 Tunneling probability vs. (a) effective mass, (b) energy barrier, and (c) tunneling distance. 23

3.1 Approximated quantum well thickness given by an effective mass m_{eff} and a first quantum state energy E_1. Numbers in the table are in nm. 47

3.2 Confined bandgap (E_g) and valence band offset (ΔE_v) with respect to valence band of InGaAs at a t_{body} = 4 nm. 53

3.3 Confined bandgap (E_g) and valence band offset (ΔE_v) with respect to valence band of InGaAs at a t_{body} = 12 nm. 56

3.4 Summary of I_{on} at $V_{GS}-V_{TH} = V_{DS} = 0.3$ V for InGaAs/GaAsSb/InAs/InP TFETs with different body thickness and doping profiles. 58

3.5 Summary of I_{on} at $V_{GS}-V_{TH} = V_{DS} = 0.3$ V for InGaAs/GaAsSb/InAs/InP TFETs with InP channel doping of (a) U.I.D, (b) P-1$\times 10^{16}$, (c) P-5$\times 10^{18}$, (d) P-2$\times 10^{19}$ cm^{-3}. 61

3.6 Finalized epi structure of InGaAs/GaAsSb/InAs/InP 3-HJ PNPN TFET. 65

4.1 The properties of sputter W gates under different growth pressure and substrate height. 80

4.2 The amount of D_{it} versus sputter W stack design that has 1 nm, 5 nm, and 20 nm of high pressure initiation layer. 80

4.3 Summarized properties of different gate metallizations on ZrO_2/AlO_xN_y with InGaAs channel 83

4.4 Process flow for TiN/Ru gate formation 88

4.5 Series of experiments testing the effect of chemical surface treatments on native oxide removal and the surface properties after treatments. 92

4.6 Summary of long gate length SS and peak g_m for MOSFETs using thermally evaporated Ni/Au gates, ALD TiN/Ru gates, and ALD Ru gates. . 103

5.1 Summary of most commonly seen TFET structures 114

5.2 The dimension of TiW foot after TiW dry etch using different substrates. The measurements are done at the center die of sample pieces. 134

A.1 MO sources used in this study 163
A.2 MOCVD growth parameters 163
A.3 ICP#1 Cr and TiW etch recipe 165
A.4 ICP#1 low power SiN_x etch recipe 166
A.5 ALD high-k and TiN deposition parameters 169
A.6 ALD Ru (with H_2) deposition parameters 169
A.7 ALD Ru (without H_2) deposition parameters 169
A.8 ICP#2 Ru dry etch recipe 170
A.9 ICP#1 BCB ash recipe 171

Chapter 1

Introduction

1.1 Background

In 1965, Gordon Moore predicted that components per unit area on microchips doubles around every two years. Higher computational speed and capabilities while decreasing in cost were expected, along with an increase in density. This well-known Moore's law is proved to be accurate for decades and has been used to set targets and long-term goals in the semiconductor industry. Meanwhile, increasing transistors density is metal oxide semiconductor field-effect transistors (MOSFETs) and other passive components scaling.

MOSFETs suffer from the short channel effect, i.e., drain-induced barrier lowering (DIBL) when the gate length is shorter. This phenomenon is owing to the degree of channel potential changes by drain bias instead of gate bias. To prevent short channel behavior, it is needed to keep a good electrostatic; a reasonable gate control over the channel. The solution is scaling body thickness and gate oxide thickness as with scaling down transistors physical gate length. However, the scaling of gate oxide thickness is

restricted. Oxide leakage current is inverse proportional to oxide thickness. Minimum gate oxide thickness has been set to prevent gate leakage current. Therefore, the limit of minimum gate length is set.

High-k gate metal gate has been incorporated into CMOS technology to improve electrostatic since 45 nm technology node. Fig. 1.1 (a) shows the cross-sectional TEM of Intel's 32 nm node P-channel MOSFETs incorporating HfO_2 [1, 2]. HfO_2 has 0.47 nm EOT, indicating a much higher gate capacitance at a given oxide thickness than SiO_2. However, scaling of traditional MOSFET beyond 28nm technology node again becomes challenging due to a degraded gate control on a short channel. High impurity scattering and dopant fluctuation resulted from a highly doped substrate reduced drive current and yield, while insufficiently doped substrate leads to subthreshold leakage.

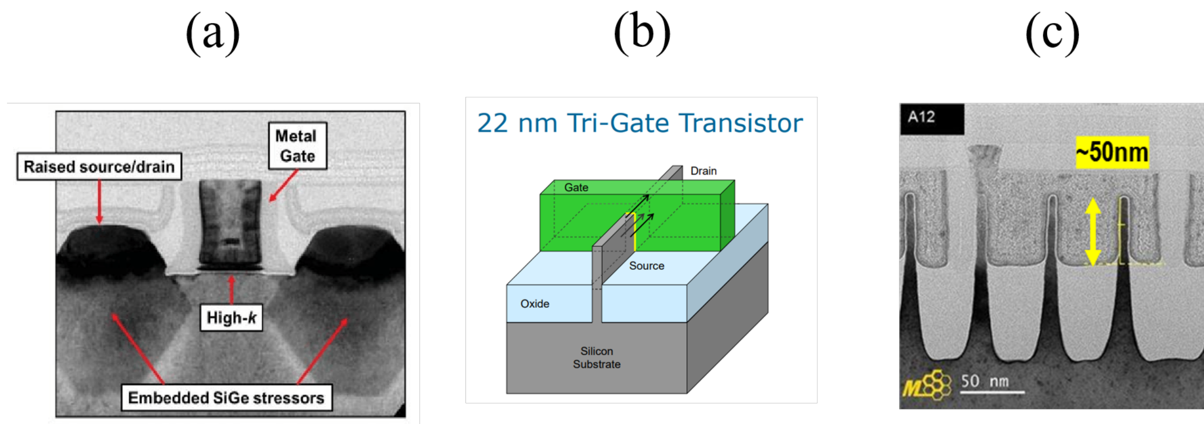


Figure 1.1: (a) Cross-sectional TEM of Intel's 32 nm node P-channel MOSFETs [1, 2]. (b) Schematic finfet structure of Intel's 22 nm node [3]. (c) Cross-sectional TEM of TSMC's 7 nm node.

Chenming Hu proposed the finfet, which flipped the planar channel 90 degrees and formed a thin fin structure. Fig. 1.1 (b) shows the schematic finfet structure of Intel's 22 nm technology node [3]. Double-side gated finfet has improved electrostatic comparing to planar MOSFETs. It has a benefit from the unintentionally doped channel, which prevents impurity scattering in highly scaled planar MOSFETs. In addition, electrons in

the channel are fully depleted at off-state, which results in a low off-state leakage current. A higher drive current and lower subthreshold leakage can be achieved in finFETs by managing fin height and fin width.

Scaling of finfets is done by increase fin height, which is gate width, and decrease fin spacing, which increases the number of parallel gates per unit length. Fig. 1.1 (c) shows the TEM image of TSMC's 7 nm technology. Fin height and fin spacing are both $\approx$ 50 nm. This technology corresponds to $\approx$ 200 MT/mm^2. Further scaling Si CMOS leads to the path to gate-all-around (GAA) fets. Fig. 1.2 shows the cross-sectional TEM image of IBM research alliance's 5 nm GAA transistors [4]. The number of transistors per unit area still increases by packing denser instead of physically scaling transistors. Thus, Moore's law is alive and well.

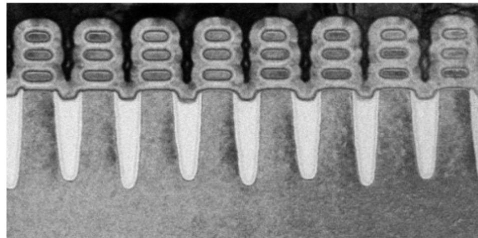


Figure 1.2: Cross-sectional TEM of IBM Research Alliance's 5nm GAA transistors [4].

1.2 Power Density Issue in VLSI

While Moore's scaling law keeps going, power consumption became a dominant constraint in nanoscale CMOS technology. Power density becomes excessive as transistors get denser. The power loss in CMOS can be divided into two parts: dynamic power and

static power, as shown below:

$$P_{dynamic} = f \times C_{wire} \times {V_{DD}}^2 \quad (1.1)$$

$$P_{static} = I_{leak} \times V_{DD} \quad (1.2)$$

Fig. 1.3 shows the operation of CMOS inverter. As input goes from 1 to 0, load capacitance C is charged by a current flow from supply through PMOS. The energy loss in charging is $\frac{1}{2}C_{wire}{V_{DD}}^2$. As input goes from 0 to 1, the stored energy lost through NMOS in discharging state is also $\frac{1}{2}C_{wire}{V_{DD}}^2$. Therefore, total dynamic power loss is given by $C_{wire}{V_{DD}}^2$. Since load capacitance must increase or keep the same as scaling because of reduced wiring area, a lower V_{DD} is needed for reducing dynamic power loss. Static power loss in CMOS, on the other hand, as shown in eq. 1.2, depends on I_{leak} and supply voltage where I_{leak} is the leakage current of a reverse bias PN junction given as follows.

$$I_{leak} = I_{off} \times \exp(\frac{-qV_{DD}}{kT}) \quad (1.3)$$

According to Eq. 1.3, leakage current increases as V_{DD} scaling down. Simply scaling V_{DD} does not reduce static power loss. Therefore, there is no explicit way to reduce power density.

Fig. 1.3 shows the operation of CMOS inverter. As Fig. 1.4 (a) shows a supply voltage versus technology node by Microwind [5]. Both inner core supply and external core supply scaled linearly before 90 nm technology node. However, the scaling of core supply voltage went sub-linearly after the 65 nm technology node. Beyond sub-10 nm technology node, inner core supply stuck at nearly 0.6 V. This suggests that the power density issue worsens as with VLSI scaling. High power density not only results in the heating problem

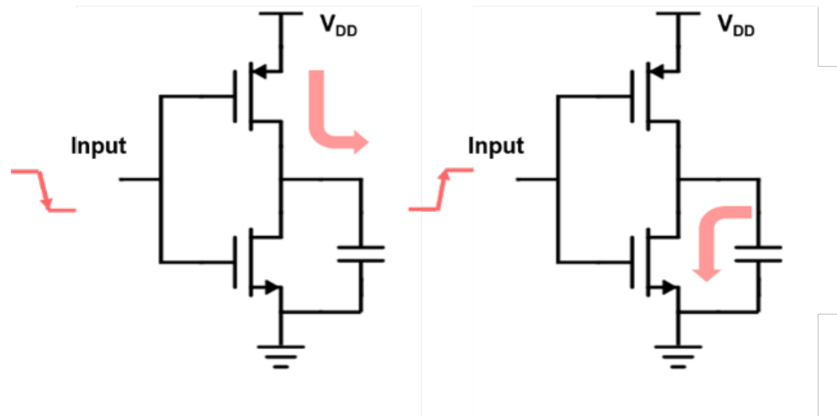


Figure 1.3: Schematic diagram showing the operation of CMOS inverter, from 1 to 0 (left) and 0 to 1 (right).

but also limits processors speed. Fig. 1.4 (b) shows the scaling of microprocessor clock frequency vs. time for the progression of CPUs created by Michael L Rieger [6]. Clock frequency increased exponentially with time before 2003 and stopped from then on due to dynamic power loss.

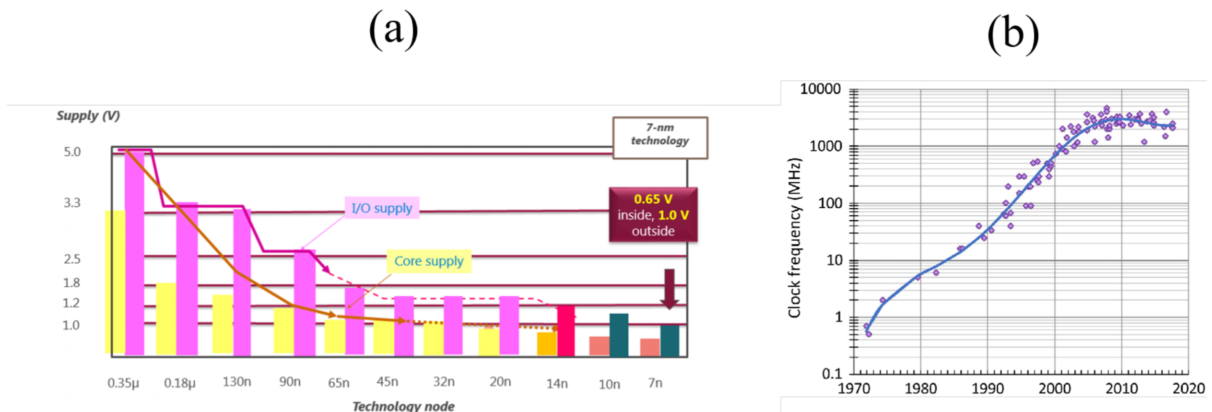


Figure 1.4: (a) Supply voltage vs. technology node by Microwind [5] and (b) Microprocessor clock frequency vs. time created by Michael L Rieger [6].

1.3 Steep Slope Transistors

To reducing power density, V_{DD} scaling needs new physics to eliminate leakage current. A device that can be turned on faster is needed. In other words, if having the same I_{OFF} with MOSFETs, this switch needs to be turned on at a smaller V_{DD}, as shown in fig. 1.5. In MOSFETs, current flow is driven by thermionic emission across over energy barriers. The driving force of this process is thermal energy. The density of carriers at given energy follows Boltzmann distribution, which suggests there is always a fermi-tail at higher energy states that will cause leakage in a subthreshold scheme, as shown in Fig. 1.6 (a).

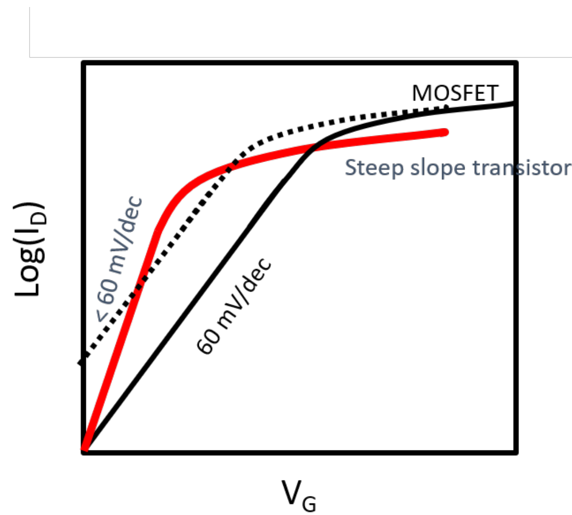


Figure 1.5: Schematic I_D-V_G characteristics of MOSFETs (with minimum SS = 60 mV/dec) and steep slope transistors (with SS < 60mV/dec).

The existence of fermi-tail gives subthreshold leakage current and limits the turn-on speed of logic devices. The amount of subthreshold leakage, which is commonly defined

by and understood as subthreshold swing (SS) for MOSFETs, is given by

$$SS = \frac{dV_{GS}}{dlog(I_D)} = \frac{dV_{GS}}{d\psi_S} \times \frac{d\psi_S}{dlogI_D} = m \times n$$

$$= \frac{C_{ox} + C_{it} + C_d}{C_{ox}} \times ln(10)\frac{kT}{q}$$

$$\cong \frac{C_{ox} + C_{it} + C_d}{C_{ox}} \times 60\frac{mV}{dec}$$

The first-term m is a voltage divider between oxide capacitance and all capacitance. The second-term n represents the change of current corresponding to the change of surface potential, which is nothing but thermal limit. Therefore, even if $C_{ox} \gg C_{total}$, minimal SS that can be achieved in MOSFET is roughly 60 mV/dec at room temperature.

There are two ways of making steep slope transistors. First, make the first term m smaller than 1. M will always be greater than one if $C_{ox} > 0$. On the other hand, $m < 1$ is possible if oxide capacitance is negative. This suggests that new physics that allows the surface potential greater than gate bias is needed. By negative oxide capacitance and matching $C_{ox} > 0$ with $C_d > 0$, sub 60mV/dec is achievable. The MOSFETs which incorporate dielectric that have negative C_{ox} are called negative capacitance field-effect transistors (NCFETs). The ferroelectric material is shown to have the property of negative capacitance in certain operating conditions.

The other way to pursue fast switch logic devices is making second term n sub 60. Subthreshold leakage in MOSFETs or PN diodes comes from an exponential decaying fermi tail with higher energy than the barrier. By adding an energy block, for example, a semiconductor with a bandgap, subthreshold leakage from the fermi tail could be filtered. Fig. 1.6 (b) shows schematic energy band diagram of N-TFETs. By modulating

tunneling barrier thickness through gate bias, devices turn on/off. Electrons tunnel from the source into the channel when the gate turns on. Subthreshold leakage of fermi tail at source region is blocked by valence band edge in this case. Therefore, well-designed TFETs make sub-60mV/dec achievable, and thus TFETs are a promising candidate for low power logic devices.

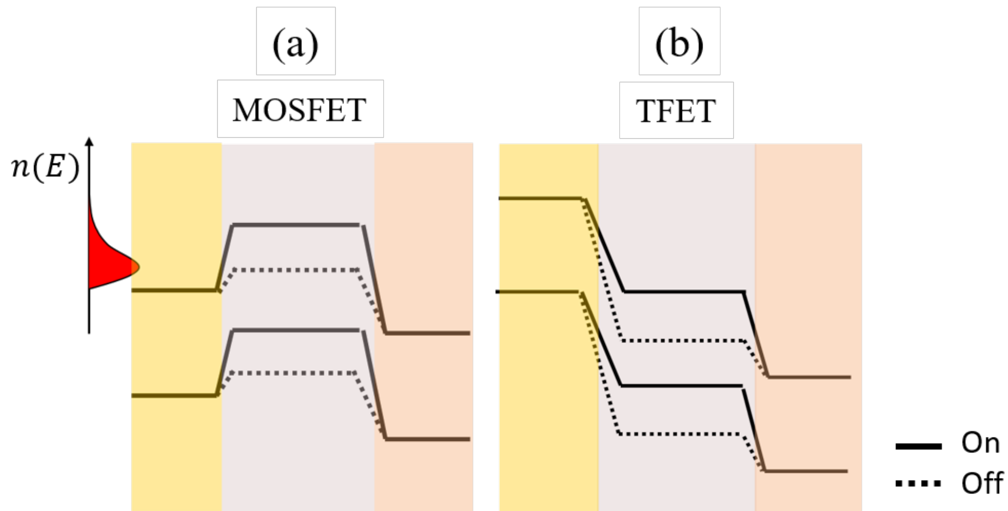


Figure 1.6: Schematic band diagram of N-channel MOSFETs (a) and TFETs (b) at on and off states.

1.4 Tunneling Field Effect Transistors

A fast turn-on, or to say, steep subthreshold swing, is necessary to reduce power loss. Indeed, SS is the most common measure used by all TFETs works. However, the importance of the demand on high drive current is usually ignored.

Even though TFETs are promising as an option for future CMOS, there are many design and processing difficulties in realization. For example, a body thickness around or below 10 nm is needed for high transmission TFETs. The fabrication of such a thin body is challenging. In order to aim at a thin body, process complexity increases dramatically,

leading to a long device turn-around and process developing time. In addition, many physics are still unknown in TFETs, such as traps and auger recombination. Therefore, long processing steps need to be tested in parts, and the potential high-k/channel damage by high-power processing steps should be considered and analyzed beforehand. In this dissertation, a steep subthreshold slope, high drive current 3-HJs TFET is designed. Vertical TFETs fabrication processing flow is developed and being tested on vertical MOSFETs structure.

1.5 Ferroelectric field effect transistors(this section needs to be modified to specify the importance of high on current in tfet

Ferroelectric materials are being claimed to have a negative capacitance in specific operation regimes [7, 8, 9]. $Hf_xZr_{1-x}O_2$ is one of them and being promising for its compatibility with CMOS technology. However, there still has been no direct evidence of capacitance amplification by negative capacitance so far. Nevertheless, the enhancement of C_{ox} is often observed in $Hf_xZr_{1-x}O_2$, which is owing to the enhanced polarizability by spontaneous electric polarization in ferroelectric material [8].

Sayyef Salahuddin's group at UC Berkeley demonstrated Si SOI MOSFET utilizing $Hf_xZr_{1-x}O_2$. An improved short channel behavior is seen comparing to HfO_2 control sample, as shown in fig. 1.7 [10]. Suman Datta's group at the University of Notre Dame presented a large dielectric constant of nearly 40 for a $Hf_xZr_{1-x}O_2$ on silicon at 2019 [11]. Even without the property of negative capacitance, a higher C_{ox} and improved DIBL behaviors are both critical in CMOS scaling. A better gate electrostatic could be expected, potentially making the scaling of physical gate length go further.

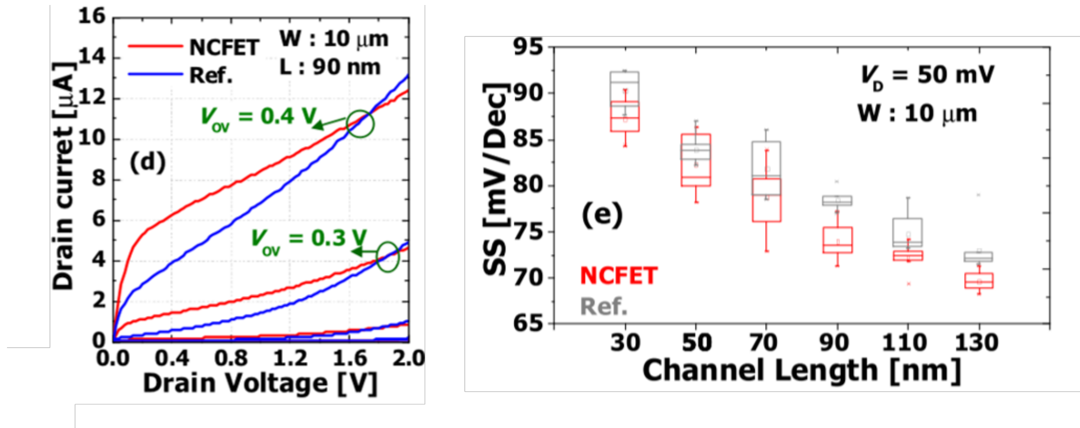


Figure 1.7: Si SOI MOSFET using ferroelectric HfZrO2 in gate dielectric by UC Berkeley research group [10]. An improved short channel behavior is seen comparing to HfO_2 control sample.

1.6 Outline

This dissertation studies the design and fabrication of III-V-based triple heterojunction TFETs. The goal is to demonstrate high on-current, high on/off ratio, and low subthreshold swing TFETs. Chapter 2 discusses TFETs design and theory. The principle of operation in more zen tunneling diodes and TFETs will be explained. The theory of tunneling in PN junction, including the effect of energy barrier, tunneling distance, and effective mass on tunneling probability, will be discussed. The importance of electrostatics in TFETs will be expressly pointed out and examined to illustrate the challenges and potential issues of TFETs.

Chapter 3 introduces doping profile and heterojunction engineering to enhance tunneling. The design of 3-HJ TFET, including material selection, the band offset preference, doping concentration, layer thickness, and other fabrication-related limitations, will be discussed in depth. Strategies for optimizing 3-HJ TFET will be explained using atomistic modeling simulation results (the Purdue team does that). A realistic 3-HJ TFET design will be shown and incorporated with the PNPN doping profile. This design is

finalized and used in the experimental pursuit of 3-HJ TFET.

The fabrication of 3-HJ TFET is divided into two parts: process development of high-quality high-k metal gate (HKMG) on InP channel and development of physical device structure and the fabrication of FETs. Chapter 4 studies the development of high-quality HKMG, including an investigation on different gate metallization techniques and interface traps recovery by post gate H_2 annealing. An ALD-grown conformal TiN/Ru gate is developed and proved on planar MOSFETs and is further incorporated in demonstrating vertical MOSFETs and TFETS.

Chapter 5 discusses the process development toward 3-HJ TFETs. Three generations of device structures that are studied will be briefly explained and summarized. Development of 3rd generation vertical top-down planarization process and the fabrication of vertical FETs will be explained in detail. Results of InP channel vertical MOSFETs and the first demonstration of 3-HJ TFETs will be shown, analyzed, and examined.

Conclusion and future work will be given out in chapter 6. The complete process flow of planar and vertical MOSFETs and all the detailed experimental parameters mentioned in the main context will be attached in the appendix.

Chapter 2

Design of TFETs

In this chapter, the basic TFET operation principle will be reviewed. Zener tunneling in a homogeneous PN junction will be derived to explain the effects of effective mass, energy barrier, and tunneling distance on tunneling probability. To maximize transmission, material design, heterojunction incorporation, as well as device orientation engineering are needed. The importance of having a thin tunneling distance will be emphasized. The tunneling distances in source/channel tunneling junctions will be addressed separately by electrostatics of TFETs. At last, challenges and potential issues of realizing high-performance TFETs will be discussed. Design considerations and fabrication difficulties will be briefly summarized.

2.1 Principles of Operation–Zener Tunneling

TFETs conduct current by electrons tunneling through an energy barrier. In a homogeneous P^+/N^+ structure without a gate, zener tunneling happens when electrons tunnel from valence band to conduction band under reverse bias, as shown in fig. 2.1 (a). A

TFET is a zener diode that has an on/off operation controlled by a gate. By aligning high-k/metal gate to tunneling junction with good gate electrostatics, tunneling barrier thickness could be controlled accordingly to boost/suppress tunneling in on/off-state. Schematic energy band diagrams of homogeneous N-TFETs at off-state and on-state are shown in fig. 2.1 (b) and (c). At off-state, tunneling is blocked by the bandgap of the channel, and electrons are fully depleted by gate electrostatics. At on-state, potential in the channel is pushed down by gate bias while tunneling barrier thickness is reduced. Electrons at the valence band of the source are thus tunneling through P^{+}/i junction into the unfilled states in the channel conduction band.

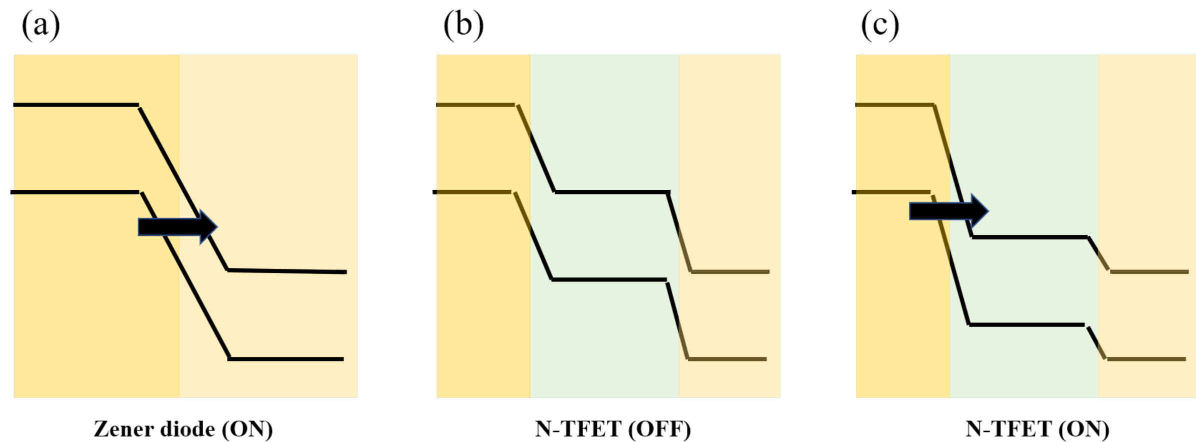


Figure 2.1: Schematic energy band structure of (a) a zener diode at on-state. (b) a N-TFET at off-state, and (c) a N-TFET at on-state.

The simplest way to understand a device is by its current-voltage relation. In this section, analytic expression of a zener tunneling current will be derived for a reverse biased P^{+}/N^{+} tunneling junction with no gate, as shown in fig. 2.2 (a). Parabolic and isotropic conduction and valence bands are assumed to simplify the calculation.

The zener tunneling current is given by integrating the product of charge flux and tunneling probability through the energy within the tunneling window, as given below

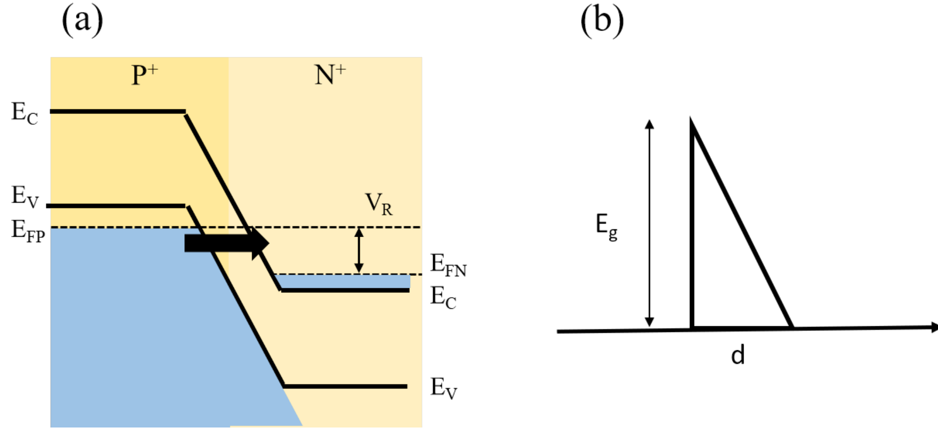


Figure 2.2: Schematic band structure of a P$^+$/N$^+$ zener diode under reverse bias (a) and the equivalent triangular energy barrier for electrons tunneling from the valence band into conduction band at a certain energy (b).

[12].

$$[h]J = \int\int qv_g(k_z)n_z(k_z)n_\perp(k_\perp)dk_x 2\pi k\perp dk_\perp \cdot (f_V - f_C) \cdot T_{WKB} \tag{2.1}$$

where $v_g(k_z) = (1/\hbar) \cdot (\mathrm{d}E_z/\mathrm{d}z)$ is the group velocity, $n_z(k_z) = 1/\pi$ and $n_\perp(k_\perp) = 1/4\pi^2$ are density of states in tunneling direction and transverse direction, respectively. k_z and $k_\perp$ are wave vector in tunneling direction and transverse direction. T_{WKB} is tunneling probability, which will be calculated using Wentzel-Kramers-Brillouin (WKB) approximation $T_{WKB} \simeq exp^{-2\int_0^d |k_z(z)|dz}$ [13]. f_v and f_c are Fermi-Dirac distribution function at valence band of P$^+$ source and conduction band of N$^+$ side.

Electrons at valence band of P$^+$ source tunnel through the forbidden gap into the conduction band of N$^+$ side. This process can be analyzed as electrons tunneling through a triangular quantum well, from which the height is bandgap energy E_g and slope is elementary charge times electric field $q\varepsilon$, as shown in fig. 2.2 (b). Thus, tunneling distance $d \simeq E_g/q\varepsilon$, is a function of bandgap energy and ε, where ε depends on doping and applied bias.

To calculate T_{WKB}, the first wave vector k_z needs to be derived. Considering a bulk direct bandgap PN junction, total energy E can be divided into transverse energy $E_\perp$ and tunneling energy E_z.

$$E = E_\perp + E_z \tag{2.2}$$

where $E_z = \hbar^2 k_z^2/2m_r^*$ is energy in tunneling direction and $E_\perp = \hbar^2(k_x + k_y)^2/2m_r^*$ is energy in transverse direction. m_r^* is reduced effective mass where $1/m_r^* = 1/m_e^* + 1/m_h^*$. Wave vector in transport direction is thus given by

$$k_z(z) = \sqrt{\frac{2m_r^*}{\hbar^2}(E_z - V(z))} \tag{2.3}$$

$V(z)$ is potential energy quantum well. As shown in fig. 2.2 (b), this problem can be simplified and analyzed as a triangle quantum well such that $V(z) = E + q\varepsilon z$ were $0 < z < d$. Thus, wave vector can be written as

$$k_z(z) = \sqrt{\frac{2m_r^*}{\hbar^2}(-q\varepsilon z - E_\perp)} \tag{2.4}$$

By calculating using WKB approximation, tunneling probability is given as [14]

$$T_{WKB} \sim exp(-\frac{4\sqrt{2m_r^*}E_g^{3/2}}{3q\hbar\varepsilon})exp(-\frac{E_\perp}{q\hbar\varepsilon/2\sqrt{2m_r^*E_g}}) \tag{2.5}$$

In thick body conditions, the denominator of the 2nd exponential term determines how much the degradation from electrons scattering in the transverse direction. Therefore, to maximize tunneling probability, a high junction field, small effective mass, and small bandgap material are needed. On the other hand, $E_\perp << E_g$ in thin body one-

dimensional condition, which results in higher tunneling probability

$$T_{WKB} \sim exp(-\frac{4\sqrt{2m_r^*}E_g^{3/2}}{3q\hbar\varepsilon}) \tag{2.6}$$

Since current flows by electrons tunneling from source into unfilled band of channel, tunneling only takes place at energy level between quasi fermi-level of source E_{FP} and channel E_{FN}. By setting E_{FN} as the starting point of integration for eq. 2.1, zener current is integrated over energy from 0 to qV_R. V_R is the reverse bias of P^+/N^+ junction. Fermi-Dirac function in channel side thus becomes $f_c(E) = [1+exp^{E/kT}]^{-1}$. At P^+ side, f_v can be written as $f_v(E) = [1+exp^{(E-qV_R)/kT}]^{-1}$. To simplify the equation, $f_v(E) - f_c(E) = 1$ is assumed here. By integrating eq. 2.1 with eq. 2.6, the resulted zener tunneling current of a homogeneous P^+/N^+ junction becomes [14]

$$J_{zener} = \frac{\sqrt{2m_r^*}q^3\varepsilon V_R}{8\pi^2\hbar^2E_g^{1/2}}exp(-\frac{4\sqrt{2m_r^*}E_g^{3/2}}{3q\varepsilon\hbar}) \tag{2.7}$$

This equation can also be written in tunneling distance (d) instead of electric field

$$J_{zener} = \frac{\sqrt{2m_r^*}q^2V_RE_g^{1/2}}{8\pi^2\hbar^2d}exp(-\frac{4\sqrt{2m_r^*}E_g^{1/2}}{3\hbar}d) \tag{2.8}$$

A small bandgap material and high junction field/ short tunneling distance are necessary for maximizing tunneling current. A higher effective mass gives higher value for the 1st term in eq. 2.8. However, current decreases exponentially with effective mass in 2nd term. Therefore, a small effective mass benefits tunneling current overall. Fig. 2.3 shows the zener tunneling current density per 1 V of reverse bias ve. electric field for homogeneous Si and InGaAs PN junctions calculated by eq. 2.7. The bandgap and reduced effective mass used in the calculation follows [14, 15, 16] ($E_g = 1.1$ eV, $m_r^* = 0.16$ for Si and $E_g = 0.75$ eV, $m_r^* = 0.023$ for InGaAs). InGaAs tunneling junction having smaller

bandgap and effective mass comparing to Si, shows much higher tunneling current.

Eq. 2.8 is also very useful in the calculation of ambipolar leakage current in TFETs. Ambipolar leakage problem in small bandgap material TFETs will be introduced in Chapter 3. We will analyze the amount of ambipolar leakage current by this equation for thick body TFETs using an InP channel.

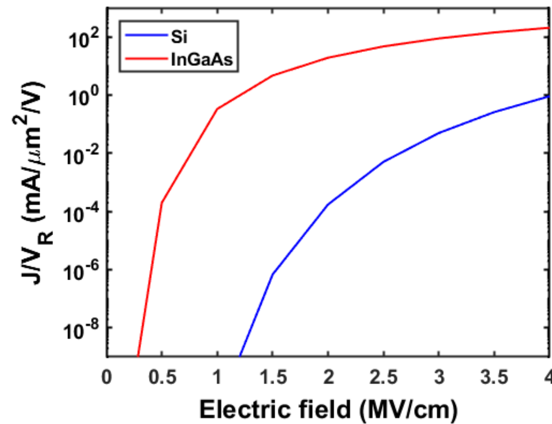


Figure 2.3: The calculated zener tunneling current density per 1 V of reverse bias using eq. 2.7 for homogeneous Si and InGaAs PN junctions.

2.2 Tunneling Probability Maximization

From the last section, it is shown that effective mass, energy barrier, and tunneling distance determine tunneling probability and on-current. In this section, we will introduce how to engineer each of them to optimize TFET performances. The amount of change on tunneling probability from engineering each of them will be calculated and compared.

	InAs	Si	GaSb	InGaAs	InP	GaAs	GaN
Bandgap (eV)	0.35	1.1	0.67	0.74	1.34	1.44	3.4
Electron effective mass (m_0)	0.023	1.09	0.047	0.041	0.07	0.063	0.2
Hole effective mass (m_0)	0.41	0.81	0.4	0.36	0.4	0.51	0.8

Table 2.1: Bandgap energy and effective mass for semiconductor materials. InGaAs listed has composition of 53% of In and 47% of Ga. [17][18][19][20][21][22]

2.2.1 Effective Mass

By picking a material with a small effective mass, the tunneling probability would be improved. Typically, a small bandgap material has a lower effective mass, i.e. InAs has a bandgap energy of 0.35 eV and electron effective mass of 0.023 m_0. Table 2.1 summarizes basic material properties of InAs, Si, GaSb, InGaAs, InP, GaAs, and GaN. Note that the material parameters listed here are for the bulk condition; both bandgap and effective mass would increase under a high confinement situation. Considering band alignment and material properties, InAs have excellent conduction band properties and GaSb has good valence band properties for tunneling.

Aside from picking a material with preferred effective mass, effective mass in a determined material could also be engineered by geometry design. Long *et al.* reported a preferable TFET orientation such that the effective mass is engineered to boost the tunneling [23]. As shown in eq. 2.2, total energy can be divided into two parts: transverse energy and tunneling energy. By device orientation design, higher tunneling energy and lower transverse energy, meaning a lower transport effective mass and a higher confinement effective mass, would enhance tunneling probability.

According to [24], a 1.8 nm InAs has highest transport electron effective mass of $0.1m_0$ at [100] transport direction with (001) confinement. The lowest transport electron effective mass of $0.05m_0$ happens at [110] transport direction with $(1\bar{1}0)$ confinement, while a 1.8 nm GaSb has the highest and lowest transport hole effective mass of $0.174m_0$

and $0.065m_0$ at the same device orientation as InAs [24]. This suggests that a more than 2 times decrease in effective mass could be achieved by aligning devices at [110] transport direction instead of commonly seen [100] transport. Fig. 2.4 shows the schematic diagram of a thin body double-gated TFETs in different device orientations (a) and their energy k-space contours (b).

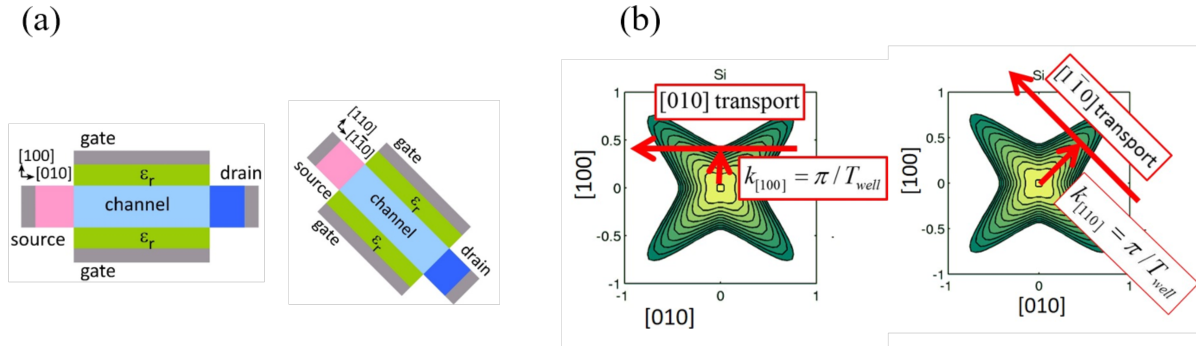


Figure 2.4: Schematic diagram of thin body double-gated TFETs in different device orientations (a) and their energy k-space contours (b). Courtesy of Purdue team.

Fig. 2.5 (a)(b) shows the band structures for conduction band of InAs and valence band of GaSb in (001) and $(1\bar{1}0)$ confinement, respectively [23]. It is clear that both energy barrier and effective mass are reduced in $(1\bar{1}0)$ confinement condition. Fig. 2.5 (c) is energy band diagram and (d) is the transmission versus energy for a InAs/GaSb TFET. Peak tunneling probability roughly doubles for $(1\bar{1}0)$ confinement with [110] transport comparing to (001) confinement with [100] transport.

2.2.2 Energy Barrier

Small barrier energy for tunneling is required to maximize tunneling probability. Tunneling probability is roughly exponentially decaying with $E_{barrier}^{3/2}$ as shown in eq. 2.6. This suggests by cutting barrier energy half, a roughly 2 times enhancement in tunneling probability can be achieved.

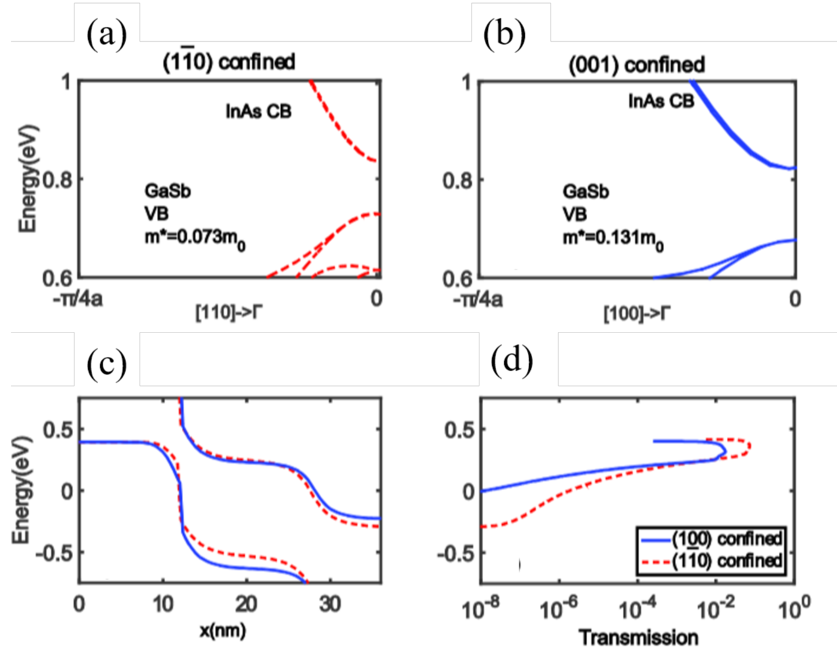


Figure 2.5: Band structures for conduction band of InAs and valence band of GaSb in (001) (a) and (1$\bar{1}$0) (b) confinement, and energy band diagram (c) and transmission vs. energy (d) for a InAs/GaSb TFET at the design orientation. [23]

InAs has the smallest direct bandgap of 0.32 eV which makes it promising in all TFET designs. Besides choosing a small bandgap material for tunneling junction, adding heterojunctions also adjusts the tunneling energy barrier. A broken gap band alignment makes a naturally thin tunneling distance and a small energy barrier.

Common TFET design uses Sb-based source and InAs channel due to their near broken gap alignment. Fig. 2.6 shows the schematic energy band diagram of a GaSb/InAs tunneling junction. The red regime is P^+-GaSb and the grey regime is the UID-InAs channel. Given this heterostructure tunneling junction, the whole tunneling process includes two parts. Electrons in the source first borrow the energy from the valence band of GaSb. After that, electrons borrow the energy from the conduction band of InAs and reach the channel conduction band. Therefore, the tunneling for heterojunction TFETs depends on both the properties of the channel conduction band and the source valence

band. In this case, T$_{barrier}$ is the tunneling distance, and ΔE is the energy barrier, which is much smaller than homojunction TFETs.

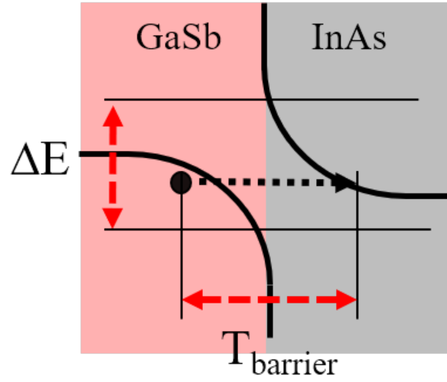


Figure 2.6: Schematic energy band structure of a GaSb/InAs tunneling junction.

2.2.3 Tunneling Distance

Tunneling probability decreases exponentially with tunneling distance. Tunneling distance $T_{barrier} = E_g/q\varepsilon$ is a function of doping concentration, build-in field, and energy barrier in zener diodes. For homojunction TFETs such as Si P-i-N TFETs, a high source doping is needed for a thin source tunneling thickness. However, Si has a relatively large bandgap of 1.1 eV which makes the tunneling distance still roughly $>$ 5nm with a higher source doping above 5×10^{19} cm^{-3}. This suggests only managing source doping concentration is not enough to get good tunneling when the energy barrier is large. Therefore, other ways of shrinking tunneling distance are needed.

Increasing build-in field by heterojunctions is a way to thin down tunneling distance. By adding heterojunctions with proper band alignment, the build-in field can be dramatically increased, and thus tunneling distance is decreased [25]. Fig. 2.7 (a) shows the energy band diagrams of InGaAs homojunction TFET and triple-heterojunction TFET with t_{body} of 3 nm. Tunneling distance is reduced from ≈5 nm (homojunction TFET)

to 2.5 nm (heterojunction TFET) and the peak tunneling probability increases by more than 10 times as shown in (b). It is noted that the large bandgap source well and channel well introduce bound states at two sides of the tunneling junction. The bound state's energy affects tunneling properties thus needs to be designed properly. Detailed design considerations of 3-HJs TFETs will be discussed in the next chapter.

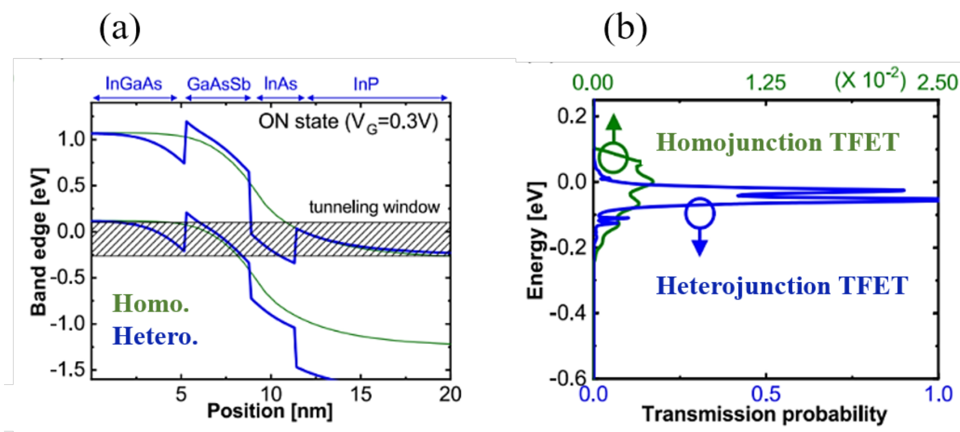


Figure 2.7: Energy band diagrams of InGaAs homojunction TFET and triple-hetero-junction TFET at t_{body} of 3 nm (a) and their tunneling probability vs. energy at V_{GS} = 0.3 V. Courtesy of Chin-Yi Chen.

Another way of thinning down tunneling distance is by modifying the doping profile. Increasing the doping concentration at the source/channel tunneling junction increases the junction field. In TFETs, a P-i-N doping profile is usually used. An unintentionally/low doped drift region in the channel is needed to hold the potential difference between the gate and drain, thus prevent the short channel effect. To reduce tunneling distance in channel, people proposed utilizing P-N-P-N doping profile [26][27][28]. Such doping profile not only improves junction electric field but also better electrostatic of the gate for a thick body. Details will also be explained in Chapter 3.

Table 2.2 compares the amount of change in tunneling probability resulted from the changes of effective mass, energy barrier, and tunneling distance. The tunneling probability is calculated using eq. 2.6. In table 2.2 (a), energy barrier of 0.2 eV and

(a)

E_b = 0.2 eV, $T_{barrier}$ = 1 nm	
m*	P
0.03 m_0	45%
0.06 m_0	33%
0.12 m_0	20%

(b)

m* = 0.06 m_0, $T_{barrier}$ = 1 nm	
$E_{barrier}$	P
0.1 eV	24%
0.2 eV	13%
0.4 eV	6%

(c)

m* = 0.06 m_0, $E_{barrier}$ = 0.2 eV	
$T_{barrier}$	P
1 nm	33%
2 nm	11%
4 nm	1%

Table 2.2: Tunneling probability vs. (a) effective mass, (b) energy barrier, and (c) tunneling distance.

$T_{barrier}$ of 1 nm are set, while effective mass changed from 0.03 m_0 to 0.12 m_0. In (b), effective mass of 0.06 m_0 and $T_{barrier}$ of 1 nm are set, while energy barrier changed from 0.1 eV to 0.4 eV. In table (c), effective mass of 0.06 m_0 and energy barrier of 0.2 eV are set, while tunneling distance changed from 1 nm to 4 nm. By doubling effective mass, tunneling probability decrease by roughly 1.5 times. Similarly, doubling energy barrier and $T_{barrier}$ gives about 2 and 3 times reduction in tunneling probability, respectively. In conclusion, tunneling probability depends on effective mass, energy barrier, and $T_{barrier}$. $T_{barrier}$ is the most critical and thus must need to be thin for a high transmission high on-current TFETs.

2.3 Electrostatics of TFETs

In the previous section, we discussed the tunneling in the P/N junction and the importance of tunneling distance on TFETs performances. However, the on/off operation of a TFET for a low power logic switch is controlled by the gate. By putting down dielectric and metal gates, while aligning them with the unintentionally doped channel region, channel potential is controlled by gate electrostatic. In this section, the electrostatics of TFETs will be discussed and what determines tunneling distance in TFETs will be analyzed.

When a positive bias is applied to the gate for N-TFETs, channel potential is driven downward. The electric field at the tunneling junction thus increases and electrons tunnel from the source valence band into the channel conduction band. At off-state, channel potential is pulled upward. Junction electric field drops and tunneling distance increases such that tunneling is blocked. In short, tunneling distance determines TFETs on/off state.

$T_{barrier}$ can be divided into two parts: tunneling barrier thickness in the source and in the channel. Take the TFET structure in fig. 2.8 (a) as an example. Black layers represent metal contact. t_{ox} and t_{body} is oxide thickness and body thickness, respectively. Perfect gate alignment at the P/i junction with gate electric field perpendicular to the junction is assumed.

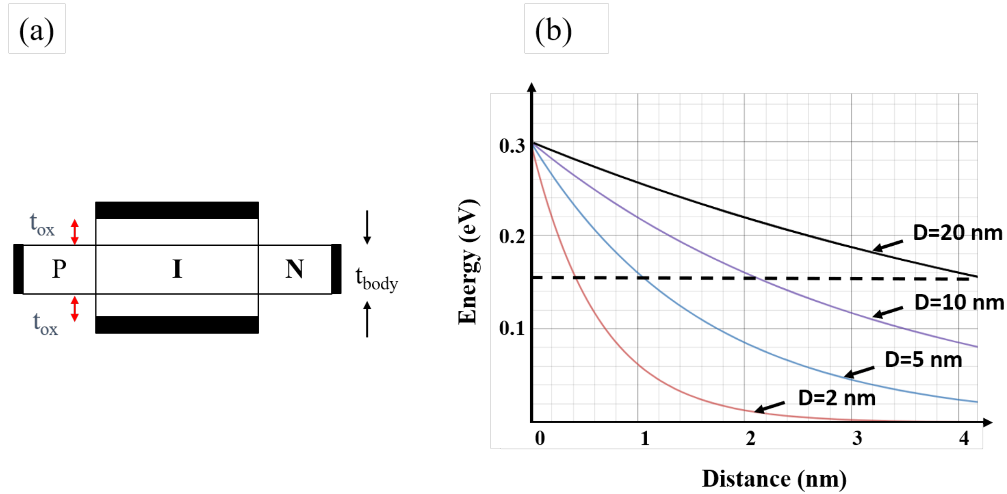


Figure 2.8: Schematic double-gated TFET structure (a) and the calculated channel potential vs. distance with different cavity thickness D (b).

At source tunneling junction, tunneling barrier thickness given by Poisson's equation

$$\nabla^2\varphi = -\frac{qN_A}{\epsilon} \tag{2.9}$$

depends on source doping concentration N_A. If N_A goes higher, depletion width of at

P-side decreases which is the same as P/N junction analysis. Therefore, a high source doping is needed for a thin tunneling barrier thickness in the source. However, the number of electrons that are available for tunneling drops when source doping goes high. For a higher doped source, fermi-level at P-side goes down which means the average energy of electrons goes down as well. If source fermi-level goes below the junction tunneling energy window, there will be limited electrons that have enough energy for tunneling through.

In addition, on-current will be sensitive to temperature if source doping is too high. As mentioned before, Fermi-level goes down far below the valence band for a highly doped source. The distribution of electrons in the source, given by Boltzmann's distribution, shows roughly exponentially decaying numbers of electrons above source fermi-level. The distribution of this fermi-tail is temperature-related. A higher temperature gives a lower exponentially decay constant, which suggests more electrons are at a higher energy state. Therefore, numbers of electrons near to tunneling energy window fluctuate with environment/junction temperature, and so as on-current.

Ideally, it is better to keep the source fermi-level closed to the top edge of the tunneling energy window to maximize the number of electrons available for tunneling and minimize the temperature effect. These two factors compensate with thinner $T_{barrier}$ given by high doping source.

On the other hand, tunneling barrier thickness at the channel side depends on gate electrostatics. To simplify the calculation, we assume there are infinite amounts of carriers in the source and no carrier in the channel. Material within two gate plates is also assumed to be homogenous. By solving Poisson's equation with boundary condition given by uniform gate bias across gate plates, the potential in the channel can be written

as

$$\psi = \psi_0 + \sum_{n=1}^{\infty} A_n \times cos\frac{x}{D/n\pi} \times exp^{\frac{-y}{D/n\pi}} \tag{2.10}$$

where ψ_0 is gate bias and $D = t_{body} + 2t_{ox}$ is the thickness of the whole cavity. From the exponential term in eq. 2.10, it's clear that the potential near to P/i junction drops exponentially with decay constant $n\pi/D$. It is noted that the more rapid potential drops, the thinner the tunneling distance is.

Fig. 2.8 shows the potential profile calculated by eq. 2.10 at $y = 0$. $V_{GS} = 0.3$ V is set as a boundary condition. For a $D = 5$ nm cavity, the tunneling distance is around 1 nm at a 50% drop of peak potential. Roughly speaking, tunneling distance at 50% of peak potential doubles by doubling D, which indicates tunneling probability, as well as on-current, are sensitive to cavity width. Therefore, a thin cavity, which includes thin oxide and body, is necessary for high transmission.

Minimum tolerable gate leakage current sets minimum gate oxide thickness. As a result, the limitation of gate oxide thickness determines the minimum tunneling barrier thickness at the channel side. In addition, body thickness also plays an important role in TFETs' performance. An ultra-thin t_{body} is needed for good electrostatics and thin tunneling distance.

Two problems raise if t_{body} gets thick. First, energy potential couldn't be bent down fast enough from the junction which makes tunneling inefficient. Second, same as the short channel effect in MOSFETs, the potential is not well controlled by the gate in the channel region far away from the high-k/semiconductor interface. Due to bad gate electrostatics, the center of the channel floats, and thus goes up and down with drain bias. This drain-induced barrier lowering (DIBL) degrades the transfer characteristics of TFETs because channel potential becomes a function of V_{DS} rather than fully controlled

by V_{GS}. Therefore, TFETs couldn't be turned on/off effectively in the floating body region.

In conclusion, we showed an effective mass, energy barrier and tunneling distance are three factors that determine tunneling properties. Effective mass could be optimized by designing a device orientation that has a low transport effective mass and high confinement effective mass. Energy barrier and tunneling distance could be engineered by adding heterojunctions. Tunneling distance in TFETs is a function of build-in field and electrostatics, thus it is difficult to be thinned down. This results in a low tunneling probability and low on-current comparing to MOSFETs.

2.4 Challenges and Potential Issues of TFETs

To achieve high-performance TFETs, it is essential to optimize the source/channel junction tunneling and consider potential design problems and fabrication constraints. For a higher tunneling probability, a small bandgap material is preferred. However, we will show in section 2.4.1 that ambipolar leakage current is a concern that is caused by employing a small bandgap material in the channel. In addition, gate misalignment, which is commonly seen in device fabrication, will be shown to degrade TFETs performances in section 2.4.2 dramatically. Trap-assisted tunneling (TAT) is another commonly seen concern, especially in heterojunction TFETs. Its impact on the current transport at on and off state in TFET will be illustrated in 2.4.3.

Last but not least, an ultra-thin body is needed to keep good electrostatics, as has been proved in section 2.3. Indeed, body thickness is set below 5 nm in most of the literature of TFETs analytical design. In terms of device fabrication, it is very challenging to realize a sub 5 nm body, especially in an academic cleanroom, due to processing difficulties/limitations. Prove concepts is hard and takes a large number of works. In

chapter 3, doping profile engineering will be shown as an alternative solution to solve this problem.

2.4.1 Ambipolar Leakage

From the design point of view for low power logic devices, the goal is getting high I_{on}, low SS and low off-state leakage current TFETs by material choose, doping profile design, geometry engineering. Significant energy barrier and thick tunneling distance limit TFETs on-current; thus, small bandgap material is favored in tunneling junction design. However, a channel with small bandgap material such as InAs causes ambipolar leakage current.

Ambipolar leakage, which is also called direct source to drain tunneling, happened at a small bandgap gate/drain junction at a reverse bias V_{GS}. A common double-side gated P-GaSb/InAs TFET structure, as in fig. 2.9 (a), is used to illustrate. The energy band diagram of this TFET at V_{GS} -0.7 V and -1 V with V_{DS} 0.3V is shown in fig 2.9 (b). To be noted that simulation is done at T_{body} = 4 nm and L_g = 15 nm, and assuming no strained effect in the structure. The confinement of this 4 nm thin body makes the energy bandgap for both GaSb and InAs much higher than bulk.

GaSb/InAs tunneling junction roughly lies at x = 12 nm (as can be seen in (b)). When V_{GS} goes more negatively, potential in the channel is pulled up, which increases the electric field between channel/drain junction. The increase of V_{DG} introduces a tunneling window between channel/drain. The transmission versus energy plot confirms the tunneling between channel/source, as shown in fig. 2.9 (c). While V_{GS} goes from -0.7 V to -1 V, the transmission increases dramatically at energy within -0.2 eV to -0.4 eV. The energy window matches the band diagram where a thinner tunneling distance between channel/drain junction could be seen. The exponentially increase in transmission leads

to a huge leakage current at low V_{GS}. Fig. 2.9 (d) shows $I_D - V_G$ characteristic. leakage current increase by more than two orders of magnitude within -0.3 V difference in V_{GS}.

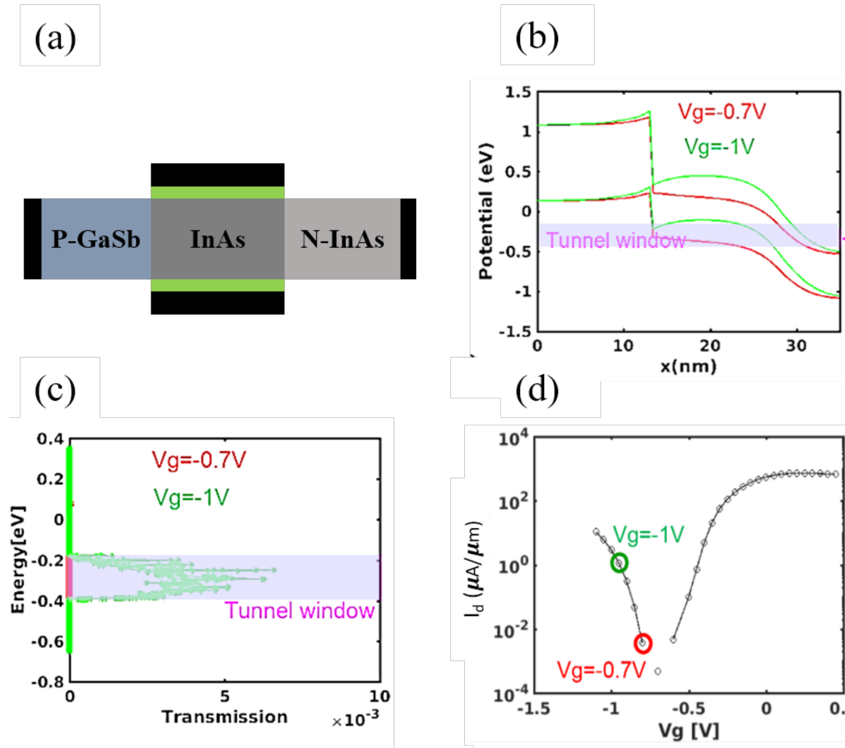


Figure 2.9: (a) Double-side gated P-GaSb/InAs TFET structure and its (b) band diagram, (c) tunneling probability vs. energy at V_g of -0.7 V and -1 V, and (d) the I_D - V_G characteristic. Courtesy of Chin-Yi Chen.

Numerous solutions have been proposed in the literature to overcome the ambipolar leakage problem. To prevent this direct source to drain tunneling, it is vital to keep the electric field between the channel/drain junction low. In other words, the slope in the energy band diagram at the channel/drain junction should be kept less steep such that tunneling leakage can be reduced.

Fig. 2.10 summarizes five different ways of improving ambipolar leakage. Normal PiN TFETs are shown in (a). By utilizing higher dielectric constant oxide near to tunnel junction and lower dielectric constant one at drain side, as shown in (b), the electric field

at channel/drain junction goes down due to a less than ideal gate electrostatics closed to drain [29]. A similar concept applies to a TFET with a uniform dielectric constant but nonuniform thickness. Thinner near to tunnel junction and thicker closed to drain works similarly to (b). In addition, people also proposed using multi-gate or local threshold voltage tuning with segmented gate metals of different work functions.

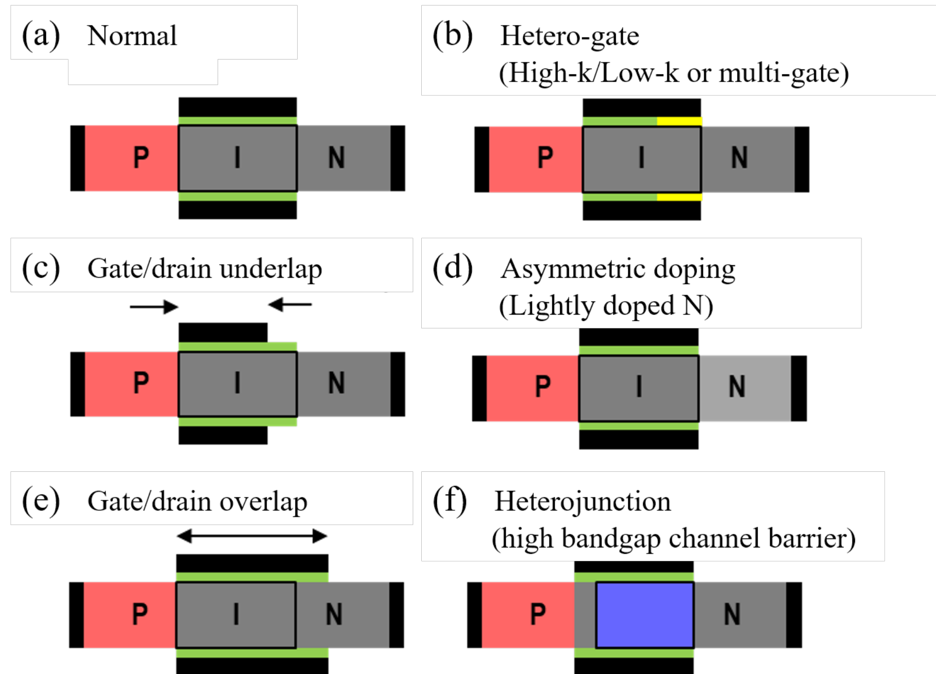


Figure 2.10: Summary of 5 ways to improve ambipolar leakage. (a) A normal P-i-N TFET. (b) Hetero-gate TFET. (c) Gate/drain underlap TFET. (d) Asymmetric doping TFET. (e) Gate/drain overlap TFET. (f) Heterojunction TFET.

Another way of overcoming the ambipolar leakage problem is the gate/drain underlapping, as shown in fig. 2.10 (c) [30]. The potential drop at the channel/drain junction is supported across a larger drift region by intentionally partial-gating. The curvature of the band diagram at this junction is thus softened, and tunneling leakage could be prevented.

Asymmetric doping of source and drain is the third way to solve this problem (fig. 2.10 (d)) [31]. As mentioned in section 2.3, tunneling barrier thickness in the source

depends on the source doping level. P^+ source usually needs to be doped above 5×10^{19} cm^{-3} for a higher transmission. On the other hand, drain does not limit doping in terms of transmission apart from drain resistance. Therefore, the drain can be lower doped such that the electric field between channel/drain is reduced.

Fig. 2.10 (e) shows the alternative geometry engineering: gate/drain overlapping [32]. By extending the high-k/metal gate onto part of the drain, the effective gate length extends. Fig. 2.11 compares the energy band diagram of a normal TFET (a) and gate/drain overlapping TFET (b) [28]. L and L_{ov} are physical gate length and extended gate length, respectively. In the overlapping case, potential drop distributes across the extended region, which reduces peak electric field and thus tunneling leakage.

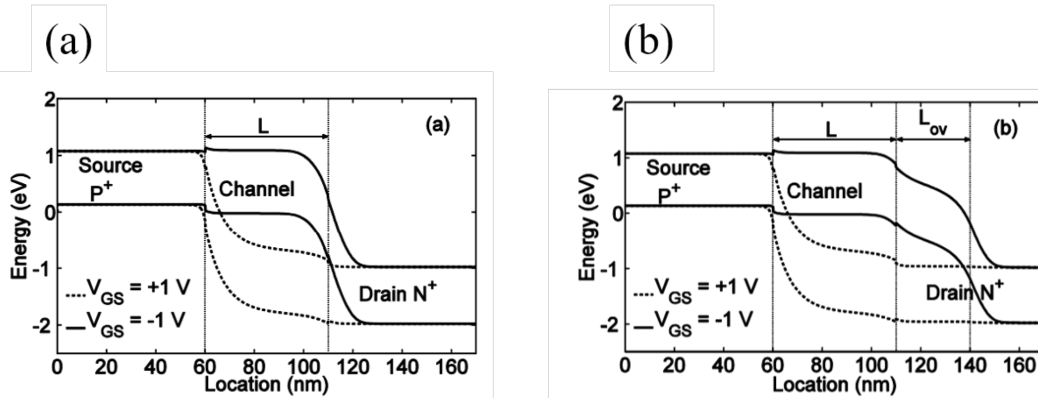


Figure 2.11: Energy band diagram of (a) a normal TFET and (b) a gate/drain overlap TFET. L_{ov} is the overlap distance, which is 30 nm in this case. [28]

The effect of the architectures mentioned above is compared in fig. 2.12 [31]. This simulation work on the ambipolar leakage reduction on Si TFETs was done by R. Narang et al. in 2012. Overall, all these structures have their pros and cons. Gate/drain overlapping and underlapping requires precise control of gate metallization processing, which is very challenging in III-V-based TFETs. Incorporating hetero-gates makes fabrication more complex and difficult. A large gate underlapping and low doped drain both show good improvements on ambipolar behavior, but gate-to-drain capacitance (C_{GD})

increases. The increase of C_{GD} hurts device operation speed.

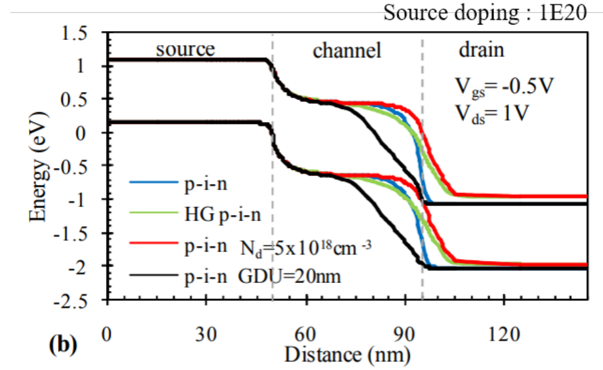


Figure 2.12: Energy band diagram comparing different solutions to ambipolar leakage problem. HG is hetero-gate and GDU is gate/drain underlap. [31]

In this work, a different way of solving ambipolar leakage is used: utilizing high bandgap material. A higher energy barrier suppresses tunneling dramatically. Meanwhile, a thin barrier must be kept to maintain the high transmission in tunnel junction at on-state for a high on-current. Therefore, a small bandgap material for channel tunnel junction is kept while a large bandgap channel barrier next to is employed. As shown in fig. 2.12 (f), this method decouples the conflicts between thin barrier for high I_{on}, and thick barrier for low ambipolar leakage current. In short, by designing the heterojunctions with proper band alignment, it is possible to achieve high $I_o n$ and low leakage simultaneously. Transmission at tunnel junctions can be improved while ambipolar leakage reduces. In our finalized 3-HJs TFETs design which will be discussed in chapter 3, a large bandgap InP channel barrier right next to InAs channel well is added. With the large bandgap InP (1.4 eV) channel barrier and drain, ambipolar tunneling leakage is sufficiently suppressed for the operation with 0.3 V supply voltage.

2.4.2 Gate misalignment

In the previous chapter, the current-voltage relation of a zener diode has been derived. Tunneling current from eq 2.8 could be roughly simplified and rewritten as

$$I = \alpha\, V_R\, \varepsilon\, exp^{-\frac{\beta}{\varepsilon}} \tag{2.11}$$

where $\alpha = A\, q^3\, \sqrt{2m_R^*/E_g}\, /\, 8\pi^2\hbar^2$ and $\beta = 4\sqrt{2m_R^*}\, E_g^{3/2}/3q\hbar$ are coefficients determined by material properties. A is the cross-sectional area of devices. Expression of subthreshold swing, which is $d\, log(I_D)/d\, V_{GS} \approx d\, log(I)/d\, V_R$ in this case, is given by [14]

$$SS = ln(10)\, [\, \frac{1}{V_R} \frac{dV_R}{dV_{GS}} + \frac{\varepsilon + \beta}{\varepsilon^2} \frac{d\varepsilon}{dV_{GS}}\,]^{-1} \tag{2.12}$$

To minimize SS, both the first and second terms of eq. 2.12 needs to be maximized. The maximization of dV_R/dV_{GS} suggests gate to source voltage should control tunnel junction bias as much as possible. It means TFETs geometry needs to be well designed, and gate dielectric needs to be thin such that good gate electrostatics can be obtained. In addition, maximizing the second term $d\varepsilon/dV_{GS}$ means tunnel junction electric field should directly controlled by V_{GS}. This is attained when a high-k/metal gate is placed right at the tunnel junction. By well-aligning the gate with good gate electrostatics, SS can be optimized in TFETs fabrication.

Fig. 2.13 shows gate misalignment analysis for GaSb/InGaAsSb/InAs nanowire TFETs done by the research team in Lund University [33]. Given the planarization process in nanowire TFETs, some degree of gate overlap on the source is inevitable and hard to solve because there is no good way to stop ashing right at tunnel junctions. Fig. 2.13 (a) shows the nanowire TFET cross-sectional schematic structure. Blue regime

on nanowire indicates U.I.D InAs channel, and dash lines indicate gate/source overlap distance. Fig. 2.13 (b) shows that the experimental $I_D - V_G$ curve matches well with the simulated gate-overlap condition and roughly drops by four times comparing to gate aligned TFETs. Fig 2.13 (c) shows the simulated transfer characteristics of nanowire TFETs with 5 nm underlap/overlap and well-aligned gates. It is clear that both gates underlap and overlap hurt TFETs SS and drive current, while gate overlap degrades performances more for the double-heterojunction TFETs.

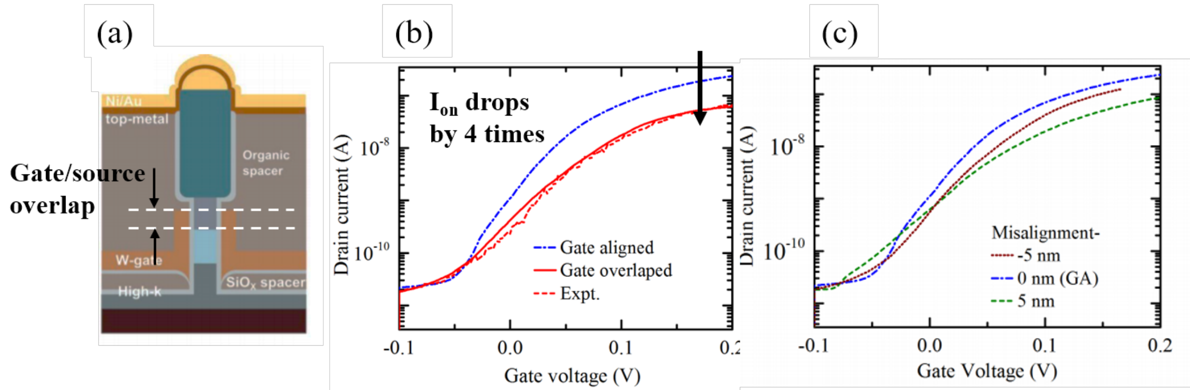


Figure 2.13: Gate misalignment analysis for GaSb/InGaAsSb/InAs nanowire TFETs done by the research team in Lund University [33]. (a) The structure of nanowire TFETs, and (b) I_D - V_G simulation fitting to the experimental gate/source overlap condition. (c) The simulated transfer characteristics of TFETs having aligned gate and 5 nm gate/source overlap and underlap gates. ©2018, ISTE® (International Society for Technology in Education), iste.org. All rights reserved.

In conclusion, the general guideline of TFET design and its potential problems are discussed. Ideally, a high source doping, an ultra-thin body, and properly aligned heterojunctions are needed to achieve a high on-current. In the next chapter, triple heterojunction TFET will be introduced. A comprehensive study will be examined on the optimization of bound states and an alternative way to a thin body.

2.4.3 Traps-assisted Tuennling

Like MOSFETs, traps at the high-k/semiconductor interface also degrade SS in TFET due to the loss of gate control over the channel. As discussed in chapter 1, ideal MOSFETs have SS of 60 mv/dec, which is set/limited by Boltzmann distribution of fermi-tail. In MOSFETs, the magnification of degradation in SS concerning interface trap density (D_{it}) can be estimated by the ratio between total capacitance to oxide capacitance, where total capacitance includes oxide capacitance (C_{ox}), semiconductor capacitance (C_d) and interface trap capacitance (C_{it}).

The loss of gate control also happens in TFETs when interfacing traps presence. In the meanwhile, the ideal SS is not limited by 60 mV/dec right now. Modern silicon finfet technology can achieve 60 mV/dec at V_{DS} = 0 to 0.7 V. It suggests that there is no more advantage of TFETs over MOSFETs in terms of power loss, if the resulted SS, that is, the magnification of ideal SS by a factor from interface traps, goes greater than 60 mV/dec in TFETs. Therefore, a low D_{it} is of extreme importance in TFET design.

InAs/InP channel is chosen in this study because of their ideal conduction band alignment that eases the optimization of the bound state in the 3-HJ TFET design (chapter 3). In the fabrication of TFETs (chapter 4 and 5), the high-k metal gate (HKMG) process with low D_{it} high-k/InAs and high-k/InP interfaces will be developed and verified as the first step towards realizing 3-HJ TFETs.

Different from MOSFETs, traps affect TFETs not only on gate electrostatics, but also through trap-assisted tunneling (TAT) [34][35][36][37]. Oxide/semiconductor interface traps, as well as heterojunction interface, traps both participate in the process. Fig. 2.14 shows the schematic band diagram of InAs/GaSb and high-k/InAs interfaces with interface traps. At heterojunction interface (a), electrons tunneling from source valence band/channel conduction band into trap states, scattering to upper/lower level

of trap states assisted by phonons, then tunneling to channel conduction band/source valence band. At high-k/InAs interface (b), electrons tunneling from channel valence band/channel conduction band into trap states, and phonon scattering to upper/lower level of states, then tunneling to channel conduction band/channel valence band [38].

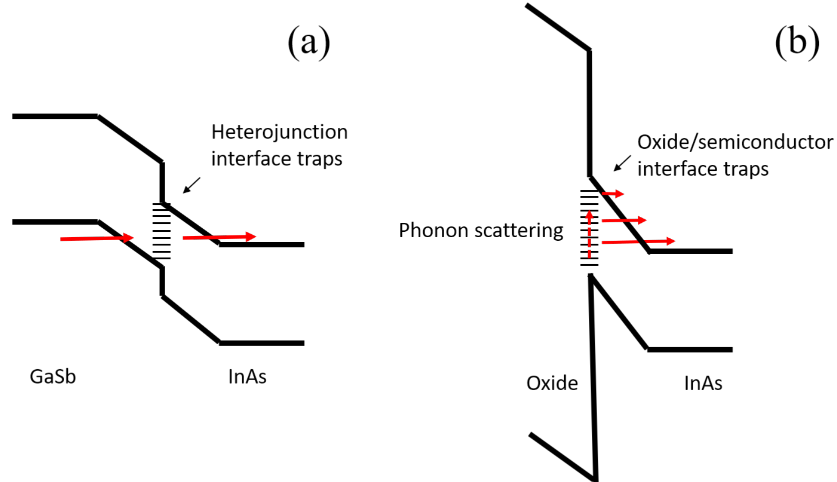


Figure 2.14: Schematic band structures showing TAT given by (a) heterojunction interface traps and (b) oxide/semiconductor interface traps in TFETs. Note that TAT happens two-way (source to channel and channel to source), dependent on the bias.

TAT happens at both on- and off-state. At off-state, the TAT current sets the floor of off-state leakage [36]. At on-state, TAT causes leakage at subthreshold regime, thus degrades SS. Therefore, it is very crucial to keep the interface trap density low to prevent TAT leakage for a low SS and low I_{off} TFETs.

The high-k/somiconductor interface will be optimized through fabrication process development. On the contrary, the interface quality in the heterojunction interface associates with epitaxy. In this work, heterojunctions in TFET are mostly designed lattice-matched to minimize the generation of defects/traps. Detailed optimization and trade-offs will be discussed in chapter 3.

In conclusion, tunneling current in zener diodes is derived, and the maximization of tunneling probability based on material design, heterojunction incorporation, and device

orientation engineering is discussed. It is also shown that electrostatics in TFETs determines tunneling distance. Challenges and potential problems in most modern TFETs are introduced. The possible solutions from the design and fabrication point of view will be proposed and explained in the following chapters.

Chapter 3

High On-Current 3-HJ TFETs with PNPN Doping Profile

We have shown the importance of a high on-current in TFET and its design considerations in chapters 1 and 2. This chapter will explain heterojunction engineering and doping profile engineering to boost device performances in detail (section 3.1 and 3.2). Triple-heterojunction (3-HJ) TFET design with resonant enhancement will be proposed to increase on-current tremendously. The way to design and optimize bound states in 3-HJ TFET will be explained in section 3.3. To relieve the fabrication burden on getting an ultra-thin body, the PNPN doping profile is employed in the 3-HJ TFET in section 3.4.

Both gate electrostatics and junction design make important roles in I_{on} in TFETs. The state-of-the-art TFETs, with highest I_{on} of 10 $\mu A/\mu m$ at V_{DS} and $V_{GS} - V_{TH}$ of 0.3 V, as well as a low SS of 48 mV/dec at V_{DS} of 0.1 V to 0.3 V, using GaSb/InGaAsSb/InAs double heterojunctions design by Lund University at 2016-2017 [39, 40]. This record high on current is still ten times smaller than in silicon CMOS in which 100 $\mu A/\mu m$ at the

supply voltage of 0.3 V is a baseline given SS of 60 mV/dec.

The low drive current can be explained by low transmission and gate overlap. In chapter 2, it has been shown that on-current drops by roughly four times for a gate-overlap case, which suggests that I_{on} is supposed to be around 40 $\mu A/\mu m$ without gate misalignment.

In addition, the transmission of a tunnel junction depends on not only electrostatics but also junction design. After all, tunneling distance needs to be reduced to maximize tunneling probability and further increases drive current. In the following sections, two different ways of reducing tunneling barrier thickness will be discussed in detail.

3.1 Doping Profile Engineering

By doping the UID region to N-type at for a PiN diode, depletion width will be reduced. The amount of thinning in tunneling distance depends on N doping level and thickness. In TFETs, channel potential becomes directly modulated by drain bias considering an N^+ channel without a low-doped drift region for charges storage. This results in the device's inability to be turned off. Therefore, V. Nagavarapu *et al.* proposed a PNPN doping profile, which incorporates an N^+ pocket doping layer at tunnel junction in channel side, and dope the rest of the channel into P well [41], as shown in fig. 3.1 (a). By engineering doping profile with proper optimization, tunneling distance can be reduced with no leakage current compensation.

Fig. 3.1 (b) (c) and (d) shows simulation results of the optimization of PNPN doping profile on Si TFETs from [28]. Body thickness is set to be 10 nm, and pocket doping is $4 \times 10^{19} cm^{-3}$. The energy band diagram at off-state with pocket length L_N ranging from 3 nm to 30 nm is shown in (d). As can be seen, an increase L_N above 4 nm makes this pocket layer start to be partially depleted. The undepleted part of pocket layers

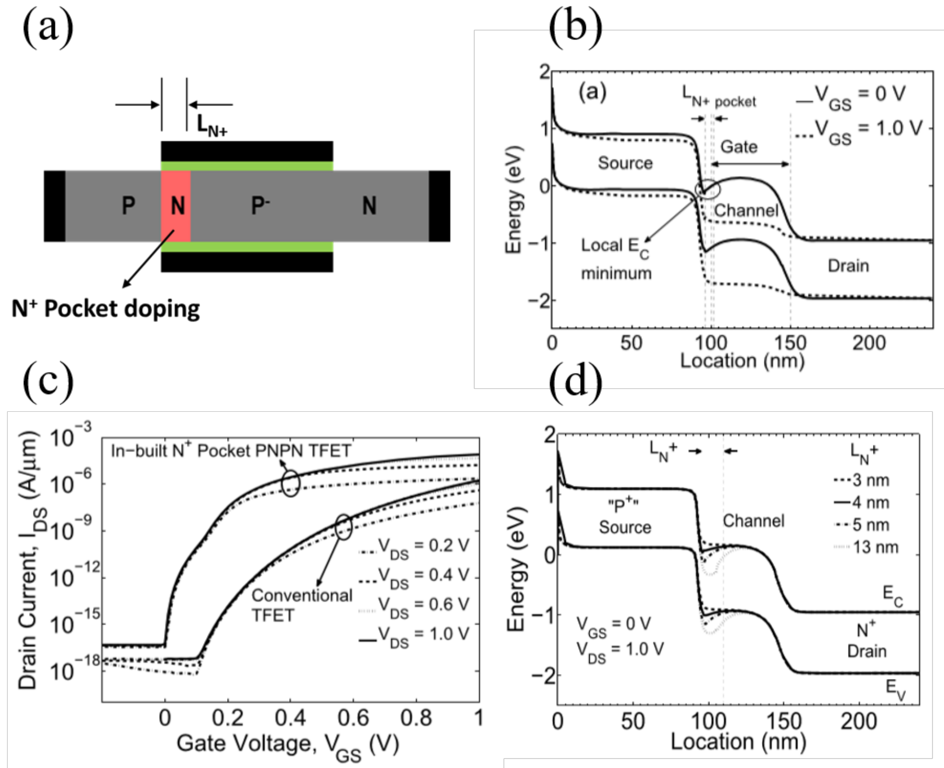


Figure 3.1: (a) Schematic structure of a PNPN doping TFET, and its band diagram at on and off states. (c) I_D - V_G characteristics of a PNPN TFET and conventional P-i-N TFET. (d) Band diagram of PNPN TFETs with pocket doping of $4 \times 10^{19} cm^{-3}$ and pocket thickness of 3 nm to 13 nm. [28]

forms a dip in-band profile, while electron concentration in this regime increases. At off-state, those electrons thermionic emits over P well. Both subthreshold slope and off-state leakage would be degraded. Therefore, it is important to keep the pocket layer fully depleted while optimizing PNPN TFETs, as shown in fig. 3.1 (b). Tunneling distance is thinned down to roughly 3 nm at on-state, while the P-doped channel bends conduction band upward at off-state at $L_N = 4$ nm. The huge improvements in this PNPN TFET's electrical performances compared to PiN TFET are shown in (c). Averaged SS is improved from 65 to 25 mV/dec in TCAD simulation.

3.2 Heterojunction Engineering

As mentioned before, heterojunction engineering is a way to reduce tunneling distance. Incorporating heterojunctions to improve I_{on} has been widely proposed and optimized in Si- and III-V based TFETs [42, 43]. In SiGe heterojunction TFETs, the mole fraction of Ge in the source region needs to be higher than 60% to achieve comparable drive current as CMOS technology [44]. Such a high composition of Ge is very challenging to be realized due to its large lattice mismatch. On the other hand, III-V TFET is promising for its flexibility of heterojunction material design which makes lattice match and thin barrier achievable simultaneously.

Long *et al.* proposed a V_{DD} 120 mV operation of multiple heterojunctions TFETs [45]. With a larger bandgap source and channel barrier, the build-in potential field in a small bandgap tunnel junction dramatically reduces tunneling distance, thus increase tunneling probability. Fig. 3.2 shows the simulation result of improved transmission by incorporating heterojunctions in TFETs. As shown in (b), the tunneling distance shrinks from 6 nm to 1.7 nm for 3-HJ TFET comparing to 1-HJ GaSb/InAs TFET. The transmission reaches $> 50\%$ over 80 meV tunneling window, as seen in (c).

Designing multiple heterojunctions in TFET is not as simple as thinning down tunneling barrier thickness. A thin barrier would result in not only higher on-current but also higher leakage. By introducing larger bandgap source and channel barriers, quantum wells are formed, and the whole tunneling process is split into three tunneling interfaces. Meanwhile, bound states in source and channel wells play an essential role in device transport, which will be discussed next.

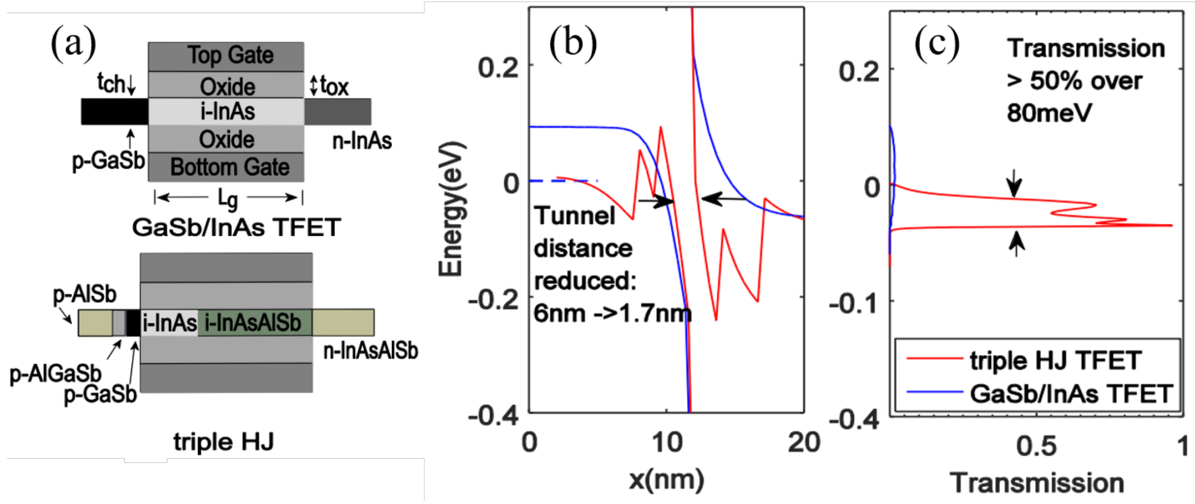


Figure 3.2: Schematic structures of 1-HJ TFET and 3-HJ TFET (a), their energy band diagram at tunneling junction (b), and the energy vs. tunneling probability plots (c). [45]

3.3 3-HJ TFETs with Resonant Enhancement

To maximize tunneling probability for high I_{on}, design of the material, device orientation, and heterojunctions are important. Decision on device orientation (section 3.3.1), material selection (section 3.3.2 and 3.3.3), layer thickness will be discussed (section 3.3.4) later. In this work, 3-hj is employed in TFETs: source barrier, source well, channel well and channel barrier, respectively, as shown in fig. 3.3. The tunneling process involves electrons tunneling through 1. source barrier/source well junction, 2. source well/channel well (main tunneling junction), 3. channel well/channel barrier. The build-in field generated by source barrier/source well heterojunction reduces the tunneling distance in the main tunneling junction (junction 2) at the source side. In contrast, the build-in field generated by channel well/channel barrier heterojunction reduces the tunneling distance in the main tunneling junction (junction 2) at the channel side. In this case, the tunneling window for the main tunneling junction (junction 2) is the region where the thinnest tunneling distance is. As marked as the grey region in fig. 3.3, the tunneling energy

window is the difference in band offsets between conduction band energy of channel well and valence band energy of source well. To be noted that a roughly 0.3 eV tunneling window best aims for an application of a 0.3 V supply voltage, such that the device could be turned on/off within 0.3 V. Given the discussion above, GaSb-based alloy for source well together with InAs-based alloy for channel well is thus an excellent choice to start with.

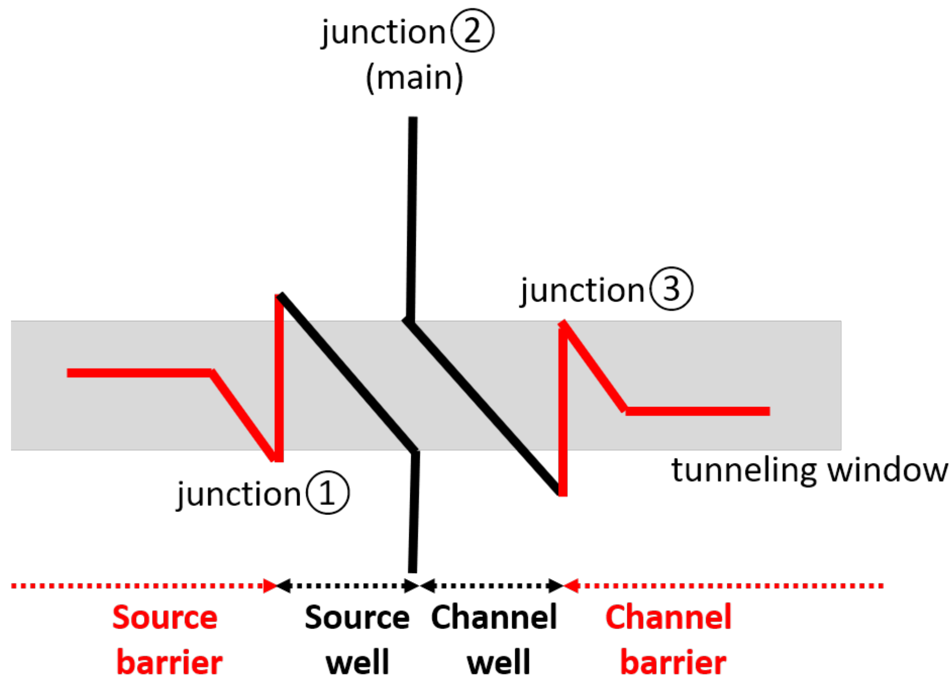


Figure 3.3: Schematic band structure of a 3-HJ TFET that has layers specified from left: source barrier, source well, channel well, channel barrier. By incorporating three heterojunctions, the tunneling interface is split into junctions 1, 2, and 3. The tunneling energy window (marked grey) is in between the conduction band edge of the channel well and the valence band edge of the source well.

3.3.1 Device Orientation

Device orientation plays an essential role in TFETs transmission due to the band structure changes in a confined condition. Its been shown in chapter 2 that [110] transport

and $(1\bar{1}0)$ confinement results in the highest tunneling for most III-V materials. However, orientation engineering is challenging from a fabrication point of view. Large amounts of time and effort on material growth are needed because the optimal orientation is not standard in existing well-developed technology. In this study, a common [100] transport direction is used even with a roughly two times drop in drive current comparing to the optimal [110] transport direction. The fabricated vertical TFET that will be shown in chapter 4 follows this designed orientation.

3.3.2 Design of Heavily-doped Source Barrier and Source Well

As discussed in chapter 2, the source of tunneling distance, given by electrostatics, is associated with source doping. The same idea happens to 3-HJ TFET design. At the source barrier/source well interface, the tunneling distance depends on source well doping concentration. A highly doped source well is needed to maximize the tunneling through junction 1. Typically, source doping of higher than 2×10^{19} cm^{-3} is needed for any TFET design for a thin enough tunneling distance in the source. However, high source doping sacrifices the number of electrons for tunneling. By doping source higher, the fermi-level is pushed further down and depletes the electrons near the source's valence band. The reduction of carriers by a high source doping limits the turn-on speed of TFET (SS) [46]. Additionally, the consequence of electron depletion would be even worse in resonant TFETs since the density of the state is now quantized. A higher SS and lower on-current would result. A source doping concentration of between 2×10^{19} cm^{-3} - 5×10^{19} cm^{-3} is commonly used in III-V based TFETs [23, 47].

The energy barrier at junction 1 depends on the valence band offset between the source barrier and source well. The higher the valence band offset, the higher $E_{barrier}$ at junction 1. However, the built-in electric field reduces the tunneling distance at junction

2 (source side). In other words, a higher valence band offset would enhance tunneling at junction 2 (source side) while diminishing the tunneling at junction 1. Indeed, adding heterojunctions is not applying magic power. It splits one knob into two knobs. One could only modify the tunneling in the source by tuning source doping in a conventional TFET. By employing a junction in source, source doping, material selection, and source well thickness could all be used for optimization.

To determine the material choice for source barrier and well, the material with proper valence band alignment is critical. Fig. 3.4 shows the bandgap heaven for most widely used III-V materials [48]. It is clear that Sb-based material with high valence band energy can be used in the source well, while As- and P-based material with low valence band energy can be used in the source barrier. Indeed, it has been discussed in the introduction of chapter 3 that, GaSb-based source well/InAs-based channel well is a good match concerning tunneling window and their low bandgap.

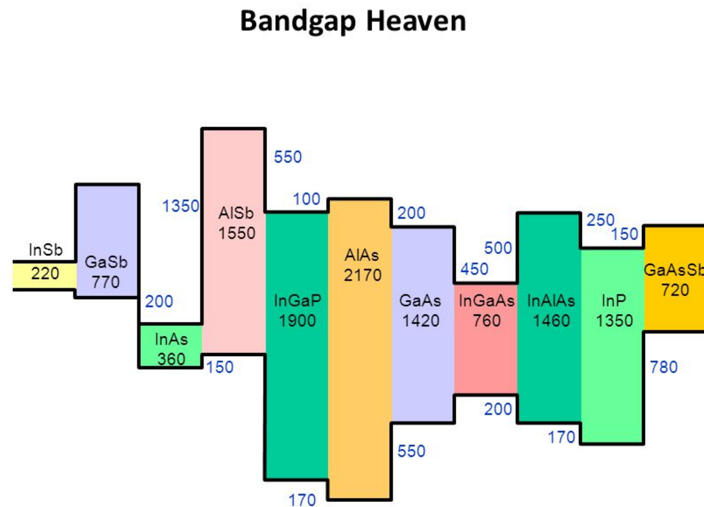


Figure 3.4: Band alignment (bandgap heaven) of widely used III-V materials. [48]

In practice, other material properties need to be considered in source design. A high

source doping of mid- 10^{19} cm^{-3} is required in addition to the proper bandgap energy and band offset. In phosphide materials, it is quite challenging to get hole concentration $> 10^{19}$ cm^{-3} due to hydrogen passivation, drop of activation at high doping level (poor electrical activation), and dopant diffusion [49, 50, 51, 52]. On the contrary, As-based materials, i.e., GaAs and InGaAs and InAlAs, have suitable dopant gaseous sources and $\approx 100\%$ activation, can get to a P-doping level above 10^{20} cm^{-3} [53, 54, 55]. Consequently, an InGaAs or GaAs source barrier is more realistic.

In conclusion, GaSb-based alloy for source well together with InGaAs or GaAs for source barrier gives proper valence band alignment, and high P-doping is doable in these materials to enhance tunneling.

3.3.3 Design of Channel Well and Channel Barrier– Substrate

The tunneling at the channel side in TFET is not merely controlled by doping as in source but also by gate electrostatics. The tunneling distance in the channel is a function of gate dielectric constant, gate oxide thickness, and body thickness. It is, therefore, tough to reduce the tunneling distance; thus, I_{on} is limited.

The same strategy as source design, heterojunctions are added into channel– one knob is split into two. Similar to the source, the channel well, and channel barrier design requires proper band alignment between each layer. As discussed in the last two subsections, a GaSb-based source well with an InAs-based channel well is suitable for the main tunneling junction of high transmission, the supply voltage of 0.3 V 3-HJ TFET design. A channel well with a higher conduction band offset with respect to InAs would increase the build-in electric field in the InAs layer, which boosts the tunneling at junction 2 (channel side). In contrast, the conduction band offset between channel well and channel barrier induces the 3rd tunneling interface– junction 3. The higher the

m_{eff}\E_1	0.4 (eV)	0.3 (eV)	0.2 (eV)	0.1 (eV)
0.1 m_0	2.91	3.36	4.12	5.83
0.2 m_0	2.06	2.38	0.84	4.12
0.3 m_0	1.68	1.94	0.93	3.36
0.4 m_0	1.46	1.68	1.00	2.91
0.5 m_0	1.30	1.50	1.06	2.61

Table 3.1: Approximated quantum well thickness given by an effective mass m_{eff} and a first quantum state energy E_1. Numbers in the table are in nm.

conduction band offset, the higher the energy barrier and thicker tunneling distance at the channel well/channel barrier interface (junction 3).

In highly confined conditions, effective mass and bandgap increase. The electron/hole effective mass of confined III-V materials lies roughly between 0.1 m_0 to 0.5 m_0, where m_0 is electron mass. Consider a band offset of 0.1 to 0.4 eV, the quantum well thickness (in nm) (approximated by the first bound state in an infinite rectangle quantum well) is summarized in table 3.1. The optimal thickness of source/channel wells roughly lies between 1 - 5 nm. It is noted that if a material with a higher effective mass is used, the optimal quantum well width could be thinned down to sub-1 nm.

Even though the channel design does not require evaluating the feasibility of high-doping as in source, the high-k/semiconductor interface has to be taken into account. High quality high-k/InAs interface has been developed at UCSB using ZrO_2 with ALD in-situ nitrogen plasma treatment beforehand [56, 57]. A near-to-ideal SS of 60 mV/dec is achieved, which suggests InAs is a good choice for channel well in 3-HJ TFET in both design and fabrication [58].

The decision on channel barrier material requires careful thoughts. First, it is part of the channel; thus, a decent high-k/channel barrier interface is also needed. Second, the thickness of the channel barrier sets gate length. In other words, channel barrier thickness will be $\approx$ 30 nm or more, which is the minimum thickness to support gate to drain potential difference for a V_{DD} of 0.3 V device [23]. At a thickness of 30 nm

or above, a lattice-matched material for channel barrier with respect to the substrate is required. Third, one of the solutions to the ambipolar leakage problem, as discussed in 2.4.1, suggests using a large bandgap material at the channel/drain junction. This specification coincides with the demand for high conduction band energy– a high bandgap material satisfies both.

3.3.4 Roles of Resonant Bound States and Optimization

The principle of adding heterojunctions in TFETs is to reduce tunneling distance. Meanwhile, states are introduced in source and channel quantum well, forming a resonant structure at the main tunneling junction. The presence of bound states in TFET affects the electron transport both at on and off-state.

At on-state, resonant states enhance transmission. The overall tunneling could be maximized by tuning the transmission of junctions 1, 2, and 3, as shown in fig. 3.3. By aligning the bound state energy in the source well to that in channel well, as well as to the thinnest tunneling distance region (where has energy lying between the conduction band of channel well and valence band of source well) at the main tunneling junction, wavefunction overlaps, thus tunneling is boosted. Figure 3.5 shows the energy band diagrams of 3-HJ TFETs with (a) well-aligned/ (b) not-aligned quantum state energies in source and channel well and the resulted tunneling probability. A near to 1 transmission in (a) is due to the excellent alignment of bound state energies, while the transmission drops to 2% in (b) in the not-aligned case.

Bound states at off-states, however, increases leakage current. The evanescent tails that extend from the source(channel) well into the channel(source) well induce tunneling leakage. Even without overlapping bound state energies, electrons could still tunnel through and jump to a higher/lower energy state assisted by phonon scattering. This

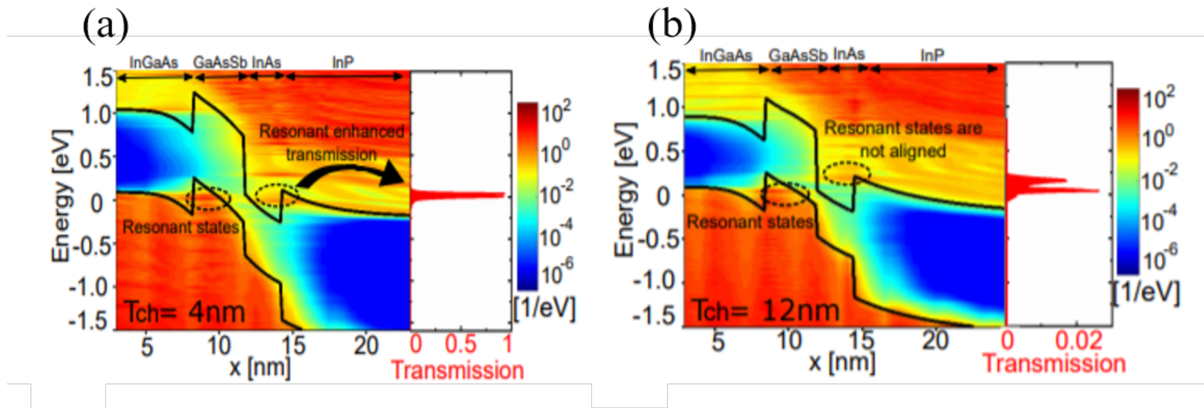


Figure 3.5: Energy band diagrams and energy vs. tunneling probability of 3-HJ TFETs with (a) aligned bound states in channel and source well, and (b) misaligned quantum state energies. © 2021 IEEE

phonon-assisted scattering process must be taken into account in the resonant TFET design.

Ideally, the bound state energies need to be designed to lie near the well edge to suppress this leakage current. Considering a band offset of 0.3 eV for both source barrier/source well and channel well/channel barrier, aligning the bound state energies to the edge of quantum well at both sides will give a 0.3 eV potential barrier at off-state. This 0.3 eV energy barrier between bound states and the existing main tunneling junction barrier is enough to eliminate phonon-assisted tunneling leakage.

In order to move the bound state energies up and down, the thickness of quantum wells needs to be adjusted correspondingly. The higher the effective mass is, the thinner the good thickness is needed for a given location of states in energy space. Based on the material property and the depth of the quantum well, the resonant states at two sides of the main tunneling junction can be optimized accordingly. Fig. 3.6 (a) shows the band diagrams of InAlAs/GaAsSb/InAs/InP TFETs with GaAsSb/InAs thickness of 1.8/2.4 nm (case 1), 2.4/2.4 nm (case 2), and 4/2.4 nm (case 3). It is clear to see that the bound states in GaAsSb are lower than that in the channel given a 1.8 nm width (case

1), while the 1_{st} bound state is aligned with the channel at a thickness of 2.4 nm (case 2). A further increase in the GaAsSb thickness to 4 nm makes quantum well deep; thus, the bound state energy in the source well increases. The channel states now align to the higher states in the source well.

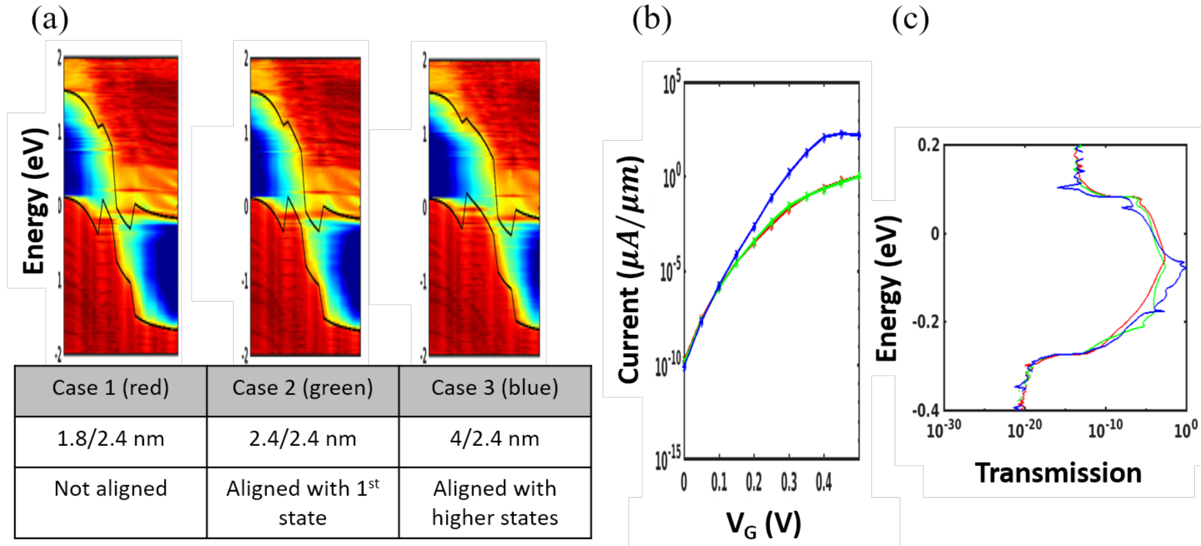


Figure 3.6: (a) Band diagrams of InAlAs/GaAsSb/InAs/InP 3-HJ TFETs with source well (GaAsSb) thickness ranging from 1.8 nm (case 1), 2.4 nm (case 2), to 4 nm (case 3). (b) and (c) the I_D - V_G characteristics and energy vs. tunneling probability for three cases. Courtesy of Chin-Yi Chen.

The I_D - V_G characteristics and transmission are shown in fig 3.6 (b) and (c), respectively. Case 1 and 2 appear similar low on-current while the I_{on} for case 3 is $\approx$ two orders of magnitude greater than others. This agrees with the transmission in (c) that case 3 has much higher transmission than cases 1 and 2. Case 2 has a lower transmission because of the junction between the source barrier/source well. As has been discussed before, tunneling in 3-HJ TFET requires tunneling through 3 junctions/interfaces. In order to enhance the tunneling through the source barrier/source well junction, the bound state energy needs to be designed near the well edge. In case 2, bound state energy in the source well roughly lies at the middle of the well, which results in a relatively large

tunneling distance at the source barrier/source well junction. It is noted that the source doping in the study in fig. 3.6 is $5 \times 10^{19}\ cm^{-3}$.

Fig. 3.7 shows the band diagram of case 3 TFET at different V_G. The V_G is already shifted by threshold condition defined at $I_{off} = 1 \times 10^{-3}\ mA/\mu m$. By decreasing gate bias, the channel potential is pulled up, and the bound state in the channel well is no longer aligned with bound states in the source. At $V_G = 0V$, the tunneling is blocked by thick barrier thickness both between GaAsSb/InAs and InAlAs/GaAsSb, even with aligned bound states. Thus, there is no leakage current, as shown in fig. 3.6 (b).

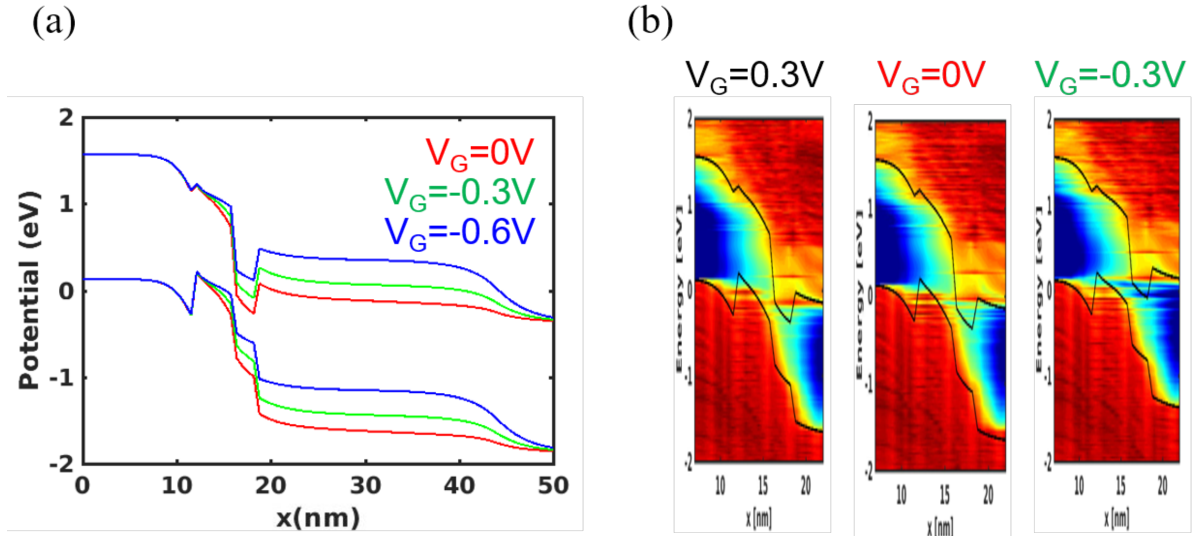


Figure 3.7: Energy band diagrams of InAlAs/GaAsSb/InAs/InP 3-HJ TFET (case 3) at V_G of 0 V, -0.3 V, -0.6 V (a), and the density of state distribution showing bound state energy alignment at V_G of 0.3 V, 0 V, -0.3 V. Courtesy of Chin-Yi Chen.

In conclusion, resonant states in source and channel well should align to maximize transmission. The bound state energy in both wells is better designed near the well edge for a low leakage at off-state. Given the effective mass of III-V materials, the quantum well thickness would lie between 2 nm to 5 nm in 3-HJ TFET. The thickness constraints in material growth and doping and high-k/semiconductor interface (3.3.2 - 3.3.3) should be considered in practical 3-HJ TFET design.

3.3.5 InGaAs/GaAsSb/InAs/InP TFET

Numbers of simulation works have been done on 3-HJ TFETs [45, 25, 23, 24]. A lot of them show high $I_{on} > 100\mu A/\mu m$ using a wide variety of III-V alloys. Most of those designs could not reach fabrication because the physical limitations in material growth and processing are not considered. Based on all discussion that has been examined so far (3.3.1 - 3.3.5), a realistic high I_{on} 3-HJ TFET design is proposed.

The 3-HJ TFET utilizes a P-doped InGaAs source barrier, a P-doped GaAsSb source well, a U.I.D. InAs channel well, a U.I.D. InP channel barrier and drain. InP substrate is chosen for its moderately high bandgap and modest lattice constant– it is lattice-matched to other III-V alloys widely studied and has well-developed technology. The lattice-matched InGaAs and GaAsSb source can not only fulfill the requirements on band offsets but also be grown on the substrate without considering Matthew Blakeslee limit– strain, relaxation, and defects [59, 60].

The Purdue team does the simulation/modeling work. Detailed discussion on methods, modeling, and simulation/experiment benchmarking could be found in Chin-Yi Chen's thesis at Purdue University. Nanoelectronics Modeling tool NEMO5 is used to optimize the 3-HJ TFET [61]. The strain layer is modeled by the atomistic tight-binding method, which is parameterized to hybrid functional calculations [62]. Quantum confinement effect, scattering between electrons, and phonon-assisted scattering are included in the simulation [63]. The mode-space approach developed in [64] is employed to model devices with a large dimension which will be discussed in section 3.4.

Table 3.2 summarizes the confined bandgap (E_g) and valence band offset (ΔE_v) with respect to valence band of InGaAs, and fig. 3.8 shows (a) the full band diagram, (b) the 3-HJ TFET structure at a t_{body} = 4 nm. The high-k is set to be 3.2 nm with a dielectric constant of 15, from experimental results. It is noted that the $In_{0.53}Ga_{0.47}As$

	InGaAs	GaAsSb	InAs	InP
E_g (eV)	0.95	0.99	0.7	1.5
ΔE_v (eV)	0	0.42	0.046	-0.38

Table 3.2: Confined bandgap (E_g) and valence band offset (ΔE_v) with respect to valence band of InGaAs at a t_{body} = 4 nm.

and $GaAs_{0.51}Sb_{0.49}$ source is lattice-matched to InP, while the InAs channel is strained to InP. Thus, the thickness of the InAs layer is kept below 2.5 nm during the optimization.

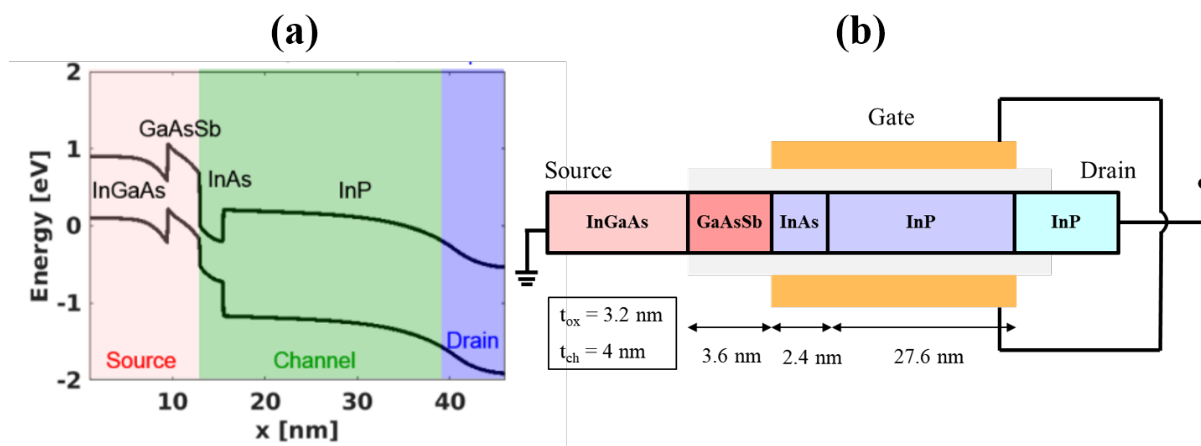


Figure 3.8: (a) Full band diagram of proposed high on-current InGaAs/GaAsSb/InAs/InP 3-HJ TFET design and (b) the structure design in the simulation set up. Channel thicknesses are set to 4 nm, and high-k thickness is 3.2 nm with a dielectric constant of 15.

The optimized doping concentration of P-InGaAs and P-GaAsSb is 5×10^{19}, and N-InP drain is 2×10^{19}. The optimized thickness for the high I_{on} is 3.6/2.4 nm for GaAsSb/InAs, respectively. The gate length is set to 30 nm, which is the minimal number ensuring good gate electrostatics at a body thickness of 4 nm in this design.

Fig. 3.9 shows the band diagram of the proposed 3-HJ TFET at on-state (V_G = 0.3 V). Due to a larger valence band offset between InGaAs and GaAsSb (0.42 eV), multiple quantum states are formed with thin quantum well width of 3.6 nm at the source. A further shrink down the thickness of GaAsSb would decrease the density of states in the

tunneling window (marked between dash lines), thus decreasing on-current. The I_{on} for this design at $V_{DS} = V_{GS} - V_{TH} = 0.3$ V is 98 $\mu A/\mu m$ considering phonon scattering effect.

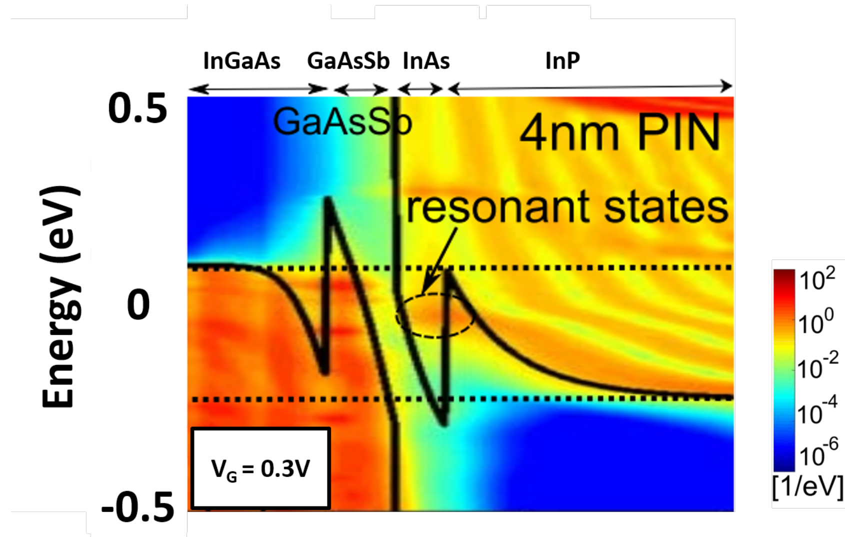


Figure 3.9: Band diagram of proposed 3-HJ TFET at on state ($V_G = 0.3$ V). The resonant state in the InAs channel well aligns to the bound states in the GaAsSb source well and lies in the tunneling window (marked in between dash lines). © 2021 IEEE

3.4 Thick Body– Doping Profile Engineering on InGaAs/GaAsSb/InAs/InP TFET

In section 3.3, an InGaAs/GaAsSb/InAs/InP TFET has been designed and optimized considering physical restrictions, including the highest doping level in different III-V materials, Matthew Blakeslee limit of strain layer epitaxy, the quality of the high-k/semiconductor interface, and device orientation. One constraint that has yet to be considered is– thick body.

As discussed in chapter 2, an ultra-thin body is needed for high tunneling, thus a

high on-current. In most works on TFET design and modeling, a t_{body} of sub- 5 nm is usually used to ensure good gate electrostatics. The fabrication of such a thin body is challenging and becomes the main obstacle of proof of concept. In this section, a different path to tackle this problem is proposed– doping profile engineering.

As shown in fig. 3.10, electrons in a thin body that is enclosed by high energy gap dielectrics at two sides can be seen as waves confined in a quantum well. Increase the body thickness results in the loss of confinement. More modes are formed, and wavefunctions leak into the bulk channel. The bound state energy would also reduce. Bad electrostatics is a result.

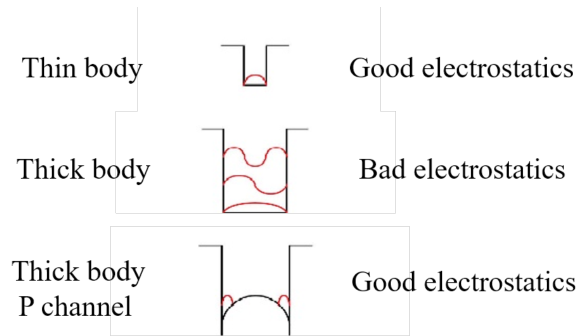


Figure 3.10: Electrostatics of wave functions in thin body and thick body wells. Electrons are highly confined in a thin well and spread over the whole body in a thick well. A P body doping supplements the loss of confinement, thus improves electrostatics.

Employing P-doping in the channel is an alternative solution. By body dope to P-type, a triangle quantum well is formed at on-state. When applying positive V_{GS}, the semiconductor surface potential is pushed downward. A higher potential at the center of the channel given by sufficient P body doping prevents the whole body from fully depleted by the gate. These triangle quantum wells provide the confinement that is lost in a thick body. Electrostatically, the higher P body doping, the steeper the triangle well edges are. In order to ensure reasonable gate control, there is minimal body doping for a particular body thickness.

In addition, the P-doped body improves the short channel effect. 30 nm gate length

	InGaAs	GaAsSb	InAs	InP
E_g (eV)	0.80	0.85	0.50	1.38
Δ E_v (eV)	0	0.43	0.067	-0.39

Table 3.3: Confined bandgap (E_g) and valence band offset (ΔE_v) with respect to valence band of InGaAs at a t_{body} = 12 nm.

is optimized on a TFET with a t_{body} 4 nm and a t_{ox} 3.2 nm. A body thicker than 4 nm might cause drain coupling to the channel. One could increase gate length accordingly (following $E \propto 1/(W_D + \lambda(T_{ch}))$, where λ is scaling length and W_D is depletion width [65]) or increase P doping level in channel [66].

By P-doped the InP channel barrier, the electric field at the main tunneling junction drops. This results in a low tunneling probability and a low current. Therefore, an N pocket doping is needed in-between P-source and P-channel– where InAs channel well is. With N-type pocket doping in InAs, tunneling distance would be thinned down by a high build-in field between P-GaAsSb/N-InAs. Since the InAs channel is quite thin (2.4 nm), a high N doping comparable with source doping is required. In this study, an N-InAs doping of 5×10^{19} cm^{-3} is used.

3.4.1 Thin body (4nm) versus Thick body (12nm)

To investigate the effect of thick body, atomistic mode space approach is applied in the simulation [64]. Table 3.3 summarizes the confined bandgap (E_g) and valence band offset (ΔE_v) with respect to valence band of InGaAs at a t_{body} = 12 nm. To be noted that a 10% to 40% reduction in bandgap comparing to table 3.2 is as a result of a lower confinement.

Fig 3.11 compares the band structure of InGaAs/GaAsSb/InAs/InP TFETs with P-i-N doping profile for a t_{body} = 4 nm (a) and a t_{body} = 12 nm (c) and with PNPN doping profile for a t_{body} = 4 nm (a) and a t_{body} = 12 nm (c) at V_G = 0.3 V. The doping

concentrations of N-InAs and P-InP are 5×10^{19} cm^{-3} and 2×10^{19} cm^{-3}, respectively.

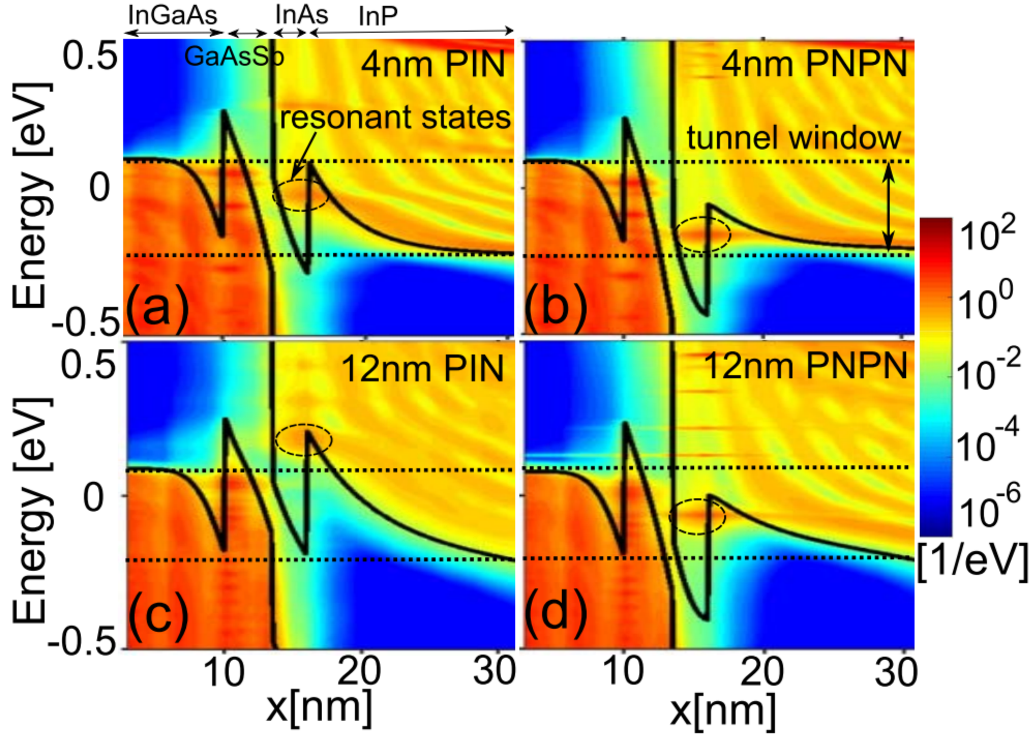


Figure 3.11: Band structures and resonant state alignment of InGaAs/GaAsSb/InAs/InP TFETs with (a) P-i-N at t_{body} = 4 nm, (b) PNPN at t_{body} = 4 nm, (c) P-i-N at t_{body} = 12 nm, (d) PNPN at t_{body} = 4 nm. ©2021 IEEE

At 12 nm, the loss in gate control results in a misalignment between bound states in source and channel well for a P-i-N structure (c). In order to turn on this device, more gate bias is needed to push down further the channel potential such that bound state energies align. By employing doping in the channel (b) and (d), channel potential is pushed down. The bound state energy becomes well-aligned with that in the source well. The enhancement in junction field by PNPN doping profile is shown in fig. 3.12. The maximum electric field is improved by ≈ 60% by applying a PNPN doping profile for a t_{body} = 12 nm. Comparing P-i-N and PNPN TFETs, a ≈ 28% and ≈ 12% improvement in E-field for t_{body} = 4 nm and t_{body} = 12 nm can be seen. This suggests that incorporating the PNPN doping profile in TFET design makes the junction field less sensitive to the

Doping Profile (t_{body})	P-i-N (4 nm)	P-N-P-N (4 nm)	P-i-N (12 nm)	P-N-P-N (12 nm)
I_{on} ($\mu A/\mu m$)	98	248	6	325

Table 3.4: Summary of I_{on} at $V_{GS}-V_{TH}$ = V_{DS} = 0.3 V for InGaAs/GaAsSb/InAs/InP TFETs with different body thickness and doping profiles.

variation to body thickness.

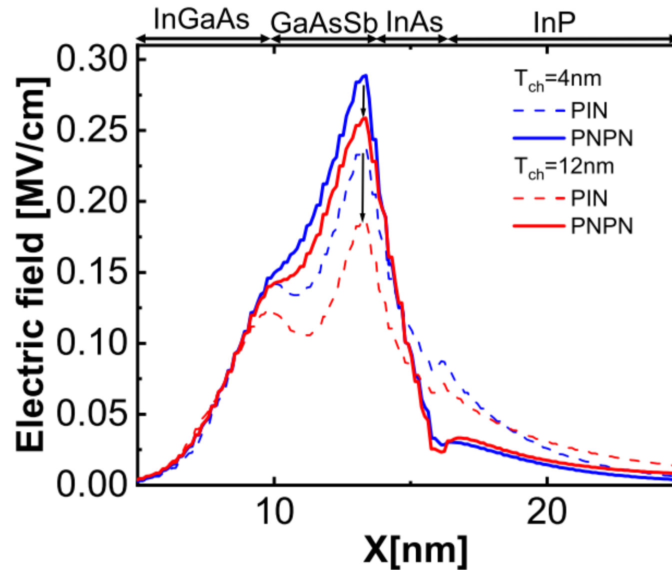


Figure 3.12: Electric field vs. distance for InGaAs/GaAsSb/InAs/InP TFETs with different body thickness and doping profiles. ©2021 IEEE

Table 3.4 summarizes the on-current of the 3-HJ TFETs with P-i-N/PNPN doping profile and t_{body} = 4/12 nm at $V_{GS}-V_{TH}$ = V_{DS} = 0.3 V. A high on-current of ≈ $300\mu m/\mu m$ for PNPN design for both 4 and 12 nm body suggests that this design is promising for the next generation low power logic application.

3.4.2 Effect of P-InP body doping level

The doping concentration in P-InP affects the confinement of electrons in the channel. In principle, the higher P doping, the lower wavefunction leaks into the center of the body. Fig. 3.13 compares the band diagram at the high-k/semiconductor interface of

3-HJ TFETs with t_{body} = 12 nm and InP channel having doping concentration of (a) U.I.D, (b) $P - 1 \times 10^{16}$, (c) $P - 5 \times 10^{18}$, (d) $P - 2 \times 10^{19}$ cm^{-3} at V_{GS} = 0.3 V. The doping concentration of N-InAs is 5×10^{19} cm^{-3}. As can be seen, bound states energy in channel well becomes more well confined as P-doping increases, such that the bound state energy aligns closer to the energy window that has thinner tunneling distance, as shown in (d).

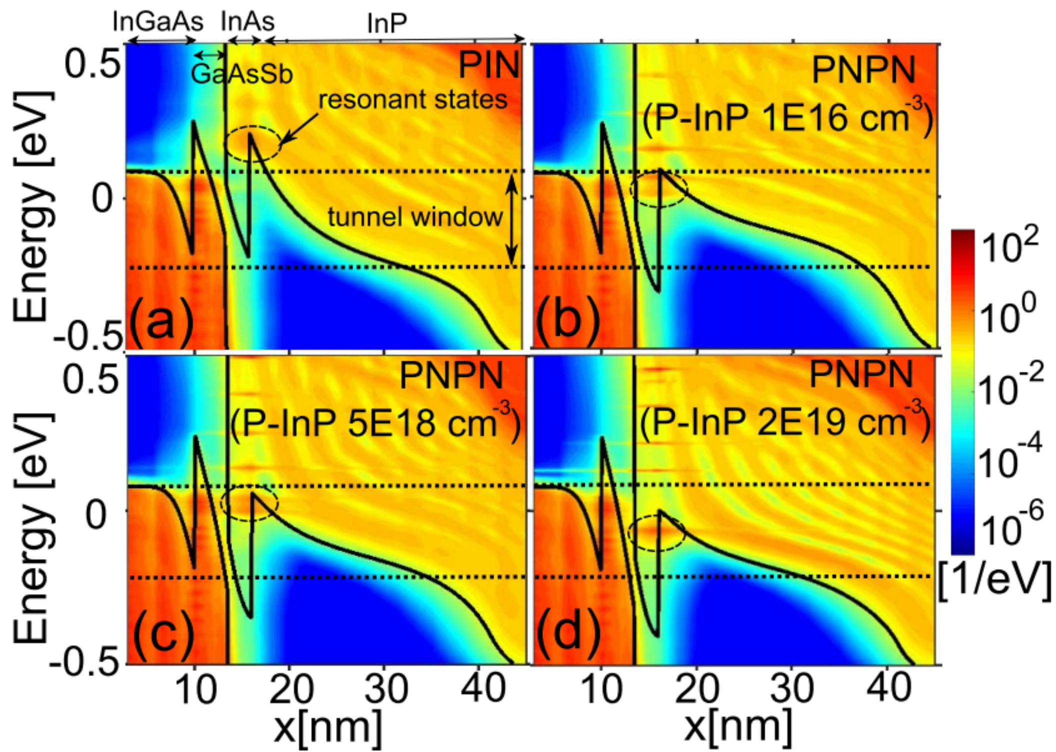


Figure 3.13: Energy band diagrams of InGaAs/GaAsSb/InAs/InP TFETs with InP channel doping of (a) U.I.D, (b) P-1× 10^{16}, (c) P-5× 10^{18}, (d) P-2× 10^{19} cm^{-3} at $V_{GS}-V_{TH}$ = 0.3 V. ©2021 IEEE

Moreover, a P-doped InP channel flattens the potential drop in the InP region. Without a high N^{+} doping in thin InAs, the electric field at GaAsSb/InAs and InAs/InP would both drop dramatically. The incorporation of N^{+} pocket doping pushes down channel potential, thus reducing tunneling distance and aligning resonant state energies.

In addition, the control of the gate to the whole 12 nm thick body can be seen in fig.

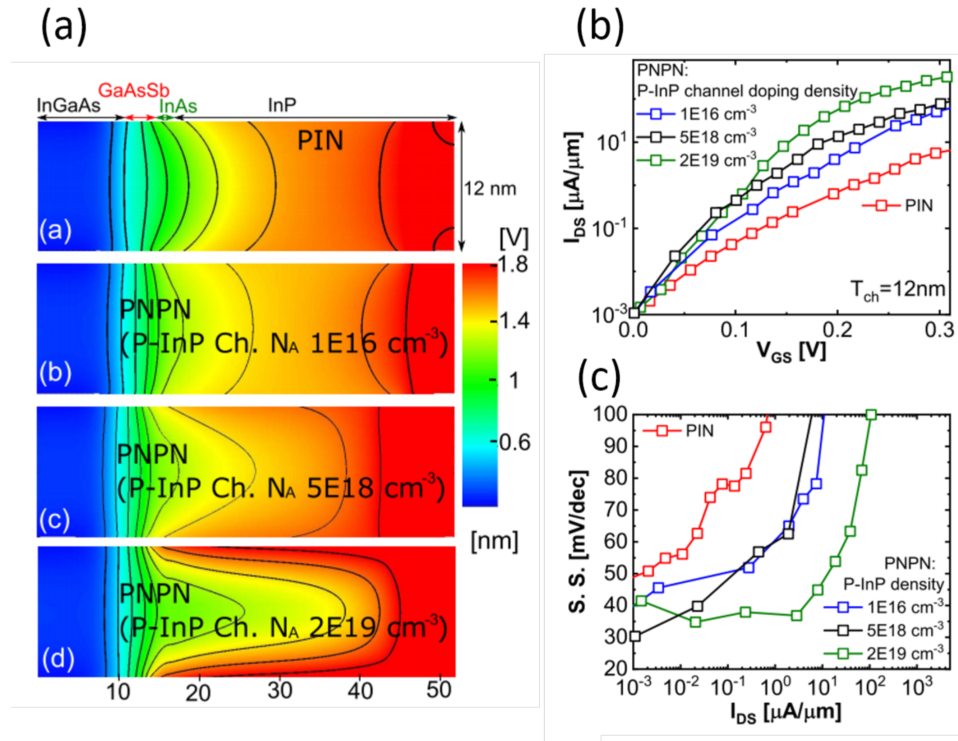


Figure 3.14: (a) Equal potential contours of 3-HJ TFETs with InP doping density of (a) U.I.D, (b) P-1× 10^{16}, (c) P-5× 10^{18}, (d) P-2× 10^{19} cm^{-3}. (b) I_D - V_G curves and (c) SS vs. I_D plots for 4 cases. ©2021 IEEE

3.14 (a). The equal energy potential contour of 3-HJ TFETs using P-i-N design shows a sparse distribution, suggesting the body's center is weakly controlled by gate bias. On the contrary, a higher P doping design shows independent potential energy at body center to surface. In addition, the much denser contours at the high-k/semiconductor interface in a highly doped channel design, meaning a high junction field, thus a thin tunneling distance. This agrees with the sharp profile seen in fig. 3.13 (d) comparing to lower doping cases (a), (b), and (c).

The I_{DS} - V_{GS} characteristic and SS versus I_{DS} of the four designs are summarized in (b) and (c), and the on-current at $V_{GS}-V_{TH}$ = V_{DS} = 0.3 V is summarized in table 3.5. High P-InP doping is needed for a high I_{on} and low SS at a thick body condition. Furthermore, the floating body for P-i-N or low doped InP design induces leakage current

Doping Profile P-InP doping	P-i-N –	P-N-P-N 1E16	P-N-P-N 5E18	P-N-P-N 2E19
I_{on} ($\mu A/\mu m$)	6	50	78	325

Table 3.5: Summary of I_{on} at $V_{GS}-V_{TH}$ = V_{DS} = 0.3 V for InGaAs/GaAsSb/InAs/InP TFETs with InP channel doping of (a) U.I.D, (b) P-1$\times 10^{16}$, (c) P-5$\times 10^{18}$, (d) P-2$\times 10^{19}$ cm^{-3}.

at low V_{GS}, while the design with doping of 2$\times 10^{19}$ cm^{-3} has the parasitic tunneling diode at the center of the body kept off. Therefore, subthreshold leakage current is minimized, and minimum SS $\approx$ 40 mV/dec ranges widely in I_{DS} ($\leq 1\mu A/\mu m$).

Fig. 3.15 investigates the transport throughout the whole body for t_{body} = 12 nm 3-HJ TFETs by cross-sectional band profiles (at $V_{GS}-V_{TH}$ = 0.3 V) cut at high-k/semiconductor interface (a) and at the center of body (b). Comparing the band diagrams of the TFET with P-InP doping density of 2×10^{19} cm^{-3} level, it is clear that the potential of InP at the channel center is pulled up. Tunneling at the channel center is thus blocked by this high bandgap InP channel well. Therefore, bulk tunneling is suppressed while the current is mainly contributed from surface tunneling in PNPN TFETs. In addition, the loss of gate control over the floating body creates parasitic tunneling diodes at the channel center. Without sufficient P body doping, tunneling leakage would result at the subthreshold regime. Once the $I_{DS} - V_{GS}$ curves are shifted by defined threshold condition (which is I_{OFF} = 1 $nA/\mu m$ in this study), the on-state current would drop according to the amount of subthreshold leakage current. Therefore, the channel potential energy for a lightly P-doped channel is higher at on-state, as can be seen in (a).

In conclusion, the PNPN doping profile improves the electrostatic control of 3-HJ TFETs. By employing a highly N^+ doped InAs channel well as well as P-InP channel, I_{on} could be improved by $\approx$ 3 times for a thin body TFET (t_{body} = 4 nm) and by $\approx$ 50 times for a thick body TFET (t_{body} = 12 nm). The enhancement of on-current is because

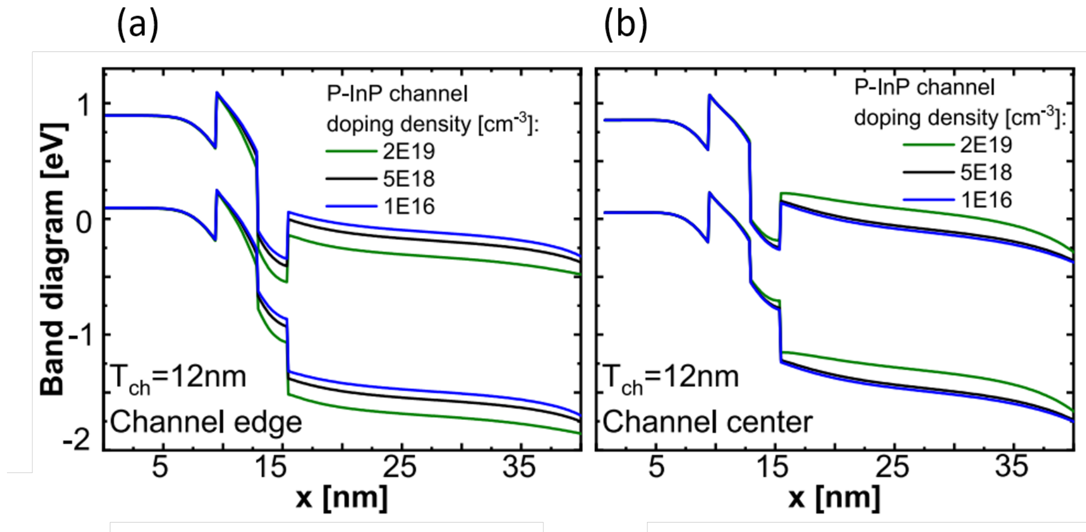


Figure 3.15: Energy band diagram of InGaAs/GaAsSb/InAs/InP TFETs at (a) body edge (near to channel/high-k interface) and (b) body center with different InP channel doping. ©2021 IEEE

of the increased electric field at the tunneling junction and the effective turning off of the parasitic tunneling diodes at the floating body region.

3.5 Design Finalization and Potential Issues

A high on-current InGaAs/GaAsSb/InAs/InP 3-HJ TFET with PNPN doping profile has been optimized in 3.3 and 3.4. In order to finalize the design before heading to device fabrication, ambipolar tunneling leakage at the P-InP channel/N-InP drain junction will be investigated in 3.5.1. The leakage current analysis is done based on the actual device dimension that will be proposed in chapter 5. The finalized epi structure will be shown at the end of the section. In addition, a big challenge of realizing a high-performance 3-HJ TFET– gate alignment will be discussed in 3.5.2, and a solution will be proposed

in chapter 5.

3.5.1 Ambipolar Leakage in PNPN deisgn

Heavily-doped N-InAs and P-InP channels are incorporated in 3-HJ TFET to increase on-current. Aside from enhancing the junction field at the source/channel junction, the electric field at the channel/drain junction is also changed. As discussed before, the floating body for a thick body TFET contributes less current at on-state while giving rise to subthreshold leakage. Even though the body tunneling leakage is suppressed by a high potential P-InP well, the tunneling leakage that happens at the P-InP/N-InP junction is actually increased. This ambipolar leakage current gets worse as the P-InP channel is doped higher.

Bandprof simulator is used to analyze ambipolar leakage current in floating body 3-HJ PNPN TFETs. It is noted that this simulation is performed under bulk conditions, which is valid since the quantum confinement effect is negligible in a body thickness $>$ 20 nm. Fig. 3.16 (a) shows the band diagrams of InGaAs/GaAsSb/InAs/InP PNPN TFETs with the doping concentration and layer thickness all the same as in section 3.4. Gate is kept floating while drain is under forward bias. As can be seen in (a), tunneling distance at P-InP/N-InP junction becomes shorter as V_{DG} increases. Even with a high potential InP barrier that effectively suppresses the tunneling leakage at the channel center through the main tunneling junction, the tunneling at the P-InP/N-InP interface becomes negligible as V_{DG} goes higher. Fig. 3.16 (c) plots the ambipolar leakage at different V_{DG}. The calculation is done using the band structures in (a) and (b), and zener tunneling current eq. 2.8. At V_{DG} = 1 V, the normalized ambipolar leakage current for a t_{body} = 100 nm TFET is $\approx$ 0.1 $\mu A/\mu m$. This leakage is within three orders of magnitude comparing to the expected on-current as shown in 3.4. Therefore, action is needed to reduce this

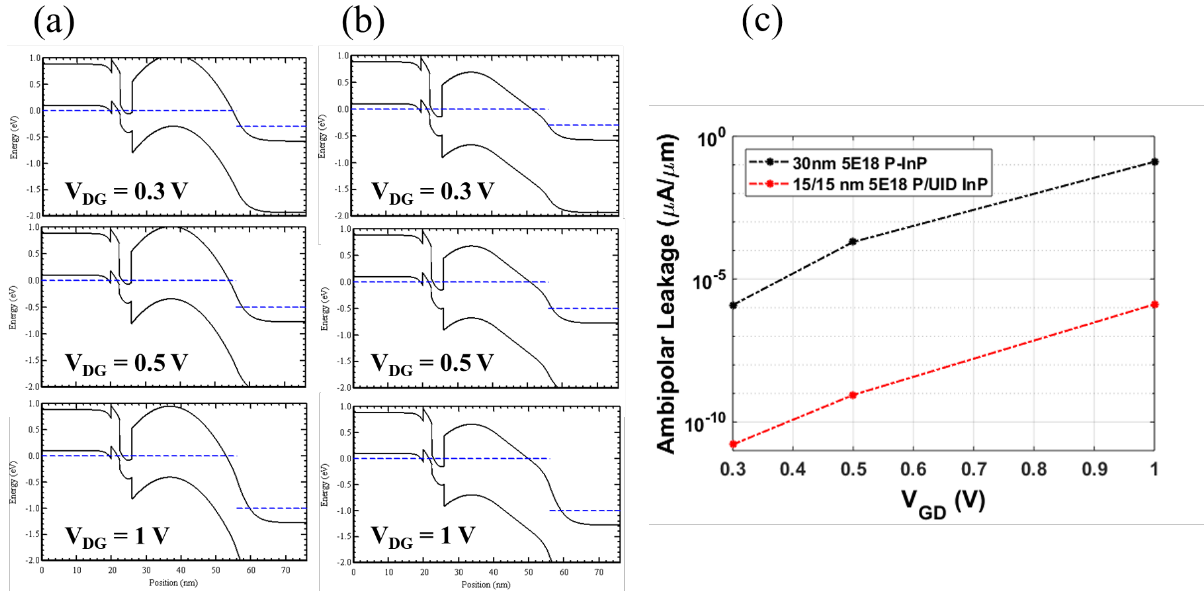


Figure 3.16: Band diagrams of InGaAs/GaAsSb/InAs/InP PNPN TFETs in bulk condition at different V_{DG} with InP channel of (a) 30 nm P-5$\times$ 10^{18} cm^{-3}, and (b) 15 nm U.I.D. + 15 nm P-5$\times$ 10^{18} cm^{-3}. (c) Ambipolar leakage current vs. V_{GD} for 2 channel designs.

P-InP/N-InP tunneling leakage.

A lower junction field is required to suppress the tunneling at the channel/drain junction. By inserting a drift U.I.D. InP layer in between highly doped PN junctions, the tunneling distance is increased. Fig. 3.16 (b) shows the band diagram of a similar structure but a 15 nm P-InP (5$\times$ 10^{18} cm^{-3}) + 15 nm U.I.D. InP channel design. The ambipolar leakage, as shown in (c), drops by $\approx$ five orders of magnitude compared to a 30 nm P-InP (5$\times$ 10^{18} cm^{-3}) design. Without increasing gate length, this method sacrifices around 0.2 eV energy barrier in the channel, which is acceptable since the relatively large bandgap of InP is already enough to block the tunneling leakage at the main tunneling junction.

Table 3.6 shows the finalized epi structure of 3-HJ PNPN TFETs. TO be noted that since it is challenging to dope P-InP $> 5 \times 10^{18}$ cm^{-3}, the actual doping concentration is on a best effort basis. The N^+ InAs is compressive strained to InP. N^+ InP drain layer

Layer	Comment	Material	x	Thickness (Å)	Dopant	Level (/cm3)	Type
9	Drain Contact	In(x)Ga(1-x)As	0.532	100	Si	>3E+19	N+
8	Drain	InP		100	Si	>2E+19	N+
7	Channel drift	InP		126			UID
6	Channel Well	InP		150	Be	best effort	P+
5	Tunnel Junction N+	InAs		24	Si	5.00E+19	N+
4	Tunnel Junction P+	GaAs(x)Sb(1-x)	0.51	36	C	5.0E+19	P+
3	Source	In(x)Ga(1-x)As	0.532	250	C	5.0E+19	P+
2		In(x)Ga(1-x)As	0.532	750	C	2.0E+19	P+
2	Etching stop	InP		30			UID
1	Buffer	In(x)Al(1-x)As	0.52	1,000			UID
	Substrate	3" InP:Fe					

Table 3.6: Finalized epi structure of InGaAs/GaAsSb/InAs/InP 3-HJ PNPN TFET.

is kept at 10 nm to ease the process while a 10 nm N^+ InGaAs ($> 3 \times 10^{19}\ cm^{-3}$) layer is grown for a low resistance contact.

3.5.2 Potential Issues– Doping Fluctuation and Trap-assisted Tunneling

Before finfet is proposed, CMOS scaling is carried out by simply increase body doping while scaling down physical gate length to ensure a good gate electrostatic. Such an increase in doping gets more challenging as scaling goes further because of doping fluctuation. When the device dimension shrinks, the number of dopants in the layer is also reduced. For a highly scaled device with dimension at the level of 10 nm and doping at the level of mid- 10^{19} cm^{-3}, a difference in 1 single dopant incorporation would make a huge difference in local transport characteristics. For example, considering a nanowire diameter of 10 nm targeting a doping concentration of 5×10^{19} cm^{-3}, ≈ 3.7 dopants needs to be incorporated into ≈ 750 atoms per nanometer of growth. More dopant incorporation per nanometer in nanowire growth would cause the local doping concentration to increase to 6×10^{19} cm^{-3}. Needless to say, the uniformity and dopant cross-section has not yet to be considered.

In a 3-HJ PNPN TFET, dopant fluctuation must be a cause for device variation. Both the thin 3.6 nm GaAsSb source well and the 2.4 nm InAs channel well require high doping. Fluctuation of doping at tunneling junctions would affect threshold voltage, junction field, tunneling current, and electrostatics. Therefore, a thick regrown highly-doped source (which is similar to the current state-of-the-art finfet/GAA FET process) will be a solution in the future. At the same time, the heavily doped channel is inevitably a trade-off between the loss in electrostatics by a thick body and doping fluctuation. From the process point of view, realizing an ultra-thin body could save TFETs from both, which is also the path silicon CMOS goes.

In addition, trap-assisted tunneling is another well-known issue that limits the performances of heterojunction TFETs [38, 37]. As has been discussed before, TAT contributes to not only off-state current but also subthreshold leakage. The interaction between electrons and traps depends on the energy of the trap state and trap capture cross-section. A study on the parameters of traps at different heterojunction interfaces is needed to include traps as a consideration in TFET design.

On top of that, the formation of defects/traps strongly depends on the process of material growth. A high amount of defects/traps and/or strain relaxation must result from growing mismatched films that exceed Matthew Blakeslee's limit. In most of the design literature of 3-HJ TFETs, strained material is usually employed for a preferred band offset and a superior material property. However, lots of efforts on avoiding strain material for minimizing defect formation are made in this study. The finalized InGaAs/GaAsSb/InAs/InP design uses all lattice-matched materials but a 2.4 nm strained InAs layer. $\approx$ 2 nm of strained InAs given a $\approx 3\%$ mismatch between InAs and InP is roughly at the Matthew Blakeslee limit.

3.5.3 Gate Misalignment in 3-HJ TFETs

As discussed in 2.4.2, a roughly four times decrease in on-current is a result of gate misalignment in a GaAsSb/InAs TFET. In 3-HJ TFET, I_{on} is even more sensitive to gate alignment due to the presence of resonant bound states. The simulation works shown in this section are credited to the Purdue team– Pengyu Long, Jun Z. Huang, and Michael Povolotskyi. Fig. 3.17 shows the schematic TFET structure having gate overlap/underlap (on the source) and gate aligned. The misalignment is ranged from ± 4 nm. It is noted that because this study is a pioneer work, the structure used here is not the finalized structure shown in the previous section.

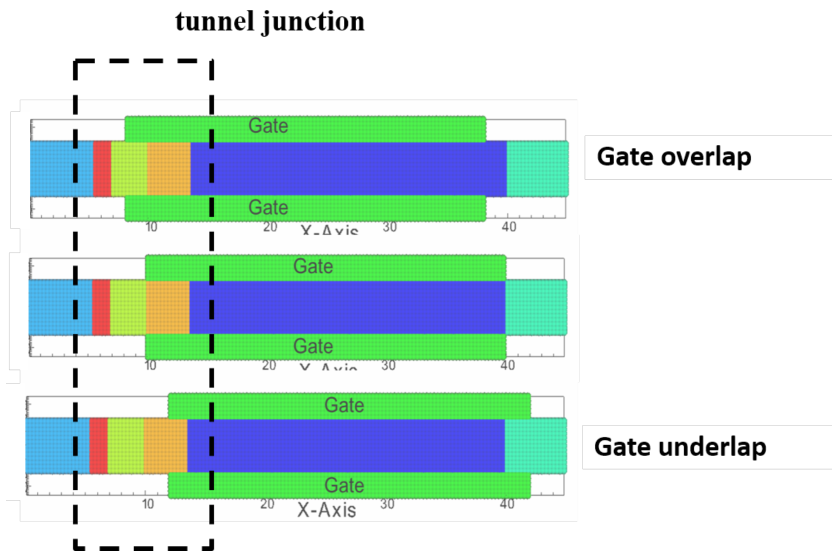


Figure 3.17: Schematic TFET structures for gate alignment analysis in 3-HJ TFETs.

Fig. 3.18 shows the band profile of 3-HJ TFET at tunneling junction for a (a) ± 2 nm, (b) − 4 nm (underlap), and (c) + 4 nm (overlap) gate alignment condition, and fig. 3.19 shows their resulted I_{DS} - V_{GS} characteristics. The red curves represent gate-aligned TFET. As can be seen from (a), the 2 nm underlap of the gate decreases the junction field between source/channel. Bound state energy in the channel increases accordingly.

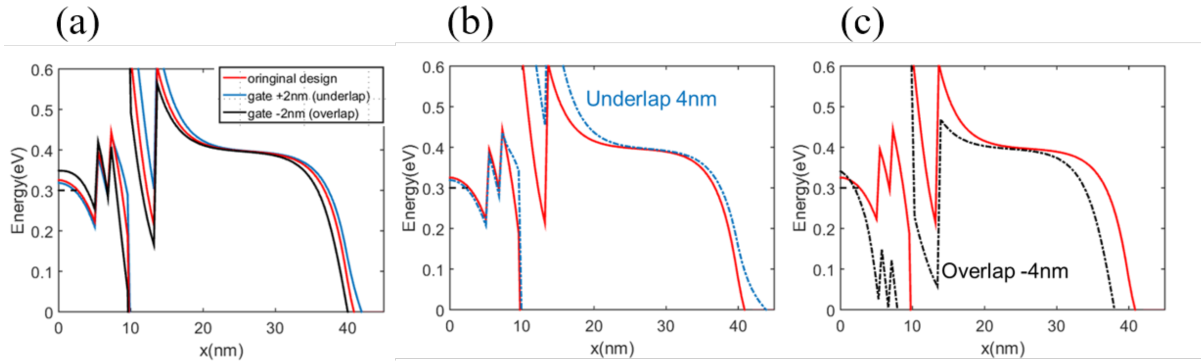


Figure 3.18: Energy band diagrams at on-state comparing gate aligned case with (a)overlap/underlap by 2 nm cases, (b) underlap by 4 nm cases, and (c) overlap by 4 nm cases. Courtesy of Purdue team.

In order to turn on the device, more gate bias is needed to push the channel potential down. Therefore, the threshold voltage is shifted positively. On the contrary, a 2 nm overlap of the gate also drops the junction field. The potential of the source well is pushed down simultaneously with channel potential, such that the bound state energy in the channel falls. A negative shift in threshold voltage is a result.

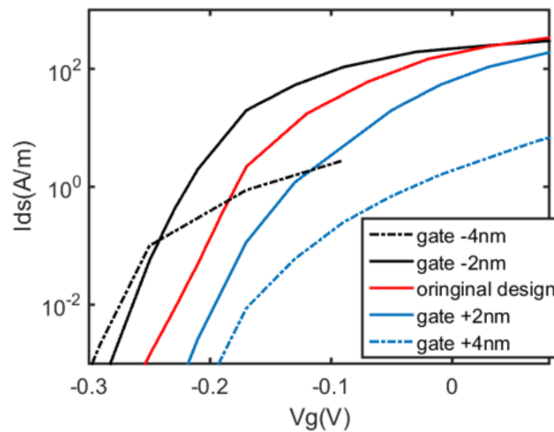


Figure 3.19: I_D - V_G characteristics of gate aligned, gate/source overlap by 2/4 nm (+2/+4 nm), gate/source underlap by 2/4 nm (-2/-4 nm) 3-HJ TFETs. Courtesy of Purdue team.

Similar to ± 2 nm, however, misaligning by 4 nm dramatically harms 3-HJ TFET performances. As shown in fig. 3.19, there is ≈ 2 orders of magnitude drops in I_{on} for

both underlap/overlap 4 nm cases. The catastrophic degradation results from the loss of gate control over bound state energy. Gate overlapping 4 nm suggests the surface potential of the source barrier moves with gate bias. The higher gate bias is, the more downward potential energy in source barrier/source well/channel well/channel barrier moves. As a result, bound state energy in source well and channel well will never align.

In general, all TFETs could not afford gate misalignment at the source/channel junction. Specifically, extremely high sensitivity to gate alignment of 3-HJ TFETs is paying for its high on-current, making it very promising and hard to achieve. The alignment tolerance of 3-HJ TFET is much smaller than the quantum well width, which is in the order of few nanometers. In chapter 5, a self-aligned gate process will be proposed to avoid misalignment. Detailed explanations on the mechanism/method as well as its pros and cons will be discussed.

In conclusion, heterojunction and doping profile engineering are introduced to improve TFET performance. Materials design, layer thickness, and doping concentration optimization for high-performance 3-HJ TFETs are discussed in detail. A realistic high on-current InGaAs/GaAsSb/InAs/InP TFET with a PNPN doping profile is proposed and optimized through atomistic modeling. At the same time, the potential issues of this design are also discussed.

Chapter 4

Development of Conformal ALD TiN/Ru Gate Metallization

4.1 Overview

The device structures proceeding to 3-HJ TFETs in this study, as will be discussed in chapter 5, includes both planar and vertical three-dimensional (3-D) design. Previous results have generally used gate metals deposited by sputtering or thermal evaporation. For nonplanar finFETs and next-generation gate-all-around (GAA) nanowire transistors, a more uniform and conformal high-k metal gate (HKMG) is necessary to ensure a low gate resistance and a consistent threshold voltage distribution [40, 67, 68].

In addition, a good gate control with high-quality high-k/channel interfaces has been shown to be extremely important in TFETs in chapter 2 and 3 for low SS and high I_{on}. The finalized InGaAs/GaAsSb/InAs/InP TFET design employes InAs and InP in channel. MOSFETs with SS approaching 60 mV/decade had been demonstrated with ZrO_2 and HfO_2 gate dielectrics on InGaAs and InAs channels at UCSB, indicating

dielectric/semiconductor interface trap densities D_{it} less than $1 \times 10^{12} cm^{-2} eV^{-1}$ [56, 57, 58]. InP channel, on the other hand, requires qualification and development of adequate surface cleaning procedures.

This chapter is divided into two parts– study on different gate metallization processes (section 4.2) and qualification of ZrO_2/InP interface as well as the development of gate metallization processes (section 4.3). Physical vapor deposition (PVD) and atomic layer deposition (ALD) metals are examined. Electrical properties, including threshold voltage and the quality of metal gate/high-k/semiconductor interface, are evaluated by capacitance-voltage (C-V) characteristics of a MOS capacitors (MOSCAPs) structure. InGaAs channel is used because of its well-developed surface cleaning procedure and proved qualitative interface with high-k. ALD TiN/Ru gate metallization process is developed. TiN/Ru gate shows decent conformity, good conductivity, suitable threshold voltage, and low D_{it}. This promising gate metallization is employed on InP channel planar MOSFETs, and the results are compared with traditional thermal evaporated Ni/Au gate and ALD Ru gate MOSFETs in section 4.3. Finally, an improved ALD Ru conformity by a lower temperature growth will be discussed in section 1.4.

4.2 MOSCAPs with Different Gate Metallization

4.2.1 Physical Vapor Deposition and Atomic Layer Deposition

PVD is commonly used to deposit thin films by a process in which materials go from a condensed phase (source/target) to a vapor/ionized phase, then back to a condensed phase (thin film) through thermal energy, ion or electron beam bombardment in a vacuum system [69]. Because of its physical mechanism and the vacuum environment, PVD coating generates condensed thin films with low conformity. If a higher conformal 3D

coating is needed, substrate rotation and low angle injection must be required. This method is not compatible with the patterning technique using photoresist lift-off. Therefore, patterning the as-deposited gate through etching is require for a higher conformal PVD gate metallization process.

Different from PVD, CVD features native higher conformity because of its chemical mechanism. CVD produces the deposition by exposing wafers to volatile precursors, which decompose and react at sample surfaces at a specific temperature depending on the reaction path. Among these, atomic layer deposition (ALD) is a popular technique in nanometer-scaled microelectronic devices due to its higher conformity, control over thickness and composition, and generally lower growth temperatures comparing to the CVD process [70]. The cyclic exposure of alternating precursors in sequence results in self-limiting processing, which gives atomic-scale precision. Materials with effective reaction pathways and suitable precursors, including high-k oxides, nitrides, sulfides, and elemental metal, can be grown by ALD [71]. Because of its moderately high-temperature growth window ($\approx$ 200 - 600 °C), the photoresist lift-off process is thus not compatible with the ALD metal gate, either.

Traditional thermal evaporated Ni/Au gate has been proved to yield a high-quality high-k/semiconductor interface because of its low damage [58]. However, Ni/Au gates could only be employed/patterned using the lift-off technique because they are the least reactive chemical elements. In other words, Ni/Au could not etch in most chemical solutions/plasma that has etching selectivity over other materials, which makes patterning challenges. A top-down etching process that blanket deposits metals following by patterning using dry/wet etching is thus not compatible with Ni/Au gate.

In addition, Au would diffuse into Ni and high-k at a moderately high temperature ($>$ 400 °C) which destroys transistors. This thermal budget limits the compatibility between different process modules. In modern transistor fabrications/VLSI manufacturing, a

refractory metal gate that could endure a high-temperature process is better appreciated. It is noted that a refractory gate will be a plus, but not the limiting factor in III-V transistors because most III-V materials (such as InAs, InGaAs, InP) melt at $\approx$ 500 °C, thus is the limiting factor of themselves.

According to the discussion above, a PVD/ALD metal gate with the feasibility of etching is required in non-planar device structures. PVD gates, including thermally evaporated Al, sputter Ti, and W, and ALD gates, including Ru, and TiN, are first investigated. A thermally evaporated Ni/Au gate is kept as a controlled sample.

4.2.2 ALD In-situ H_2 Annealing

The investigation is done using the MOSCAP structure. The fabrication of MOSCAP starts from metal-organic chemical vapor deposition (MOCVD) growth of 200 nm U.I.D. $In_{0.53}Ga_{0.47}As$ on N^+ (100) InP substrate. Solvent clean and one cycle of digital etch (UV ozone for 10 minutes and HCl: DI=1:10 for 1 minute) is done, followed by BHF dip for 2 minutes to remove surface oxide before loading into ALD system. It is noted that InGaAs surface must be hydrophobic in DI after BHF passivation such that water does not attach and further oxidize the surface in a high-temperature ALD chamber. High-k deposition includes nine cycles of AlO_xN_y initiation layer ($\approx$ 1 nm) that help the nucleation of the following high-k growth and 40 cycles of ZrO_2 layer ($\approx$ 3 nm). After that, samples are annealed in forming gas (5 % H_2/95 % N_2) at 400°C for 15 minutes to passivate surface dangling bonds. Large capacitive ground contact is used instead of ohmic contact to simplify the process. Such a high capacitance (that is generated by the large area ground) is negligible because it's in parallel with a small capacitance ($C_{ox}//C_s$) that is of interest. Gate and large-area capacitive ground contacts are formed in one go. The controlled thermal evaporated Ni/Au gate is deposited and lifted off while others

are deposited then dry/wet etched.

Fig. 4.1 shows the C-V characteristics of InGaAs channel thermal evaporated Ni/Au gate MOSCAPs with (a) as-deposited, (b) 3 minutes of ALD post gate H_2 annealing, (c) 9 minutes of ALD post gate H_2 annealing. The measured frequency is 1k, 10k, 100k, and 1M Hz. The large increase in capacitance in accumulation at low frequencies is a measurement artifact due to gate leakage, arising when G/ω becomes comparable to the capacitance. The as-deposited Ni/Au gate shows a high capacitance bump ($\approx$ -0.75 V) and a higher frequency dispersion at the depletion region ($\approx$ 0.25 V). They result from interface traps states near to mid-gap, and band edge of InGaAs, respectively. The frequency dispersion as in (a) is much greater than the old results published by UCSB [57], suggesting that a high amount of D_{it} is generated in the process. After carefully debugging every process step, it is concluded that this D_{it} is created during Ni deposition, indicating UV radiation could damage high-k during the metallization process.

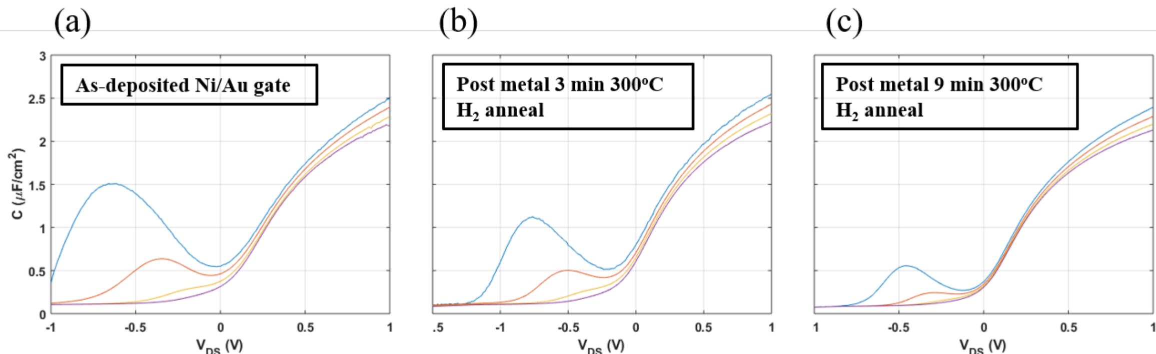


Figure 4.1: C-V characteristics of InGaAs channel MOSCAPs using thermal evaporated Ni/Au gate with (a) as-deposited, (b) 3 minutes of ALD post gate H_2 annealing, (c) 9 minutes of ALD post gate H_2 annealing at 300°C.

Post-metal annealing is tried to recover the damage. Fig. 4.1 (b) and (c) shows the reduced mid-gap and band edge D_{it} after ALD H_2 annealing at 300°C. It is noted that a longer annealing time above 9 minutes does not make more difference. This dispersion is comparable with that in [57], suggesting the damage to the high-k/channel interface

that is generated from Ni deposition is fully recovered. The annealing also causes a small negative threshold voltage shift.

The mechanism of ALD H_2 annealing and forming gas annealing (FGA) is the same– H_2 passivation of dangling bonds. Meanwhile, H_2 post-metal annealing is done at 300°C, which is lower than FGA at 400°C. It is possible that a higher amount of H_2 gas could enable complete passivation at a lower temperature. Therefore, H_2 anneal at 300 - 350°C is tried to replace FGA.

Fig. 4.2 shows the C-V characteristics of InGaAs channel thermal evaporated Ni/Au gate MOSCAPs with (a) 15 minutes of ALD post gate H_2 annealing at 300°C, (b) 15 minutes of ALD post gate H_2 annealing at 325°C, (c) 30 minutes of ALD post gate H_2 annealing at 350°C. It is noted that no FGA is performed after high-k deposition. It is clear that a large amount of D_{it} still presences in HKMG with H_2 annealing at 300°C for 15 minutes, while a $>$ 325°C annealing for more than 15 minutes passivates most of dangling bonds. By comparing these results to fig. 4.1 (c), it is concluded that a lower temperature H_2 annealing is effective to replace FGA at 400°C. Starting from 4.2.3, the standard FGA after high-k deposition is replaced by 30 minutes of ALD H_2 annealing at 350°C (that is done after high-k deposition), and an extra 30 minutes post gate metal H_2 annealing at 350°C (that is done after gate metallization) to recover the damage is set as a standard process for any gate process.

4.2.3 C-V characteristics of Different Gate Metallization and Metal Comformity

Initially, the sputter W gate is of particular interest because of its refractory property and well-developed and highly selective dry etching over high-k and III-V materials using SF_6/Ar plasma. In order to minimize the ion bombardment damage to the high-

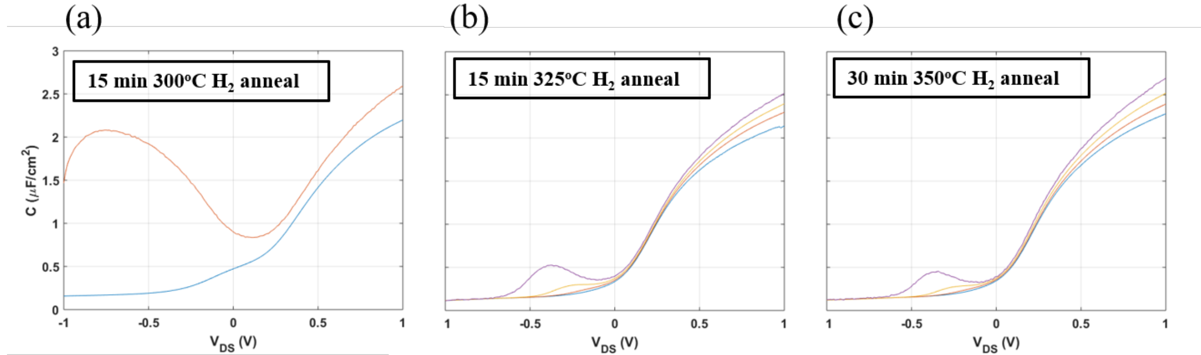


Figure 4.2: C-V characteristics of InGaAs channel MOSCAPs using thermal evaporated Ni/Au gate with (a) 15 minutes of ALD post gate H_2 annealing at 300°C, (b) 15 minutes of ALD post gate H_2 annealing at 325°C, (c) 30 minutes of ALD post gate H_2 annealing at 350°C.

k/InGaAs interface during sputter W deposition, low power of 100 watts is used. 40 nm of W film is sputtered with a growth pressure of 3 mtorr, and no substrate bias nor heat is added. After W sputtering, the sample is patterned, and W is dry-etched by inductive plasma etching (ICP) with source power of 600 watts, a bias power of 200 watts, and etch pressure of 1 Pa. To be noted, post-metal H_2 annealing is done for all W gates shown/compared in this section.

Fig. 4.3 shows the C-V characteristics of InGaAs channel MOSCAP using sputter W gate. As can be seen, the capacitance goes high at negative bias. It is observed from the red curve (measured at 100k Hz) that two different mechanisms contribute to the high capacitance. Analyzing by measuring over MOSCAPs with different area/perimeter, it is found that the high capacitance per unit area at $\approx$ -1 V is proportional to the perimeter of MOSCAPs, meaning it is generated during W dry etch. In comparison, the capacitance bump per unit area at $\approx$ -0.4 V is proportional to the area of MOSCAPs, indicating it is related to sputter W growth.

To reduce the damage on high-k by dry etching and sputtering, sputter W pressure is increased from 3 mtorr to 15 mtorr, while ICP bias power is reduced from 200 watts

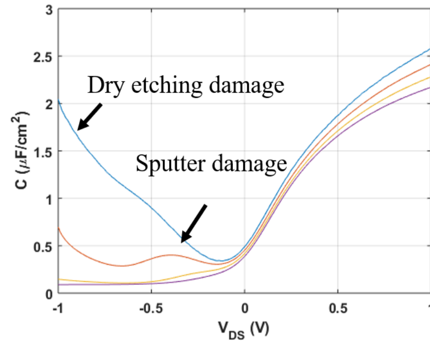


Figure 4.3: C-V characteristics of InGaAs channel MOSCAP using sputter W gate. Sputter W: power: 100 W, pressure: 3 mtorr. ICP W etch: SF_6/Ar: 20/5 sccm, power: 600/200 W, pressure: 1 Pa.

to 50 watts. However, the increase in growth pressure results in poor film conductivity. As a result, a 20 nm high-pressure W film that protects the high-k/InGaAs interface and a 20 nm low-pressure W film that serves as the conductive gate is grown in gate stacks. Fig. 4.4 (a) shows the resulted C-V characteristics. It is noted that the results are shown in fig. 4.3 and 4.4 are MOSCAPs with the same size such that the comparison in capacitance is not affected by MOSCAP area or perimeter.

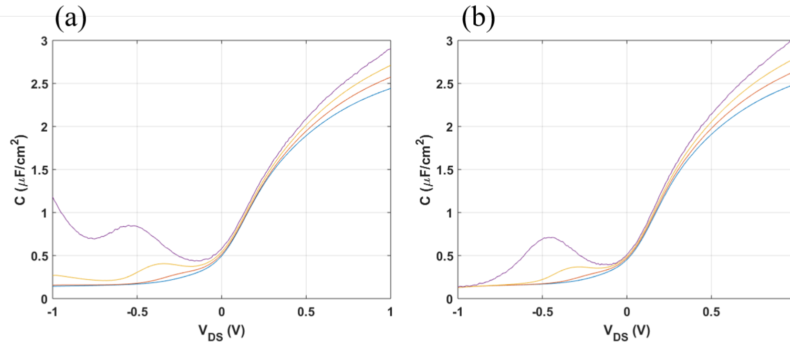


Figure 4.4: C-V characteristics of InGaAs channel MOSCAP using sputter W gate. (a) W stacks: 20/20nm with pressure: 15/3 mtorr. ICP W etch: SF_6/Ar: 20/5 sccm, power: 600/50 W, pressure: 1 Pa. (b)W stacks: 40/20nm with pressure: 15/3 mtorr. ICP W etch: SF_6/Ar: 20/5 sccm, power: 600/50 W, pressure: 5 Pa.

It is clear that a high-pressure growth of W and a low power etching both improve the D_{it} degradation. The further reduction in etching power results in a huge drop in

etch rate, owing to redeposition. Therefore, instead of reducing power, etch pressure increases from 1 Pa to 5 Pa to minimize dry etching plasma damage. In the meanwhile, the high-pressure W film thickness is increased from 20 nm to 40 nm. The resulted C-V characteristics show the damage to the high-k/InGaAs interface is minimized, and the mid-gap D_{it} of this optimized W gate is comparable to controlled Ni/Au gat, as shown in fig. 4.4 (b).

The conformity of sputter W is evaluated by sputtering on a test structure with a vertical sidewall and undercut feature and scanning electron microscope (SEM) imaging. Fig. 4.5 shows the cross-sectional SEM images of the test structures that have blanket W sputtered on at (a) a low pressure of 3 mtorr and (b) a high pressure of 15 mtorr. The upper images show the whole field, and the lower images focus on the deposition on sidewalls. Red arrows indicate sputter W films, which are bright under SEM due to a higher conductivity than substrates. The growth of W is perpendicular to the ground plane (as indicated by white arrows). This generates interfaces between W on vertical sidewalls and planar deposited W films. Such interfaces merge when W is getting thick for a low-pressure deposition condition (a). On the contrary, high-pressure sputtering results in a columnar structure as in (b). Those boundaries are formed initially during the nucleation stage and do not merge as the film is grown thicker. This phenomenon agrees with the C-V results shown before– the high-k/InGaAs interface could be protected from sputter damage (that is generated from low-pressure growth) by a thick enough high-pressure deposited W film.

In addition, there is no apparent difference between high/low-pressure growth on the conformity of W films. The growth rate on sidewalls is roughly 1/3 to 1/2 to the growth rate in-plane (depending on substrate height and target tilt). Undercuts in structure generate voids during sputtering, which would merge as W getting thicker, but still might be an issue when employed in different devices. Conformity test of W gate on

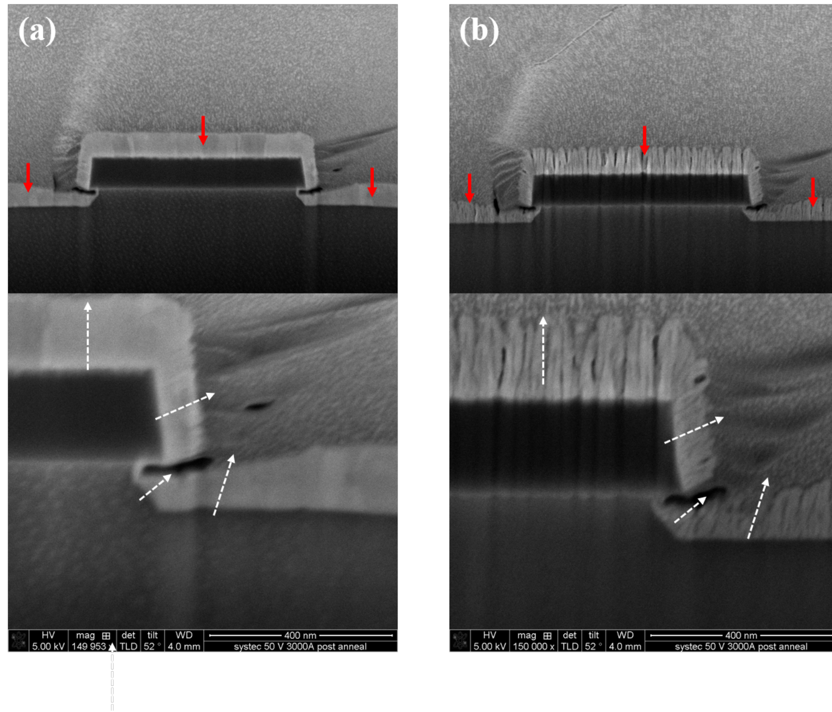


Figure 4.5: Cross-sectional SEM images of the test structures that have blanket W sputtered at a pressure of (a) 3 mtorr and (b) 15 mtorr. Red arrows indicate sputtered W layers.

device structure will be discussed later in the section.

Table 4.1 summarizes the W film properties versus different growth parameters, and table 4.2 shows the effect of high pressure initiation layer thickness on MOSCAP D_{it} (in form of C_{it}). At a low growth pressure, the film is smooth and conductive. At a high growth pressure, the film is columnar, and the sheet resistance (given by four-point measurement) is > 50 times greater than that of the film deposited at 3 mtorr. In contrast, a high-pressure growth of W minimizes the damage to the high-k/semiconductor interface. Concluded from table 4.2, W gate stack with > 5 nm high pressure (15 mtorr) initiation layer following a low pressure (3 mtorr) conductive layer is the optimized design.

In addition, MOSCAPs using thermal evaporated Al gate and sputter Ti gate show a high amount of leakage current, suggesting metal diffusion into high-k. ALD TiN gates, on the other hand, show decent D_{it} after post-metal H_2 annealing, as can been seen in fig.

Pressure\Height	2.5 (Wafer to Target Closer)	1.5 (Wafer to Target Farer)
3 mtorr	Film: smooth (conductive) Better conformal Midgap C_{it}~1.1* C_{ox}	Film: smooth (conductive) Less conformal
7 mtorr	Film: smooth Better conformal Midgap C_{it}~0.68* C_{ox}	–
15 mtorr	Film: columnar (resistive) Better conformal Midgap C_{it}~0.28* C_{ox}	Film: columnar (resistive) Less conformal Midgap C_{it}~0.28* C_{ox}

Table 4.1: The properties of sputter W gates under different growth pressure and substrate height.

W Stacks Design	D_{it} (Midgap Cit)
15mT initiation 1nm 3mT 75nm	~1.1* C_{ox}
15mT initiation 5nm 3mT 75nm	~0.28* C_{ox}
15mT initiation 20nm 3mT 75nm	~0.28* C_{ox}

Table 4.2: The amount of D_{it} versus sputter W stack design that has 1 nm, 5 nm, and 20 nm of high pressure initiation layer.

4.6. In TiN gate MOSCAPs, 25 nm of TiN is deposited by ALD using TDMAT with an N_2- and H_2-plasma (400-watt ICP power) at 300°C. A thick thermal evaporated Au layer is patterned and lifted off to ensure a uniform potential distribution over the resistive TiN layer. TiN is dry-etched by ICP with Cl_2/Ar = 20/5 scmm at a source power of 500 watts, bias power of 25 watts, and pressure of 1 Pa. Plasma damage is reduced because it is remotely generated in ALD TiN deposition. Nevertheless, the high-k/InGaAs interface is still damaged comparing (a) with the controlled thermal evaporated Ni/Au gate (as in fig. 4.2 (c)) By post-metal H_2 annealing, mid-gap D_{it} is successfully minimized, which makes TiN gate promising. However, the etch of 25 nm thick TiN requires both chemical reaction (Cl_2/BCl_3) and physical bombardment (Ar). This gas mixture has a much higher etch rate for III-V materials than TiN itself. As a result, TiN etching could not

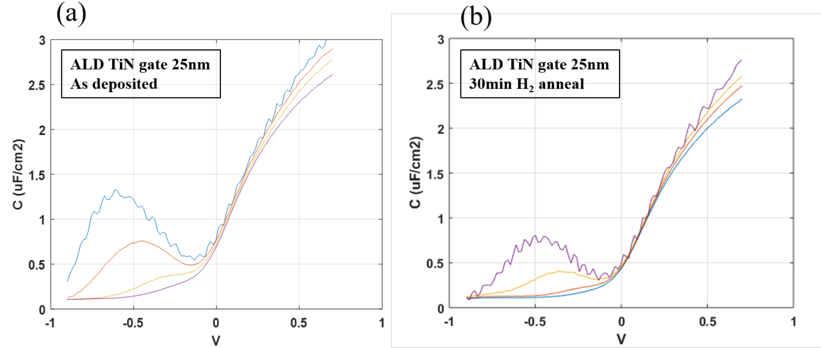


Figure 4.6: C-V characteristics of InGaAs channel MOSCAP using 25 nm TiN gates with (a) as-grown and (b) post-metal H_2 annealing at 300°C for 30 minutes.

be stopped and would go down into substrate in device fabrication.

In addition, ALD TiN has a significant sheet resistance of 188 Ω/□ for a 25 nm thick film. The resistive TiN film could be owing to the similar reason of the resistive sputter W film deposited at a high pressure, where the growth front of TiN is perpendicular to the formed grain boundaries, suggesting a bad conductivity along the wafer topography surface. The resistivity could be reduced by increasing source power [72]. However, higher power means more D_{it} is sacrificed. An alternative solution for a resistive TiN gate is a conductive metal layer on top of a thin TiN layer. The proposed TiN/Ru gate metallization will be studied and discussed in 4.2.4.

ALD Ru gate has different behaviors and usually shows significant variations over a small sample. In Ru MOSCAP fabrication, a 30 nm of Ru is deposited by ALD using (ethylbenzene)(1,3-cyclohexadiene)ruthenium (EBCHDRu) and O_2 cycles at 300°C. Ru film is then patterned and dry-etched by ICP with Cl_2/O_2 = 49.5/5.5 sccm at a source power of 500 watts, a bias power of 50 watts, and pressure of 2.5 Pa. Fig. 4.7 shows the C-V characteristics of InGaAs channel MOSCAP using 30 nm Ru gates at 1k (blue) and 1M Hz.

The mid-gap D_{it} of Ru gate, indicating by the capacitance bump at ≈ -0.25 V, is as low as in the controlled Ni/Au gate. This is reasonable because the ALD growth

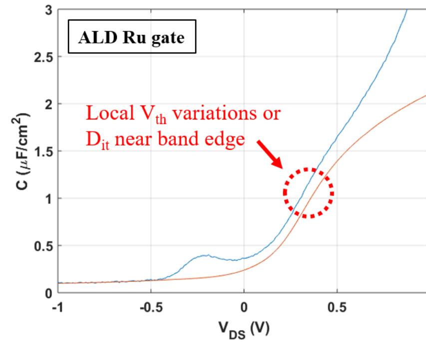


Figure 4.7: C-V characteristics of InGaAs channel MOSCAP using 30 nm Ru gate. Blue and red curves are measured at 1k and 1M Hz, respectively.

of Ru does not include any plasma process, meaning it does not cause any damage to the high-k/InGaAs interface in principle. However, there is frequency dispersion in C-V characteristics (as indicated by red arrows) at 0 - 0.5 V. This behavior in the Ru gate is different from what has been seen on other gate metallizations. In general, a dispersion at depletion region is mainly owing to a high amount of D_{it} near to band edge. This explanation might not be proper because the chemical reactions that take place during Ru deposition should not damage high-k, and Ru does not interact with or diffuse in ZrO_2 [73]. Another possible reason– a long ($\approx$ 3 hrs) and relatively high temperature (300°C) Ru deposition process is over the thermal budget of the high-k grown in-house, has been tested and proved it is not the case.

Local threshold voltage variation could explain this C-V dispersion. With threshold voltage variation, C-V curves are flattened, and the capacitance is averaged and spread out. Nucleation and growth initiation of ALD Ru appears to vary with substrates used. Rough and non-uniform growth on SiO_2, low-k dielectrics, and TaN surfaces have been reported [74, 75, 76]. Indeed, non-uniform substrate color after Ru deposition on ZrO_2 is sometimes observed at UCSB, suggesting a less than ideal Ru nucleation on ZrO_2. This bad nucleation could cause an insufficient surface coverage of Ru at Ru/ZrO_2 interface, which results in local threshold voltage variation [77]. No matter what exactly leads to the

	Ni (control)	Al	Ti	W	TiN	Ru
Technique	Thermal evaporator	Thermal evaporator	Sputter	Sputter	ALD	ALD
Preparation	Lift-off	Wet etch	Wet etch	Dry etch	Dry etch	Dry etch
V_{th} (V)	0.2	-1.3	-1.25	-0.2	-0.1	-0.2
C_{it} peak (μF/cm^2)	0.5	–	–	0.75	0.7	0.5
Application	Planar	–	–	Some non-planar	Non-planar	Non-planar
Note		leaky	leaky	columnar resistive film	resistive	Dit at band edge

Table 4.3: Summarized properties of different gate metallizations on ZrO_2/AlO_xN_y with InGaAs channel

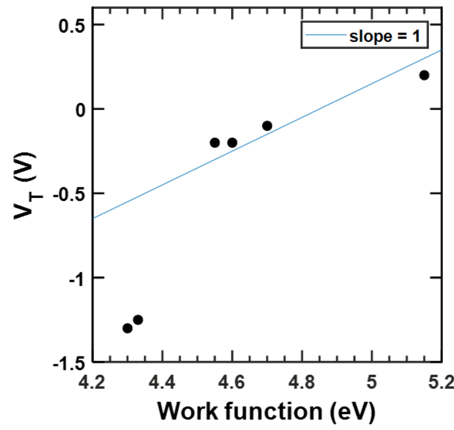


Figure 4.8: Threshold voltages of different gates vs. metal work functions.

dispersion, its effect on MOS transistors is similar to interface traps. A degradation in SS will result in MOSFET and TFET due to a less sharp subthreshold turn-on. Therefore, surface treatment or other technique to improve Ru nucleation is needed.

Table 4.3 summarizes different gate metallizations on ZrO_2/AlO_xN_y on InGaAs channel, and fig. 4.8 shows threshold voltage of InGaAs MOSCAPs versus metal work function. Al and Ti metals diffuse into high-k; thus, the leaky gates make the extrapolation of V_{th} inaccurate. Excluding Al and Ti, threshold voltage roughly increases with increased metal work function. A slope of one meaning there is no fermi-level pinning between the gate metal and high-k.

Concluding from the discussion above, thermally evaporated Ni gate and optimized

sputter W gate are promising for planar devices, while resistive ALD TiN gate and bad nucleated ALD Ru gate might be promising if the issues could be resolved. W gate could also be used in some non-planar structures that do not ask for excellent conformity. In the next section, an ALD TiN/Ru gate will be proposed and investigated

4.2.4 ALD TiN/Ru Gate

As discussed in 4.2.3, the drawback of the ALD TiN gate is a high sheet resistance, while the disadvantage of the ALD Ru gate is its bad nucleation on high-k that results in C-V dispersion at depletion regime. A more conductive TiN film could be achieved by using a higher source plasma power but sacrificing more significant damage on the high-k/channel interface. Similar to the development of sputter W, it is challenging to achieve high-quality film (which is related to its density, crystallinity, microstructure, and conductivity) and a low D_{it} at the same time. Therefore, Ru gate is focused, and methods to improve Ru nucleation are discussed in this section.

ALD Ru growth using metalorganic precursors with a metal valence of 2 or 3 shows a low resistivity, high step coverage (in high aspect-ratio structures), low impurity level (C and O), but a long incubation cycle which indicated a low growth rate at growth initialization stage [78]. An almost zero incubation time was achieved by NH_3 plasma-enhanced Ru growth, but a plasma process limits its application, such as MOS transistors [79]. Kim *et al.* reported a negligible incubation; thus, better nucleation and smooth surface using zero-valent Ru precursors [78].

At UCSB, a zero-valent Ru precursor (ethylbenzene)(1,3-cyclohexadiene)ruthenium (EBCHDRu, C14H18)(Hansol Chemical, Korea) and O_2 are used for ALD Ru growth. Ru precursor is dosed and held for 2.5 seconds at 1500 mtorr, while O_2 is dosed for 5 seconds at 500 mtorr. The growth rate is 0.6 A/cycle at 300°C.

In order to carefully study the generation of frequency dispersion in Ru gates, a different channel material (InP) is used in this section. Fig. 4.9 shows the C-V characteristics of InP channel MOSCAPs measured at 1k - 1M Hz using (a) thermal evaporated Ni/Au gate and (b) Ru gate. Comparing (a) with fig. 4.2, the controlled Ni/Au gate MOSCAP with InP channel shows no high capacitance D_{it} bump closed to mid-gap in C-V characteristics. This could be owing to a relatively greater bandgap of InP (1.4 eV) over InGaAs (0.75 eV) such that the characteristic frequency of interface trap states near to mid-gap is lower than the measured frequency [80]. In other words, electrons could not respond to the mid-gap traps in a high bandgap material within 1k - 1M Hz. The C-V dispersion in InP channel MOSCAPs lies mainly at the depletion region (≈ 0.3 - 0.6 V), as indicated by red arrows in (a). This suggests a higher D_{it} distributes nearly to band edge.

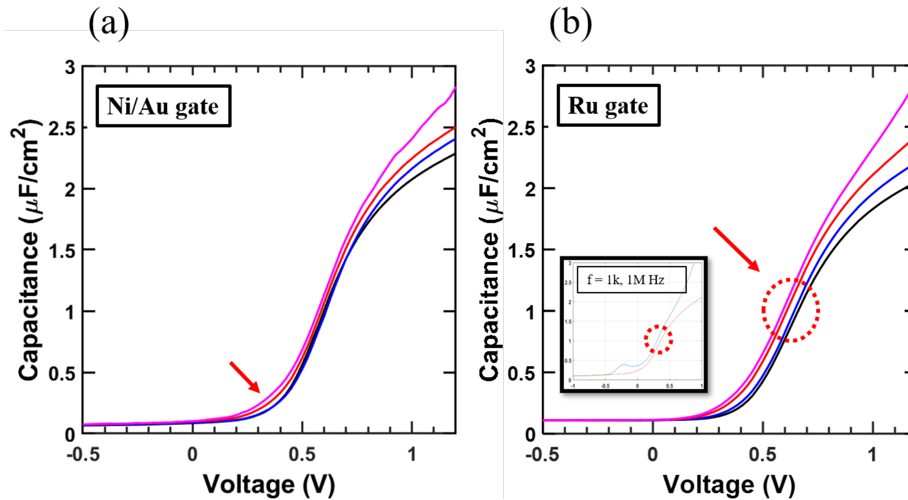


Figure 4.9: C-V characteristics of InP channel MOSCAPs measured at 1k - 1M Hz using (a) thermally evaporated Ni/Au gate and (b) Ru gate.

Again, Ru gate (as in (b)) shows a more significant frequency dispersion at depletion regime compared to Ni/Au gate. The inert figure in (b) is the C-V characteristics of Ru gate InGaAs channel MOSCAP. As specified by the red circles, similar dispersion at

depletion happens at both InGaAs and InP channels, meaning the cause is unrelated to channel material. Nevertheless, the dispersion is electrically equivalent to interface trap states near to band edge. Therefore, it is crucial to solving this dispersion issue for a low SS in Ru gates MOSFETs and TFETs.

As discussed in 4.2.3, local threshold voltage variation caused by a less-than-ideal Ru nucleation on ZrO_2 could explain this dispersion. Yim *et al.* reported ALD Ru with improved nucleation on a thin SiN_x deposited or heavily NH_3-plasma-activated (50-minute exposure) SiO_2 surfaces according to TEM analysis [74]. Moreover, M. Zhang *et al.* demonstrated improved ALD Ru nucleation with an Al_2O_3 surface layer over SiO_2 [81]. Heo *et al.* also showed an enhanced Ru nucleation behavior of ALD Ru film by UV-O_3 treatment on low-k dielectrics [75]. Overall, it was suggested that one could improve the nucleation of Ru growth by surface energy engineering, which helps the adsorption of Ru precursors [82]. However, inserting an Al_2O_3 interlayer and plasma treatment before Ru deposition in HKMG/channel interfaces may degrade devices' electrical performance due to increased effective-oxide-thickness (EOT), increased interface trap density, and a shift in the MOSFET/TFET threshold voltage. Therefore, it is important to consider those trade-offs in HKMG design and optimization.

In this study, a thin TiN layer ($\approx$ 2nm) is deposited as a nucleation/stiction layer in Ru gate devices [83]. The $\approx$ 2nm TiN is deposited by ALD using TDMAT, N_2, and H_2 plasma at 300°C. TDMAT is dosed and held for 0.25 and 0.75 seconds, respectively at 80 mtorr while the mixture of plasma is generated/held by $N_2/H_2 = 12/4$ sccm at a source power of 400 watts, the pressure of 2.5 mtorr for 20 seconds. Ru growth parameters are the same as before. The pattering of Ru and TiN is done separately. Ru is first patterned, and dry etch using ICP $Cl_2/O_2 = 49.5/5.5$ sccm at a power of 500/50 watt (source/bias) and a pressure of 2.5 Pa.

TiN etching could be done in two ways. Dry etching using $BCl_3/Cl_2/Ar$ plasma by

ICP system is adequate to etch TiN film. However, the physical bombardment using this gas mixture results in the etch of TiN and high-k and semiconductor. The etch selectivity between them is not high enough, such that the dry etch is hard to control and repeat. Therefore, a wet etch of TiN is investigated.

The TiN film deposited in-house by the ALD system is stable, hard, and refractory. Typically, a mixture of ammonium hydroxide, hydrogen peroxide, and DI water at above 50°C can be used to etch TiN. However, this etchant also attacks InGaAs with etch rate much faster than TiN, meaning liquid etchant would go through the pinholes of high-k and etches the InGaAs underneath in device fabrication.

A wet etch of thin TiN by BHF could be achieved once the TiN layer is oxidized. In other words, there is no etching of as-grown TiN film in BHF. To oxidize TiN thin film, 1 minute of oxygen plasma exposure is done in Technics PEII asher. Plasma power is set to 100 watts, and pressure is kept at 300 mtorr. After oxidation, 1 minute of BHF etch is done to clean up both TiN and high-k. It is noted that the resist/hard mask for Ru dry etch was kept until finishing the TiN etch process, such that Ru is fully protected from oxygen plasma and BHF.

Table. 4.4 summarizes the process flow for TiN/Ru gate formation and fig. 4.10 shows the C-V characteristics of InP channel MOSCAPs using (a) TiN (1 nm)/Ru gate and (b) TiN (2 nm)/Ru gate. As can be seen, the accumulation capacitance C_{acc} (at 1 V) in TiN (1 nm)/Ru gate is $\approx$ 2 $\mu F/cm^2$, which is smaller than $\approx$ 2.25 $\mu F/cm^2$ in TiN (2 nm)/Ru gate. This is a result of the non-negligible capacitance given by a thin resistive 1 nm TiN layer. The C_{acc} for a TiN (2 nm)/Ru gate is almost the same as in Ni/Au gate (after threshold shift), indicating the capacitance of a 2 nm TiN is negligible. Therefore, 2 nm is the minimum thickness of the TiN layer in TiN/Ru gate design, given its high resistivity.

After comparing the results in (b) with Ru gate as in fig. 4.9 (b), it is clear that

TiN/Ru Gate Formation:
ALD TiN/Ru deposition
In-situ post metal H_2 annealing
Resist/SiN mask patterning
Ru dry etch in Cl_2/O_2 plasma
Thin TiN oxidization by O_2 ashing
BHF dip to remove oxidized TiN & high-k
Resist stripping/hard mask removal

Table 4.4: Process flow for TiN/Ru gate formation

the frequency dispersion at the depletion region is tremendously reduced. In addition, uniform color on any arbitrary substrates is realized by depositing the thin TiN layer prior to Ru. Therefore, it is concluded that the nucleation of Ru is improved by the addition of a thin ALD TiN layer, and the corresponding C-V dispersion in the Ru gate is fully solved. In addition, the dispersion in TiN/Ru gate is comparable to that in Ni/Au gate (as in fig. 4.9 (a)), suggesting the damage to high-k/semiconductor interface generated from the thin TiN deposition could be fully recovered by post-metal H_2 annealing.

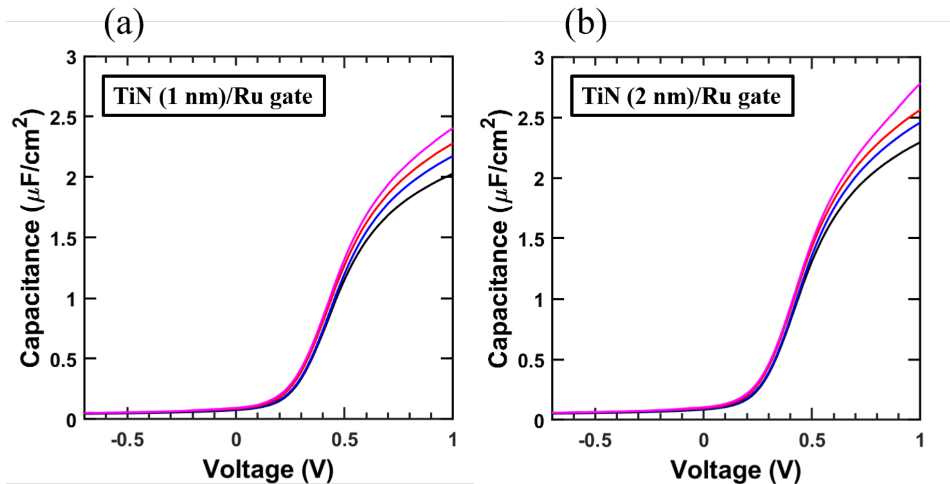


Figure 4.10: C-V characteristics of InP channel MOSCAPs using (a) TiN (1 nm)/Ru gate and (b) TiN (2 nm)/Ru gate.

Furthermore, a phenomenon is observed on InP channel MOSCAPs using any gate. Fig. 4.11 shows the (a) first sweep and (b) the second sweep of C-V characteristics of

InP channel MOSCAPs using TiN/Ru gate at 1k - 1M Hz. The order of measurement (all sweeping from -1 to 1 V) is: (a) 1M, 100k, 10k, 1k, (b) 1M, 100k, 10k, 1k. As can be seen, the first swept curve (blue) in (a) has a threshold voltage slightly smaller than other curves, while the blue curve in (b) has the same threshold voltage as in others. This indicates there might be a trapping/de-trapping action happening at the first sweep of MOSCAP. Once the trap state is ionized/deionized, the rest of the sweeps are not affected anymore. This phenomenon is not observed on the InGaAs channel but InP channel using any gate metallization. As a result, it is due to a defect state related to the high-k formation on the InP channel, specifically.

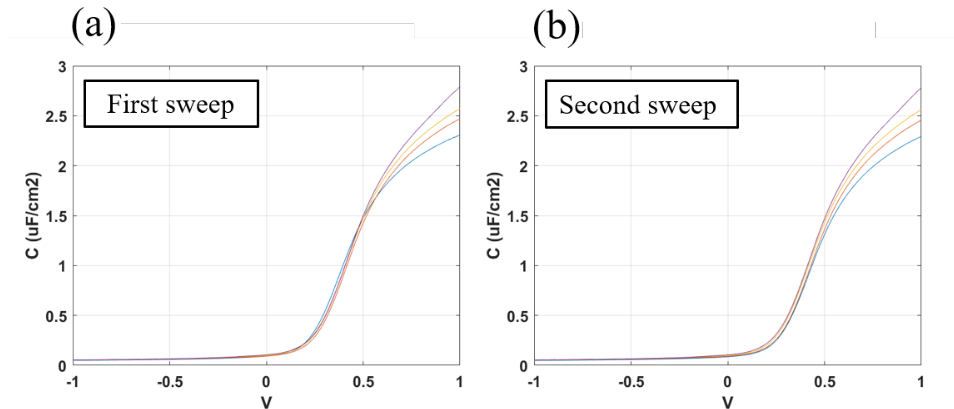


Figure 4.11: (a) First sweep and (b) second sweep of C-V characteristics of InP channel MOSCAPs using TiN/Ru gate.

In conclusion, a high-quality HKMG process using ALD TiN/Ru gate is developed, and the C-V characteristics suggest it is promising both on InGaAs and InP channels. Because of the chemical and layer-by-layer nature of ALD, the deposited TiN/Ru is assumed to be conformal. In the next section, TiN/Ru gate will be employed on a recessed gate planar MOSFET structure. The step coverage/conformity will be discussed based on TEM analysis.

4.3 InP Channel Planar MOSFETs with ALD TiN/Ru Gate

Low InP/dielectric interface trap density D_{it} will enable low SS in tunnel FETs (TFETs) using InAs/InP heterojunctions for increased tunneling probability [47, 84]. The quality of the HKMG/InP interface depends on the growth of high-k dielectrics and the InP surface clean before growth. 4.3.1 will introduce the surface cleaning of InP that is different from arsenide materials. 4.3.2 will show the process flow of recessed gate planar MOSFETs fabrication. The qualification of D_{it} at the high-k/InP interface is done by measuring SS over long gate length MOSFETs using controlled Ni/Au gates. InP channel planar MOSFET results comparing different gate metallization and channel doping will be discussed in 4.3.3 and 4.3.4.

4.3.1 InP Surface Cleaning Procedure

InP channel design is widely employed in electronic and photonic devices due to its moderate bandgap and lattice constant. To achieve a high-quality metal/InP (M/S) or metal/oxide/InP (MOS) interface, the removal of semiconductor surface oxide is fundamental. In general, most oxides dissolve in HF-based solution, suggesting that the use of HF in surface cleaning requires careful evaluation such that it does not attack other dielectrics on the substrates. This limits the incorporation of HF as a standard surface cleaning process. In III-V planar MOSFETs, fabrication is quite simple, and BHF dip is usually done before high-k deposition as a surface clean-up. However, HF is not compatible with the fabrication of a vertical MOSFET structure (which will be discussed in chapter 5) because it would attack the SiN_x dielectric sidewalls that surround the fin. The necessity of adding SiN_x sidewalls in the vertical FET process will be explained in

chapter 5. Therefore, another chemical treatment is needed, and HCl-based cleaning will be focused on in this study.

Effective surface oxide removal by HF, HCl, and H_2SO_4 solutions is seen on most of the arsenide materials, including InGaAs, InAs, and GaAs [85, 86]. An oxide-free and hydrophobic surface after treatment is a result of elemental As build-up at the surface [87]. The elemental As also protects the surface from the attachment of water molecule, which would further oxidize the InP surface in the high temperature ALD chamber. At UCSB, a standard 2 minutes BHF dip or one cycle of digital etch (UV ozone 10 minutes then dip in HCl/DI = 1:10 for 1 minute) is done right before loading samples into the ALD chamber. Those surface treatment/preparation gives a beautiful hydrophobic surface and is proved to be effective for high-quality high-k/InAs and InGaAs interfaces, given the almost ideal SS of 60 - 65 mV/dec in MOSFETs [58]. InP, on the other hand, behaves differently as tested in-house at UCSB.

A hydrophilic surface is observed after dipping InP into HF/BHF solution. Meanwhile, there is rare oxide left on the surface after HF/BHF dip (given by ellipsometry analysis), meaning InP native oxide is actually etched in HF/BHF. Sun *et al.* suggests surface In atoms were bonded to fluorine, and the high ionic energy in F-terminated surface does not prevent water from attaching onto the surface, thus generating a hydrophilic surface [87]. In addition, a standard digital etch (DE) (UV ozone 10 minutes then dip in HCl/DI = 1:10 for 1 minute) also gives a hydrophilic InP surface and the amount of native oxide left on the surface is not consistent. Table 4.5 shows a series of experiments testing the effect of surface treatments on native oxide removal and its surface property. The native oxide thickness is measured by ellipsometer using simple Cauchy -SiO_2 set-up in the fitting. It is for sure that the fitted absolute thickness is not accurate, but the relative number between samples could be used to understand the relations and effectiveness of the treatment.

Sample No.	Preparation	Surface property	Native oxide thickness
1	InP epi-ready wafer	Hydrophobic	1.09 nm
2	Sample 1 + InP etchant	Hydrophobic	0.64 nm
3	Sample 2 + Standard DE	Hydrophilic	0.55 nm
4	Sample 3 + Standard DE	Hydrophilic	0.62 nm
5	Sample 4 + Standard DE	Hydrophilic	0.61 nm
6	Sample 5 + 10 min UV ozone	Hydrophilic	1.47 nm
7	Samle 6 + 1 min HCl:DI=1:10	Hydrophilic	0.66 nm
8	Sample 7 + Standard DE	Hydrophilic	0.61 nm
9	Sample 8 in DI 5min	Hydrophilic	0.69 nm
10	Sample 9 in DI 5min	Hydrophilic	0.76 nm
11	Sample 10 in HCl:DI=1:10 3min	**Hydrophobic**	0.56 nm
12	Sample 11 + 10 min UV ozone + >4min HCl:DI=1:10	**Hydrophobic**	
13	Sample 12 in DI 1 min up	Start to attach some water	

Table 4.5: Series of experiments testing the effect of chemical surface treatments on native oxide removal and the surface properties after treatments.

As can be seen, an epi-ready InP substrate is hydrophobic and has 1 nm native oxide on it. After putting the substrate into InP etchant for 8 seconds (which etches $\approx$ 60 nm InP), the surface is still hydrophobic while the thickness of surface oxide is reduced to 0.6 nm. After that, one cycle of standard DE is done, which leaves a hydrophilic surface. If the sample is treated by more standard DE, the surface would always be hydrophilic, and the surface oxide thickness after each cycle remains the same ($\approx$ 0.6 nm).

In order to quantify the oxidation by UV ozone exposure, the sample is put in a UV ozone reactor for 10 minutes, and the resulted surface oxide is 1.4 nm. After that, a 1-minute dip in HCl: DI=1:10 is done and leaves a 0.6 nm oxide as expect. This suggests a $\approx$ 0.8 nm oxide is etched in dilute HCl solution. If putting the hydrophilic InP sample (that is ended with 0.6 nm thick oxide by a standard DE) under flowing DI, the thickness of surface oxide would increase with time, suggesting the InP surface is slowly oxidized. Ten minutes in DI results in a $\approx$ 0.15 nm increase in oxide thickness.

Surprisingly, a $>$ 4 minutes dip in dilute HCl after 10 minutes of UV ozone gives a hydrophobic surface. This longer dip in a dilute acid might serve the same as a short dip in a more concentrated acid as shown in [87], in which elemental P atoms at the surface are terminated by H and leaves a hydrophobic surface. It is noted that the dipping time to get a hydrophobic surface would depend on the intensity of UV light. There is no control over light intensity in the UV ozone reactor at UCSB. The increased intensity is observed as the lamp is turned on longer in time. Therefore, the minimum dipping time in dilute HCl is designed to be $>$ 3 minutes for surface oxide clean up and $>$ 5 minutes for a hydrophobic surface.

A hydrophobic surface prevents water from attaching to the surface, which would later become an oxygen source for surface oxidization in a higher temperature ALD chamber. Even without getting a hydrophobic surface, a $>$ 3 minutes dip in HCl: DI=1:10 is enough to clean up the surface oxide on the InP substrate. It is tested that the electrical performances (of MOSCAPs) are exactly the same in hydrophobic and hydrophilic InP channels, possibly thanks to the 9 cycles of nitrogen plasma/TMA growth initiation before ZrO_2 deposition.

In conclusion, a longer ($>$ 3 minutes) dip in dilute HCl is needed to clean up the surface oxide on the InP channel. A modified standard process for InP surface clean that contains 10 minutes of UV ozone exposure and $>$ 3 minutes dip in HCl: DI=1:10 is designed and employed in the fabrication of devices that will be discussed in the following sections/chapter.

4.3.2 Process Flow

Planar MOSFETs were fabricated on semi-insulating (100) Fe-doped InP substrates. The fabrication started with epitaxial growth of the channel by metal-organic chemical

vapor deposition (MOCVD). The MOCVD channel growth was performed at 600°C and consisted of a bottom 9 nm Zn-doped P-InP ($8 \times 10^{17} cm^{-3}$) layer to compensate for the donor impurities at the growth interface and a top 9 nm U.I.D. InP layer, as shown in fig. 4.12 (a).

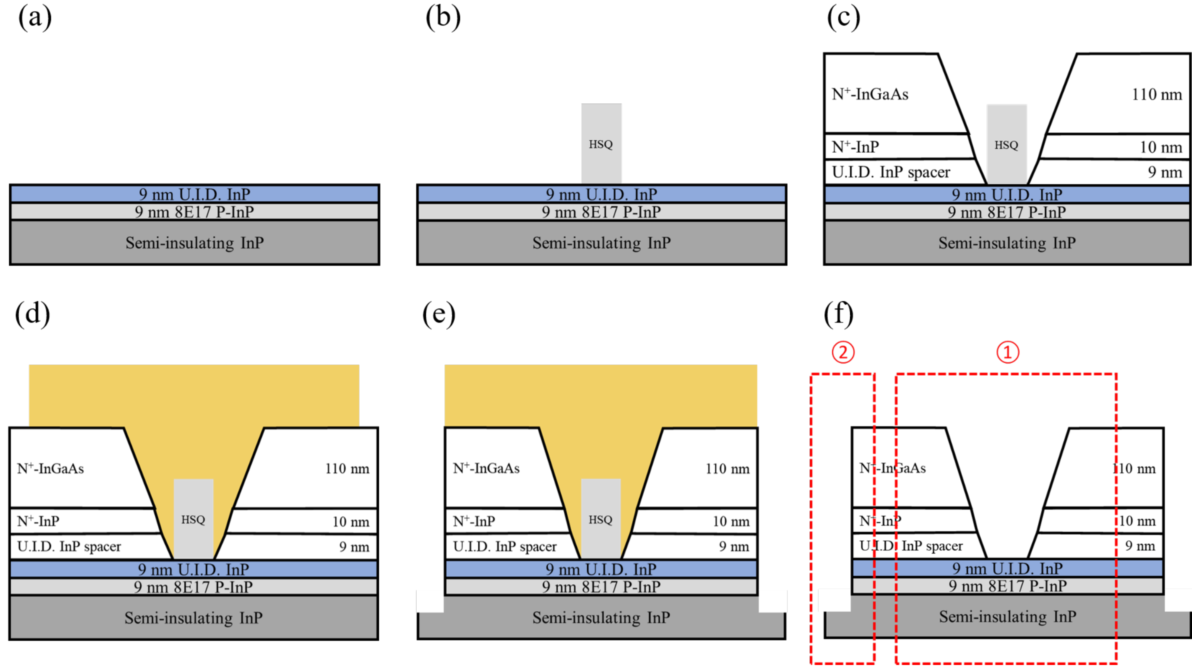


Figure 4.12: Planar MOSFET fabrication process before gate metallization

After channel growth, a dummy gate was defined by electron beam lithography (EBL) using hydrogen silsesquioxane (HSQ) resist for subsequent self-aligned raised source/drain MOCVD regrowth (fig. 4.12 (b)). A 1 minute dip in HCl:DI=1:10 was done prior to the regrowth. For the source/drain, a 9 nm U.I.D. InP spacer, 10 nm Si-doped N^+-InP ($2 \times 10^{19} cm^{-3}$), and 110 nm Si-doped ($4 \times 10^{19} cm^{-3}$) N^+-$In_{0.53}Ga_{0.47}As$ layers were grown at 600°C (fig. 4.12 (c)). Devices were then isolated by selective wet etch using H_3PO_4:H_2O_2:DI=1:1:25 and HCl: H_3PO_4=1:4 for InGaAs and InP, respectively (fig. 4.12 (e)). To be noted that fig. 4.13 shows the rest of process steps specifying MOSFET active region.

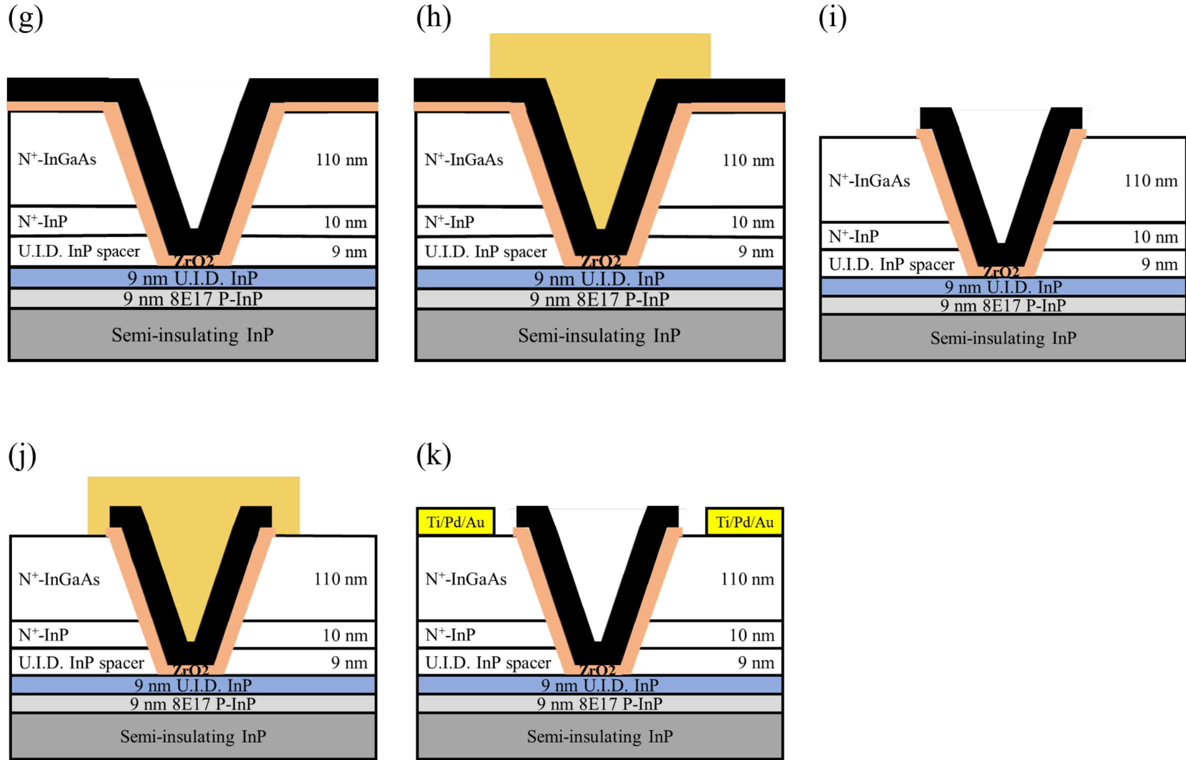


Figure 4.13: Planar MOSFET fabrication process after gate metallization

After a two-minute BHF dip followed by one cycle of an HCl-based digital etch (10 minutes of UV ozone and 3 minutes of HCl: DI=1:10) to remove the dummy gate and surface oxides, high-k and metal gate layers were deposited in an Oxford FlexAL ALD system. The high-k deposition process includes initial surface passivation using 9 cycles of alternating N_2-plasma and trimethylaluminum (TMAl) dosing ($\approx$1 nm AlO_xN_y) followed by 40 cycles of H_2O and tetrakis(ethylmethylamido)zirconium (TEMAZ) dosing at 300°C [56, 57]. In order to further passivate surface dangling bonds, a 30-minute ALD in-situ H_2 annealing was performed at 350°C . Metal gate deposition starts from 35 cycles of TiN nucleation layer ($\approx$2 nm) deposited using Tetrakis(dimethylamido)titanium (TDMAT) with an N_2- and H_2-plasma at 300°C (400-watt ICP power). 500 cycles of ALD Ru ($\approx$30 nm) was then deposited using (ethylbenzene)(1,3-cyclohexadiene)ruthenium (EBCHDRu), and O_2 cycles at 300°C. EBCHDRu is a zero-valent organometallic precur-

sor (Hansol Chemical, Korea). In-situ post-metal annealing in H_2 ambient at 350°C was employed for 30 minutes to recover plasma damage at the high-k/InP interface (fig. 4.13 (g)). The Ru metal gate patterns were dry-etched using an O_2-based ICP etch (catalyzed with a small amount of Cl_2) [88]. After oxidizing the thin TiN layer through 1-minute exposure in oxygen plasma, the TiN and the high-k dielectric layers were etched via a 1-minute BHF dip (fig. 4.13 (i)). Source and drain metal contacts (Ti/Pd/Au) were deposited by lift-off (fig. 4.13 (k)). Devices were finally passivated using Al_2O_3 (≈3 nm) deposited by ALD. It is noted that detailed processing parameters and recipes are summarized in the appendix.

Because Ru is conformally grown but anisotropically dry-etched, the TiN/Ru gate metal surrounding mesa edges remain after gate patterning, as shown in fig. 4.14 (b). This wrap-around gate on mesa edges creates gate leakage current (I_g). The amount of I_g generation depends on the perimeter of mesa edges. A higher normalized I_g will result if the ratio between mesa perimeter to gate width is high.

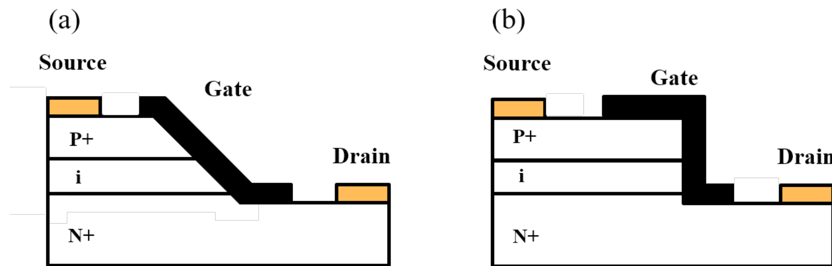


Figure 4.14: (a) The as-grown schematic diagram showing TiN/Ru gate converage on mesa edges. (b) The surrounding TiN/Ru gate metal (over mesa edges) remains after gate patterning.

4.3.3 InP MOSFETs Using Thermal Ni/Au, ALD TiN/Ru, ALD Ru Gates

Fig. 4.15 shows (a) SEM images of an L_g=30 nm planar MOSFET using thermal Ni/Au gate, and (b) the schematic top view MOSFET layout. The recess structure and U.I.D InP spacer are used to suppress band-to-band tunneling (BTBT) happened at the high-field region near the drain end of the channel [89]. Note that the effective gate length is the horizontal channel length plus twice the vertical spacer thickness.

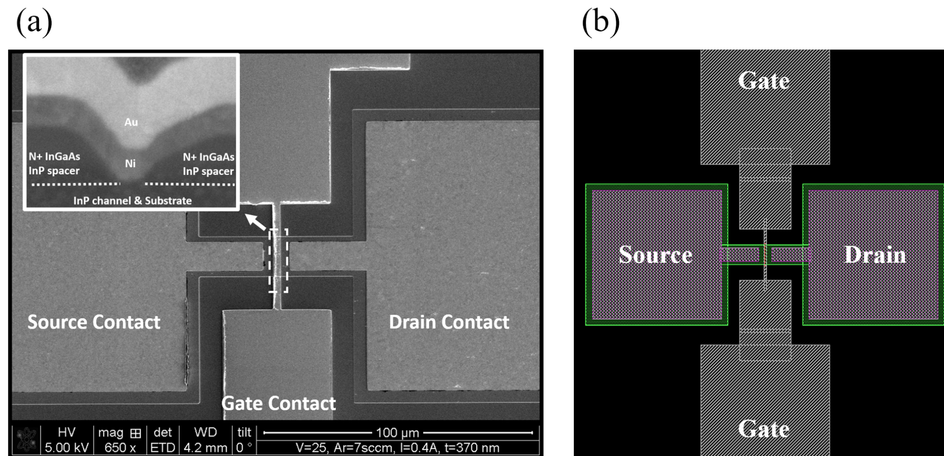


Figure 4.15: (a) SEM images of an L_g=30 nm planar MOSFET using thermal Ni/Au gate and © 2019 IEEE (b) schematic top view MOSFET layout.

The gate length of fabricated planar MOSFETs ranges from 30 nm to 2 μm. Fig. 4.16 shows the transfer characteristics of 40 nm, 100 nm, and 1 μm-L_g FETs using thermal evaporated Ni/Au gates. The on/off ratio is $>10^5$ for L_g > 50 nm. Due to the 14 nm thick channel, the output characteristics, fig. 4.17 (a), show short-channel effects at 40 nm L_g. I_{on} is 0.09, 0.1, 0.05 mA/μm for 40 nm, 100 nm, 1 μm-L_g, respectively at V_{DS} = 0.6 V and V_{GS}-V_{TH} =0.5 V. 100 nm-L_g devices exhibit 0.31 mS/μm peak g_m at V_{DS} =0.8 V. The N^+ S/D layer sheet resistance is 14 Ω, while the S/D contact resistivity is 7 Ω-μm^2, these determined from TLM measurements.

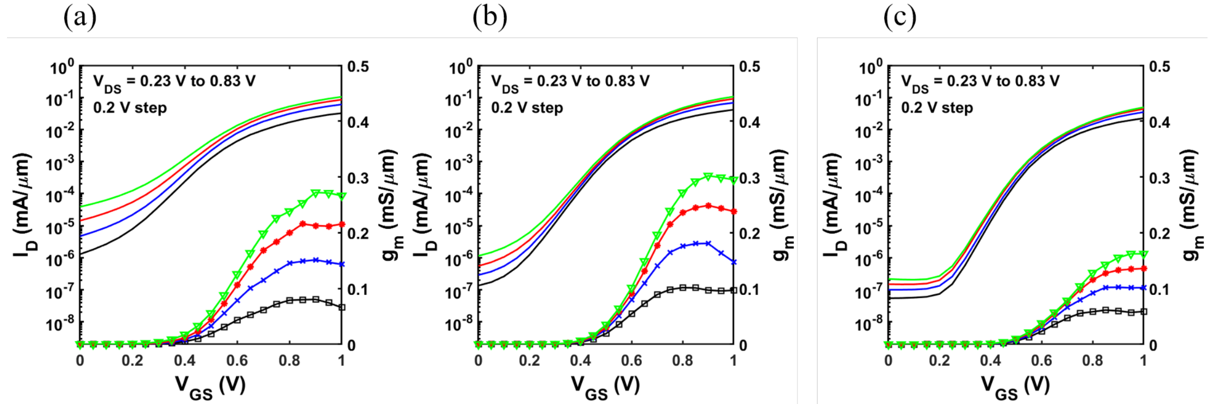


Figure 4.16: Transfer charateristics of 40 nm, 100 nm, and 1 μm-L_g MOSFETs using Ni/Au gates. © 2019 IEEE

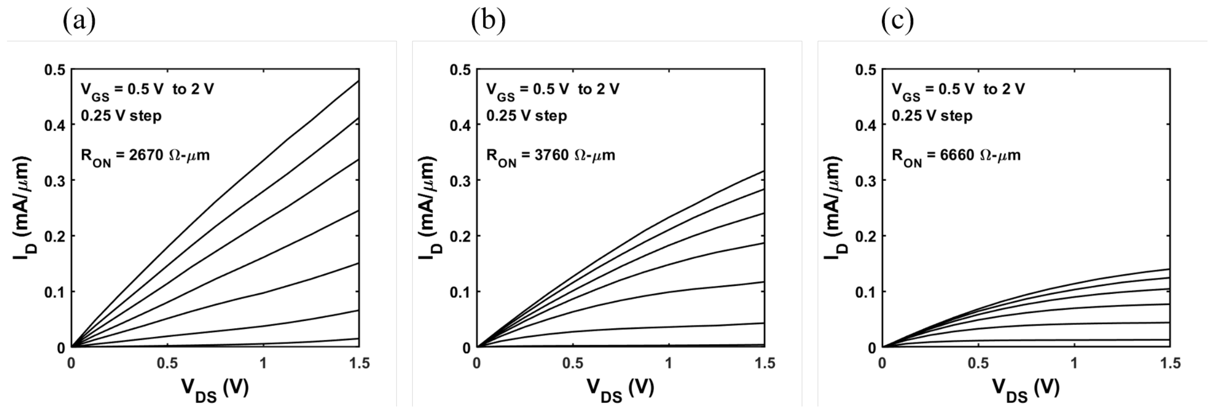


Figure 4.17: Output charateristics of 40 nm, 100 nm, and 1 μm-L_g MOSFETs using Ni/Au gates. © 2019 IEEE

Fig. 4.18 shows (a) SS vs. L_g at V_{DS} =0.1 and 0.6 V; (b) peak g_m vs. L_g at V_{DS} =0.6 V; (c) DIBL vs. L_g at V_{TH} =1 $\mu A/\mu m$. The low SS of 70±3 mV/dec at V_{DS} =0.1 V is measured over 13 different 2 μm-L_g devices. D_{it} calculated from this SS is ≈ 3× 10^{12} $cm^{-2}eV^{-1}$. Peak g_m reaches ≈0.25 mS/μm at L_g <100 nm. The maximum g_m happens at L_g in between 60 nm to 100 nm, which is ≈4 times of channel thickness of 14 nm. Further reducing gate legnth does not enhance g_m due to a relative thick channel.

As a controlled gate metallization technique, MOSFETs using thermal evaporated Ni/Au gate prove a reasonably low D_{it} at ZrO_2/InP interface. It is concluded that InP

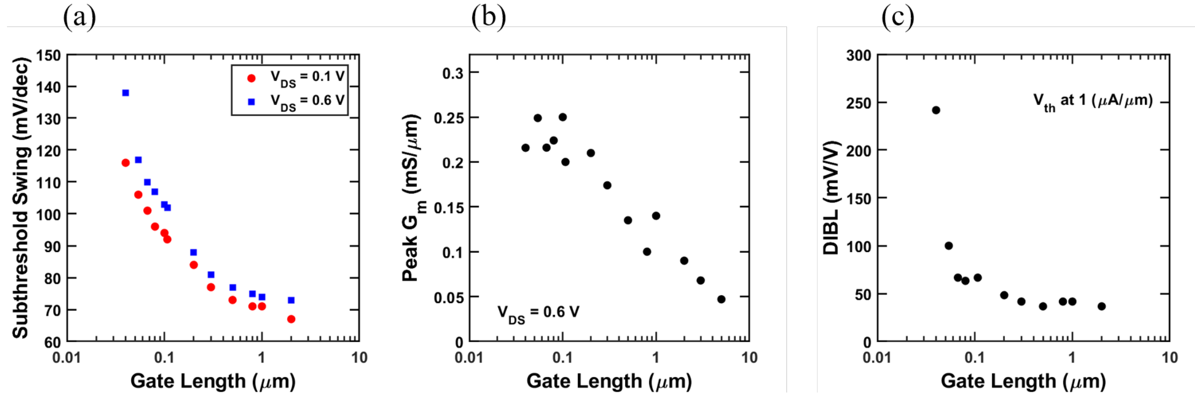


Figure 4.18: (a) SS vs. L_g at V_{DS} =0.1 and 0.6 V; (b)peak g_m vs. L_g at V_{DS} =0.6 V; (c) DIBL vs. L_g at V_{TH} =1 $\mu A/\mu m$ of MOSFETs using Ni/Au gates. © 2019 IEEE

layers can be incorporated into MOS device structures without interface defects causing significant degradations in DC characteristics. Later, MOSFETs with Ni/Au gates will be used as a reference for analyzing MOSFETs with ALD TiN/Ru gates.

The STEM image of an L_g = 30 nm MOSFET with TiN/Ru gate is shown in fig. 4.19. The regrowth facet with HSQ mask shows a ≈30°incline at the channel after InP growth and a ≈54°incline after InGaAs growth. The corresponding InP spacer thickness at the channel is roughly 5 nm. The inner highlight in fig. 4.19 is high-angle annular dark-field imaging (HAADF) STEM image. The interfacial layers contain ≈1/2.5/2 nm AlO_XN_Y/ZrO_2/TiN, as indicated by layers 1, 2, and 3, respectively. In addition, the thickness of Ru film in the channel is ≈10 nm, which is much thinner than thickness in-field (≈25 nm). The thickness variation (bad conformity) does not hurt the DC performances of planar MOSFETs. However, it could potentially increase the gate resistance or result in gate open-circuit in a more complex vertical transistor structure that requires metal gating/contacting the sidewall channel. Therefore, a more conformal Ru deposition is needed for non-planar/vertical MOSFETs/TFETs. Detailed processes to improve Ru conformity will be discussed later.

Fig. 4.20 and 4.21 shows the transfer and output characteristics of 30 nm, 80 nm,

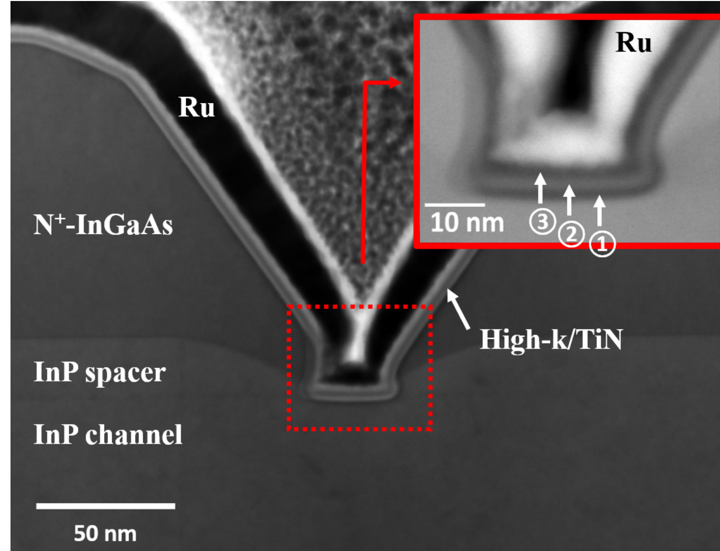


Figure 4.19: The STEM image of an L_g = 30 nm MOSFET with TiN/Ru gate. The inner plot in (b) is the HAADF-STEM image highlighting the structure at InP channel. Layer 1, 2, and 3 represent 1 nm AlO_xN_y, ≈2.5 nm ZrO_2 and ≈2 nm TiN, respectively.

and 200 nm-L_g FETs using TiN/Ru gates. The threshold voltage is ≈ 0.2 V smaller than Ni/Au gates determined by linear extrapolation. At a long $L_g > 200$ nm, a more than 5 orders of magnitude on/off ratio is achieved. The maximum g_m of ≈0.75 mS/μm for a $L_g = 80$ nm MOSFET at $V_{DS} = 0.6$ V is reached. This number is the highest g_m reported for InP channel MOSFETs. Due to a comparably thick channel of ≈18 nm, MOSFETs with $L_g < 80$ nm suffer from the short channel effect. Stronger DIBL results in a degradation of on/off ratio to 3 orders of magnitude for $L_g = 30$ nm MOSFET and a lower transconductance. Moreover, the high gate leakage current $I_g \approx 10^{-4}$ mA/μm at $V_{DS} = 1$ V in TiN/Ru gates is owing to the parasitic gated source/drain. As shown in fig. 4.14 (c), conformal TiN/Ru gates surround mesa edges. The total gated area is equivalent to more than 100 times the MOSFET active gate region area. Note that the active region means TiN/Ru gated InP channel and InGaAs S/D. By normalizing based on the estimated total gated area, the gate leakage current density in planar MOSFETs

is similar to the gate leakage of MOSCAPs, which is ≈30 mA/cm^2 at $V_g = 1$ V. The issue of high I_g could be eliminated by either employing a mesa-last (gate-first) process or conformal Ru etch. Note that Ru wet etch via Transcene RU-44 (ceric ammonium nitrate/ nitric acid) is not compatible with the fabrication process of planar MOSFETs because the etchant attacks underlying InGaAs much faster than etching Ru itself.

In addition, this relatively high I_g, which is within one to two orders of magnitude comparing to I_d, would result in an inaccurate acquisition of SS. The total drain current $I_d = I_{ds} + I_{dg}$, is a function of gate-to-drain leakage current. The modulated current I_{ds} is the measured I_d plus gate-to-drain leakage I_{gd}. When gate leakage is not negligible with respect to measured I_d, the subthreshold swing SS = dV_{gs}/dI_d would also be affected. Typically, a leaky gate results in an over-estimated low SS. It is therefore important to correct the SS by excluding I_{gd}. The correction of SS follows

$$I_{ds} = I_d + I_{gd} = I_d + xI_g \tag{4.1}$$

where x is the portion of gate leakage flowing to drain and is dependent on physical device structure, gate, source, and drain bias as well as the V_{gs} at where minimum SS happens. The corrected SS will be SS = dV_{gs}/dI_{ds} rather than dV_{gs}/dI_d. All the SS that is discussed in this thesis is corrected using this method. Due to the insignificant gate leakage, the corrected SS is roughly 1-2 mV/dec higher than the un-corrected one in MOSFETs using TiN/Ru gates.

Peak g_m at V_{DS} =0.6 V vs. L_g (a) and SS vs. L_g at V_{DS} =0.1 V (b) of MOSFETs using thermal evaporated Ni/Au gates, ALD TiN/Ru gates, and ALD Ru gates are compared in fig. 4.22. ALD TiN/Ru gate has the highest g_m of 0.75 mS/μm for an L_g = 80 nm transistor. On the contrary, Ru gates show the lowest transconductance, which agrees with a large frequency dispersion in the C-V characteristic, as in fig. 4.9 (b). In

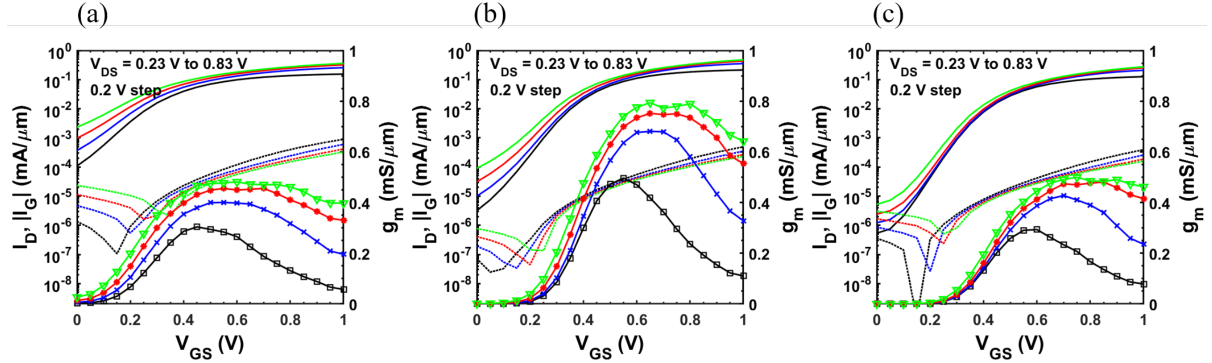


Figure 4.20: Transfer charateristics of 30 nm, 80 nm, and 200 nm-L_g MOSFETs using TiN/Ru gates.

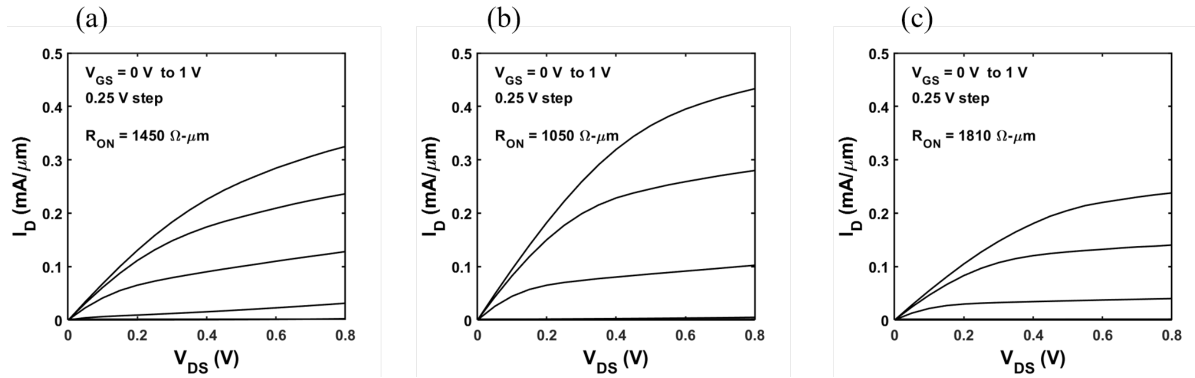


Figure 4.21: Output charateristics of 30 nm, 80 nm, and 200 nm-L_g MOSFETs using TiN/Ru gates.

addition, Ru gate MOSFETs have high device variation and a poor yield, which could be due to Ru's bad nucleation on high-k. As a result, the SS of Ru gate MOSFETs with long gate length is > 80 mV/dec, and its SS vs. L_g does not show a clear trend.

Table 4.6 compares the averaged SS at V_{ds} = 0.1 V for long gate length devices and maximum peak g_m at V_{ds} = 0.6 V for thermal evaporated Ni/Au gates, ALD TiN/Ru gates and ALD Ru gates. As can be seen, a record low SS of 68 mV/dec is achieved by ALD TiN/Ru gate metallization, and the corresponding D_{it} is $\approx 1 \times 10^{12} cm^{-2} eV^{-1}$ at the current EOT (≈ 1 nm), which is 3 times smaller than that in Ni/Au gates.

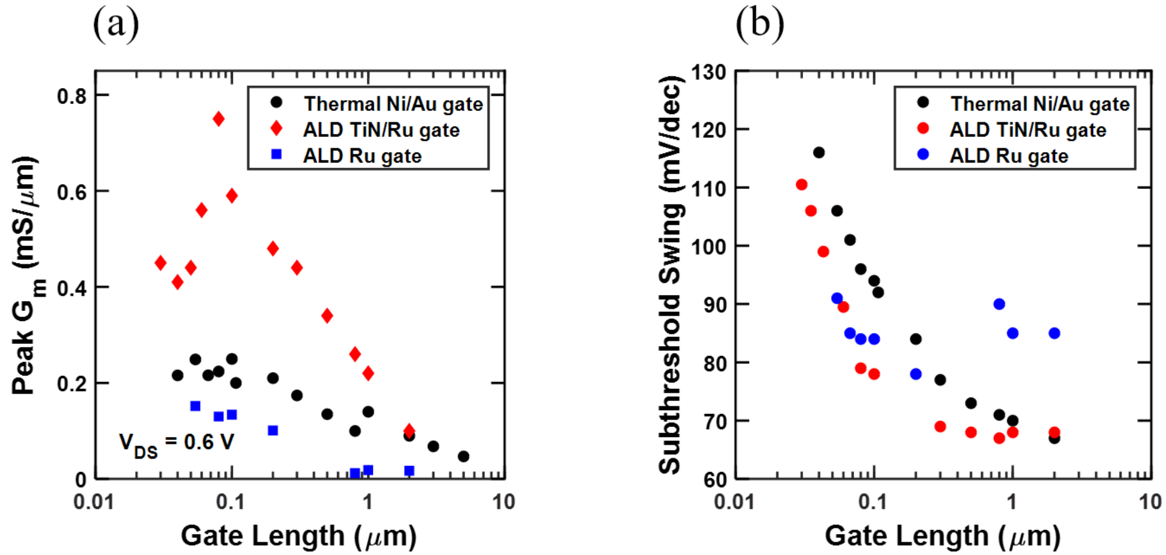


Figure 4.22: (a) Peak g_m at V_{DS} =0.6 V vs. L_g and (b) SS vs. L_g at V_{DS} =0.1 V (b) of MOSFETs using thermal evaporated Ni/Au gates, ALD TiN/Ru gates and ALD Ru gates.

Gate Metallization	Thermal Ni/Au	ALD TiN/Ru	ALD Ru
Avg. SS at $V_{ds} = 0.1$ V for $L_g > 200$ nm	70 mV/dec	68 mV/dec	80 mV/dec
Max. peak g_m at $V_{ds} = 0.6$ V	0.75 mS/μm	0.25 mS/μm	0.15 mS/μm

Table 4.6: Summary of long gate length SS and peak g_m for MOSFETs using thermally evaporated Ni/Au gates, ALD TiN/Ru gates, and ALD Ru gates.

In conclusion, InP channel planar MOSFETs using thermal evaporated Ni/Au gates, ALD TiN/Ru gates, and ALD Ru gates are fabricated and compared. Controlled Ni/Au gate MOSFETs prove a decent ZrO_2/InP interface by a low averaged SS of 70 mV/dec for long gate length MOSFETs. MOSFETs using ALD TiN/Ru gates show record-high peak g_m of 0.75 mS/μm at V_{DS} = 0.6 V and a record low averaged SS of 68 mV/dec at V_{DS} = 0.1 V, indicating a defect-free gate metallization process. Despite its chemical and layer-by-layer growth mechanism, a more than 66% variation in Ru thickness over the channel is observed via TEM analysis. An improved Ru conformity is needed for complex non-planar/vertical device structures and will be discussed in section 4.4.

4.3.4 U.I.D. Channel versus P-doped Channel

TBD

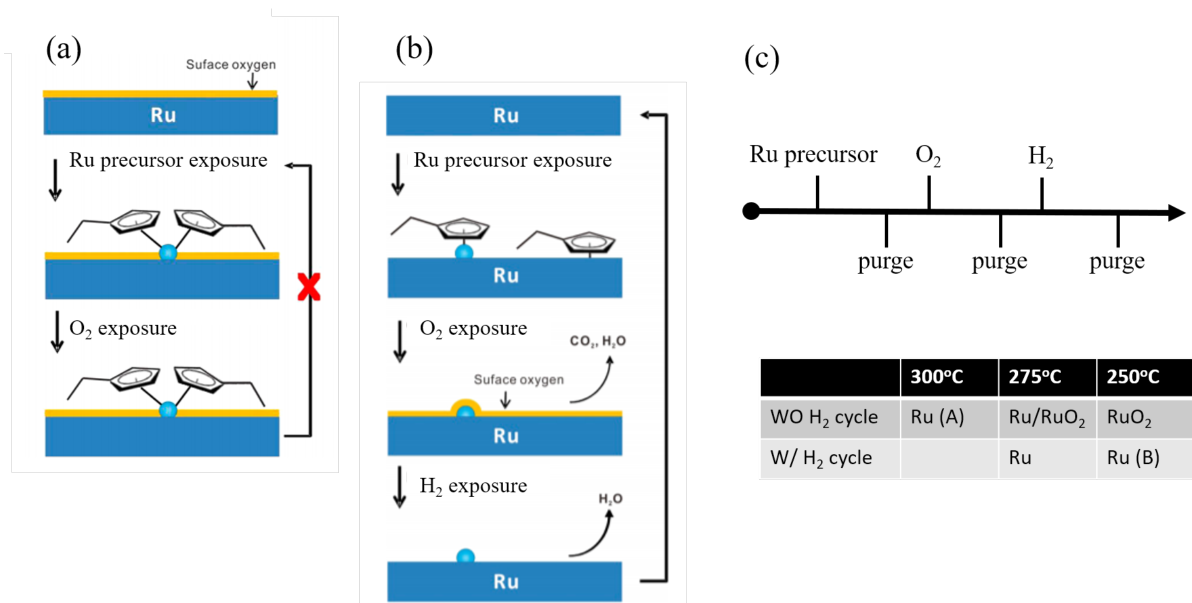
4.4 Low Temperature Conformal ALD Ru Growth

As shown earlier in fig. 4.19, even though the metallic Ru film grown at 300°C has excellent conductivity ($\approx 20\ \Omega/\square$ for a 30 nm film), it is compromised by an inferior conformality on sidewalls due to a small amount of thermal decomposition of the precursor [90, 91]. The poor conformity (thickness variation) does not degrade the DC performance of planar MOSFETs but could be a problem for a more complex device structure. Fig. 4.24 (a) shows the TEM images of a vertical MOSFET using TiN/Ru gate, in which Ru is deposited at 300°C.Note that this vertical MOSFET has undercut and double-sided sidewall channel. Red arrows point to Ru metal (black) gating on the channel. As can be seen, the thickness of Ru on sidewalls is < 10 nm, which is much thinner than Ru deposited in-field (≈ 25 nm). At the bottom and upper corner of the undercut channel, Ru thickness is even smaller. Vertical MOSFETs that use Ru grown at 300°C in gates show gate open-circuit, which agrees with the TEM analysis.

In order to improve Ru conformity, a lower temperature growth is performed. However, simply lowering the temperature below 300°C (to avoid the CVD channel) results in little to no metallic deposition via the standard EBCHDRu and oxygen ALD cycle. It has been found that surface RuO_2 is formed as part of the ALD Ru process at a lower substrate temperature. The RuO_2 is stable and prevents further adsorption of the Ru precursor, thus inhibiting growth [92]. It could also be explained by the formation of partly dehydrogenated carbonaceous species on the surface, which block further catalytic combustion and dehydrogenation reactions in the Ru process [93]. Therefore, the

pulsed EBCHDRu following pulsed O_2 at a low temperature causes a self-limited process [92, 93].

Lu *et al.* reported an ABC-type low-temperature Ru growth and demonstrated a metallic Ru film grown at as low as 150°C using H_2 as a coreactants, and fig. 4.23 shows the schematic diagram explaining the process of Ru growth using (a) traditionally AB-type growth (in which O_2 is used to combust) with a RuO_x starting surface, and (b) the ABC-type growth with a Ru starting surface [92]. In the AB-type Ru growth, surface oxide stops the ligand removal of Ru precursor, while in ABC-type Ru growth, surface oxygen is removed by H_2 exposure. The role of H_2 in the process is to reduce surface oxygen by forming H_2O. The sequence of pulsing reactants, as shown in fig. 4.23 (c), follows dosing Ru precursor, hold, pump, and purge; dosing O_2, pump, and purge; then dosing H_2, pump, and purge.



	300°C	275°C	250°C
WO H_2 cycle	Ru (A)	Ru/RuO_2	RuO_2
W/ H_2 cycle		Ru	Ru (B)

Figure 4.23: Schematic diagram explaining the process of Ru growth using (a) traditionally AB-type growth (Ru precursor-O_2) with a RuO_x starting surface, and (b) the ABC-type growth (Ru precursor-O_2-H_2) with a Ru starting surface, and (c) the sequence of ABC-type reactants exposure. (a) and (b) Reprinted with permission from [92]. © 2015 American Chemical Society.

This ABC-type growth is incorporated in our process, and the detailed growth parameters are summarized in the appendix. The growth rate is ≈ 0.37 A/cycle at 250°C, which is ≈ 40% lower than 0.6 A/cycle at 300°C because of fewer surface reactions due to energy reduction. The conformity of Ru grown at 250°C is shown in fig. 4.24 (b). As pointed by red arrows, Ru layers (black) grown at 250°C shows no variation in thickness on sidewall channel as in-field. Comparing to Ru deposited at 300°C (as in (a)), it is clear that improved conformity is achieved by a lower temperature ABC-type Ru growth.

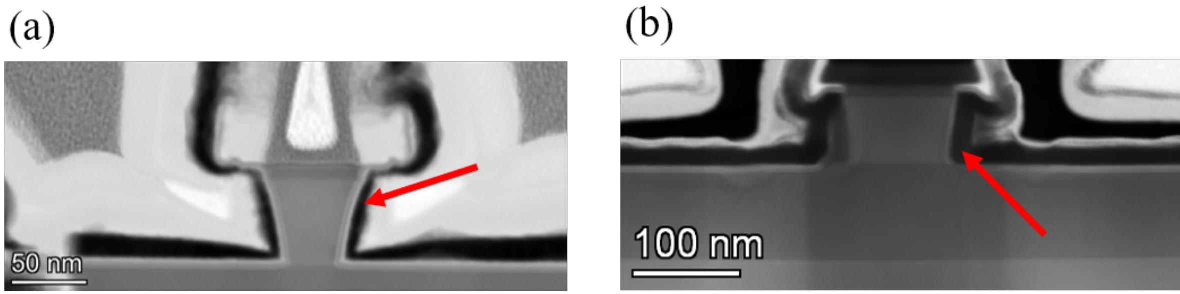


Figure 4.24: TEM images of vertical MOSFETs using TiN/Ru gates in which Ru layer is grown using (a) AB-type growth at 300°C and (b) ABC-type growth at 250°C.

In conclusion, several gate metals are explored. Using ALD TiN/Ru gate metallization, MOSFETs show record low SS, high peak g_m on the InP channel, indicating a high-quality defect-free HKMG/InP interface. The variation in Ru thickness is observed on recessed-gate planar MOSFET and vertical MOSFET structure, which is then improved by a lower temperature ABC-type Ru growth.

Chapter 5

Process Development and First Demonstration of 3-HJ TFETs

To realize high-performance 3-HJ TFETs, a suitable device structure and according to process development are required. In section 5.1, commonly seen TFET structures in literature are reviewed. Their compatibility with the 3-HJ TFET design is discussed, and the pros and cons for each structure are detailed. Section 5.2 shows the three structures investigated at UCSB– 1. confined epitaxial lateral overgrowth (CELO), 2. anchor-shape top-down vertical process, 3. top-down vertical planarization process. Main challenges in gen. 1 and 2 structures are addressed. A thorough explanation of process modules development/transfer and the integration is given in 5.3. Results of vertical InP channel MOSFET and vertical 3-HJ InGaAs/GaAsSb/InAs/InP TFET are shown and discussed in 5.4 and 5.5, respectively.

5.1 Structure Overview– From Planar to Vertical TFETs

Similar to MOSFETs, TFETs are surface devices that electrons/holes flow near the surface of materials. Carriers move as the channel potential is modulated by external voltage bias. No matter in any device structures, the modulated current flows only in the channel close to the high-k/semiconductor interface. In the fabrication of III-V-based TFETs, both planar and vertical structures are employed, and each of them has its pros and cons.

Fig. 5.1 shows a kind of vertical TFET structure that uses a top-down approach and has its gate on mesa edges. Depending on different ways of etching mesa, channels could be formed perpendicularly to junctions or with a slope (like a pyramid), as shown in fig. 5.1 (a) and (b). The fabrication process for gated mesa edge TFET is quite simple, including mesa etching, source, drain lift-off, and HKMG growth and patterning. However, the on-current is limited by the relatively long tunneling distance given a μm size body. The body leakage would also result in short channel effects and degrade SS.

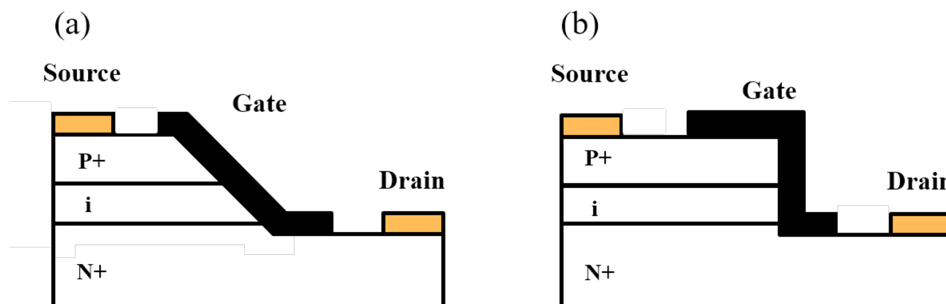


Figure 5.1: Schematic gated mesa edge vertical TFET structures (a) with pyramid-shape, and (b) with channel perpendicular to tunneling junction.

Improved methods of fabricating a vertical TFET using a similar top-down approach have been reported in [94] and [95], and their schematic structures are shown in fig. 5.2 (a)

and (b). The difference between them is the order of process steps: drain contact first (fig. 5.2 (a) or drain contact last (fig. 5.2 (b)). In the drain contact first process, drain metal is deposited initially and used as the mask for the subsequent etching. Semiconductor layers are wet etched in 2 steps: 1. anisotropic etch through N^+ and i layers to expose P^+ source; 2. isotropic etch to undercut the features for a thin body [94]. To probe the small dimension drain, planarization following by large-area metal pads lift-off is needed. On the other hand, drain contact last process could achieve a thin body by dry etching. Alian *et al.* reported top-down dry-etched TFETs with a low SS of 47 mV/dec at a body diameter of 30 nm and gate length of 100 nm [95]. Drain contact last process starts with semiconductor dry etching. After that, gate and source metal are deposited and patterned. In order to contact a small dimension N^+ layer, planarization is again needed. Drain metal contact and large area metal pads are formed at the same time. Note that the spacer specified in fig. 5.2 (a) and (b) is spin-on dielectric/resist, which provides good insulation between metal layers and is commonly applied in the planarization process.

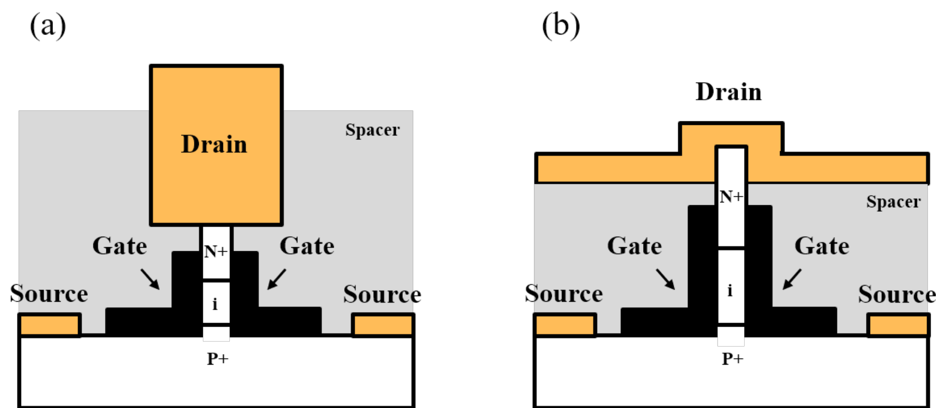


Figure 5.2: Schematic vertical TFET structures using top down approach with (a) drain contact first and (b) drain contact last.

Nanowire growth is another solution to small dimensional devices aside from forming the thin body through wet/dry etch. Memisevic *et al.* from Lund University first reported nanowire TFETs at 2016, with SS of 48 mV/dec and up until now the record high I_{on}

of 10 $\mu m/\mu m$ at $V_{DS} = V_{GS} = 0.3$ V [39]. Fig. 5.3 (a) shows the schematic structure of nanowire TFETs. The fabrication of nanowire TFETs starts from Au discs array deposition and patterning following by nanowire growth of designed semiconductor layers on the substrate, where the substrate contains a highly N^+ doped drain layer on a resistive material. Nanowire growth of N^+, i, and P^+ layers is done in sequence, and the dimension of nanowires are shrunk by cycles of digital etch after growth. Note that the bigger head in P^+ region in fig. 5.3 (a) is formed due to a different etching rate of digital etch in materials [39]. A similar technique is used to form drain and source metal contact and HKMG formation. However, to remove gate metal surrounding the bigger head P^+ region, planarization is needed. A thick spacer layer is spined on/deposited and etched back to expose the head, and gate metal is removed in the exposed region. In order to contact the source and prevent short circuit source/gate, the second planarization is done. Once another spacer is spined on/deposited and etched back to expose the top of P^+ layers, the source contact, and measurement pad are lifted off in one go. Overall, the fabrication process of nanowire TFETs requires two times planarization to contact the drain layer and effectively isolate gate/drain metal, which is similar to the improved vertical top-down approach. The growth of nanowire with dimension shrink by digital etch enables a thin body of $\approx$ 10 nm [39].

Tomioka *et al.* reported core-multishell nanowire TFETS with an improved SS of 25 mV/dec and I_{on} of 2.4 $\mu A/\mu m$ at $V_{DS} = V_{GS} = 0.3$ V, which is owing to the formation of two dimensional electron gas (2DEG) throughout the channel [96]. The core-shell TFET structure is shown in fig. 5.3 (b). The fabrication process is almost identical to the nanowire TFETs as in (a). After the growth of i and N^+ layers, semiconductor shells are grown surrounding the whole wires. In addition, the substrate for the growth of nanowires has a P^+ source layer on it. The tunneling junction for N-TFET operation, which lies between P^+ and i layers, is defined at the nanowire growth interface. This method enables

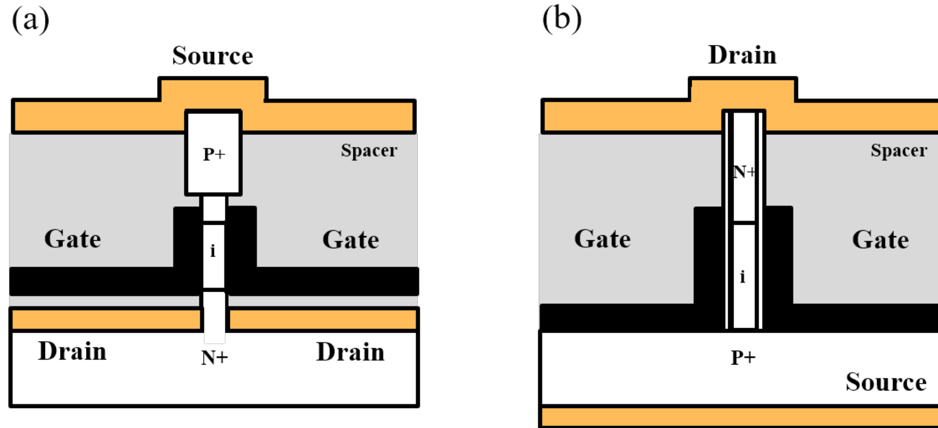


Figure 5.3: Schematic vertical nanowire TFET structure (a) and multi-coreshell nanowire with source at the bottom (b).

a self-aligned gate which is critical for achieving high-performance TFETs, as discussed in chapter 3.

Because of the requirement of a thin body for a high I_{on}, TFETs use a planar device structure that must be fabricated on a platform/template, which has semiconductor layers on a dielectric. In silicon-based TFETs, an SOI substrate is ideal and a good starting point for pursuing a thin body. In III-V based materials, TFETs fabricated on a box using direct wafer bonding (DWB) technique has been achieved by Convertino *et al.*, showing SS of 49 mV/dec and I_{on} of 0.2 $\mu m/\mu m$ at $V_{DS} = V_{GS} = 0.3$ V [97]. The device planar DWB TFET structure is shown in fig. 5.4. The fabrication for this planar DWB TFETs requires selective regrowth of P^+ and N^+ layers. After double regrowth, a recessed metal gate is formed, and source/drain contact metal is lifted off. This process is compatible with the fabrication of Si finFETs: both of them use P^+ and N^+ layers regrowth and recessed gate formation. In addition, the channel sheet (i layer) could be patterned and etched to form parallel nanowires to improve gate electrostatics. The fabrication of planar DWB TFET is the simplest by far (excluding gated mesa edge TFETs) and is promising for its proven feasibility of integration with Si CMOS [97].

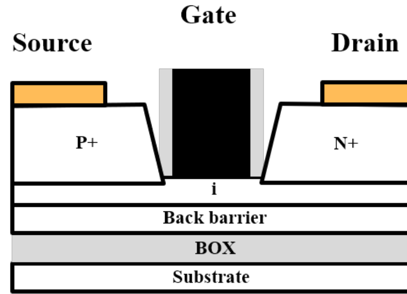


Figure 5.4: Schematic lateral TFET structure using direct wafer bonding and source/drain selective regrowth.

Another approach towards a thin body planar TFET is template-assisted selective epitaxy (TASE). Cutaia *et al.* reported the first lateral nanowire TFETs using TASE technology [98]. Note that TASE has also been phrased as confined epitaxial lateral overgrowth (CELO). The fabrication of TFETs using TASE structure could be divided into three parts: fabrication of templates/boxes, material (including N^+, i, P^+ layers) growth in the templates/boxes following by top dielectric removal, HKMK, and source/drain metal contact formation. As shown in fig. 5.5 (a), the body thickness and shape of TFETs are defined by the inner boundary of the template/box. A lying fin or a lateral nanowire TFETs are both achievable using TASE structure. Another TFETs structure could be formed using TASE, as shown in fig. 5.5 (b). In this case, the material growth inside the template/box contains a uniform i layer. After confined growth, the top dielectrics are removed. N^+ drain and P^+ source layers are selectively regrown. The rest of the process steps are identical to what has been discussed in (a). The advantage of the structure in (b) over (a) is the avoidance of growing arbitrary materials and heavily doped layers in boxes, which could result in parasitic nucleation and defect generation. These challenges of employing TASE technology in pursuing TFETs will be studied and detailed discussed in 5.2.1.

Table 5.1 summarizes the most commonly seen TFET structures that are discussed

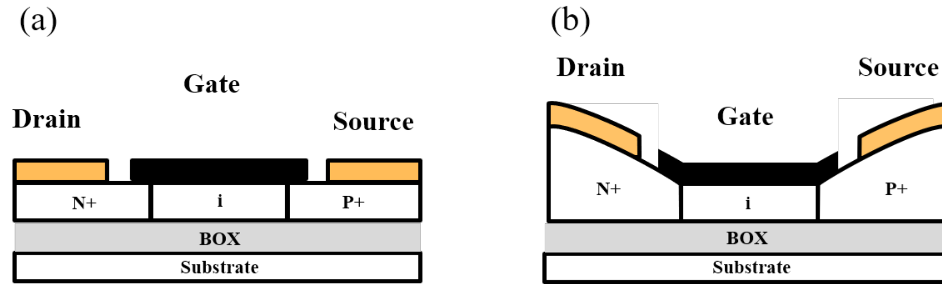


Figure 5.5: Schematic lateral TFET structures using TASE technique with (a) all layers grown in confined box, and (b) only U.I.D. channel grown in confined box and source/drain selective regrown.

before. Structure 1 has the most uncomplicated fabrication process but is compromised with its body leakage as well as low on-current is given by the poor electrostatics. A thin body is achievable in structures 2 - 5, while their fabrications are all not simple. Material growth in confined boxes in 5 and nanowire in 3 requires a lot of effort and investigation. The development of wafer bonding technology is also challenging and time-consuming. Planarization process with high accuracy makes a device batch turn-around much longer. Overall, none of them are simple, and a large amount of effort in process development is needed.

The goal is to demonstrate 3-HJ InGaAs/GaAsSb/InAs/InP TFETs in this work. The compatibility of 3-HJ TFET with the device structure needs to be considered. Heterojunctions perpendicular to the TFET current flow direction is favored for a maximized junction electric field. In addition, structures that require source/drain regrowth are less preferable because of potential defects/impurities accumulation at the regrowth interface that could affect heterojunction TFET performance.

CELO technique is firstly chosen to pursue 3-HJ TFETs (5.2.1). After 1 - 2 years of process development and growth investigation, an anchor-shaped top-down vertical structure has come up (5.2.2). In the next two years, this structure is then proven to be problematic. In the 3rd generation of structures (5.2.3), the anchors are removed,

Ref.	[99]	[94, 95]	[39, 96]	[97]	[98]
TFET Structure	Gated Mesa Edge	Vertical Etch	Nanowire	Wafer Bonding	Template Assisted Selective Epitaxy
Current Flow	Vertical	Vertical	Vertical	Lateral	Lateral
Type	Top-down	Top-down	Bottom-up	Bottom-up	Bottom-up
Epi	Epitaxial layers on blank substrate	Epitaxial layers on blank substrate	Nanowire growth of all layers	N^+ and P^+ selective regrowth	Confined growth of layers in small dimension boxes /employing with N^+ and P^+ selective regrowth
Ways of Probing	No special process needed	2 times planarizations	2 times planarizations	No special process needed	No special process needed
Body Thickness	$>\mu m$	20 - 100 nm	down to 10 nm	down to 10 nm	20 - 50 nm
Pros	Simple process Well-established bare wafer III-V growth	Well-established bare wafer III-V growth	Good electrostatics (Record TFETs)	Good electrostatics CMOS integrable	Good electrostatis
Cons	Body leakage Bad electrostatics	Complex process	Complex process Growth effort	Wafer bonding technology development	Growth effort Template formation technology development

Table 5.1: Summary of most commonly seen TFET structures

while one planarization process is added. Based on each device structure, InP channel MOSFETs are fabricated to prove the process. At last, the top-down planarization process is used to pursue vertical 3-HJ TFETs.

5.2 Process Development Towards 3-HJ TFETs

5.2.1 Gen. 1–Confined Epitaxial Lateral Overgrowth

Brian Markmann does detailed studies on the formation of CELO template, and comprehensive investigations on growth parameters, heterojunctions, ternary materials, and defects in InP CELO are done by Simone Tommaso Šuran Brunelli and Aryana Goswami at UCSB. A thorough discussion on CELO development could be found in their theses. CELO structure is first brought up to demonstrate the 3-HJ TFETs because it

enables perpendicular junctions concerning tunneling direction. The planar structure of CELO has quite simple processing steps for the fabrication of FETs once material growth is done. This would speed up the batch turnaround, which is essential in developing new device structures and transistor iterations.

CELO requires the formation of growth templates. Fig. 5.6 shows the schematic fabrication process flow for making a template. Fabrication of template starts from a blanket semiconductor substrate, and InP substrate is mostly used in this study (fig. 5.6 (a)). A 3 nm ALD grown Al_2O_3 is first deposited as an etch stop layer to protect the InP surface from the upcoming plasma damage by dry etching. A dielectric material (SiO_2 or SiN_x) is blanket deposited, patterned and dry-etched to form seed holes (fig. 5.6 (c)). A sacrificial layer, which could be amorphous silicon or resist, is then blanket deposited, patterned, and etched (fig. 5.6 (e)). The outline of the pattern of sacrificial material defines the cavity of the box. Template fabrication is finished by either blanket deposition of top dielectric layer (SiO_2) or spin-coated a spin-on dielectric, patterned and dry-etched to form source holes (fig. 5.6 (g)), following selectively removal of sacrificial layer (fig. 5.6 (h)). The removal of sacrificial amorphous silicon or resist is done by XeF_2 gas or resist stripper, respectively. Prior to loading into the MOCVD system, the 3 nm Al_2O_3 etch stop layer is etched by $>$ 5 minutes dip in 300MIF developer, and one cycle of digital etch and 0.3% HF dip for 10 seconds are done to clean up InP regrowth surface.

During MOCVD growth of designed materials in the cavity, MO precursors get into the box through source holes. They further diffuse inside the box and reach the semiconductor surface at seed holes. With the template shown in fig. 5.6, the material growth is initially vertical and then lateral throughout the box after making a turn. Hetero- or doping- junctions would be perpendicular to the growth direction. The growth rate of materials in confined boxes is much lower than in-field due to the limited amount of MO and MO diffusion delay inside the cavity. In addition, the diffusion constant is different

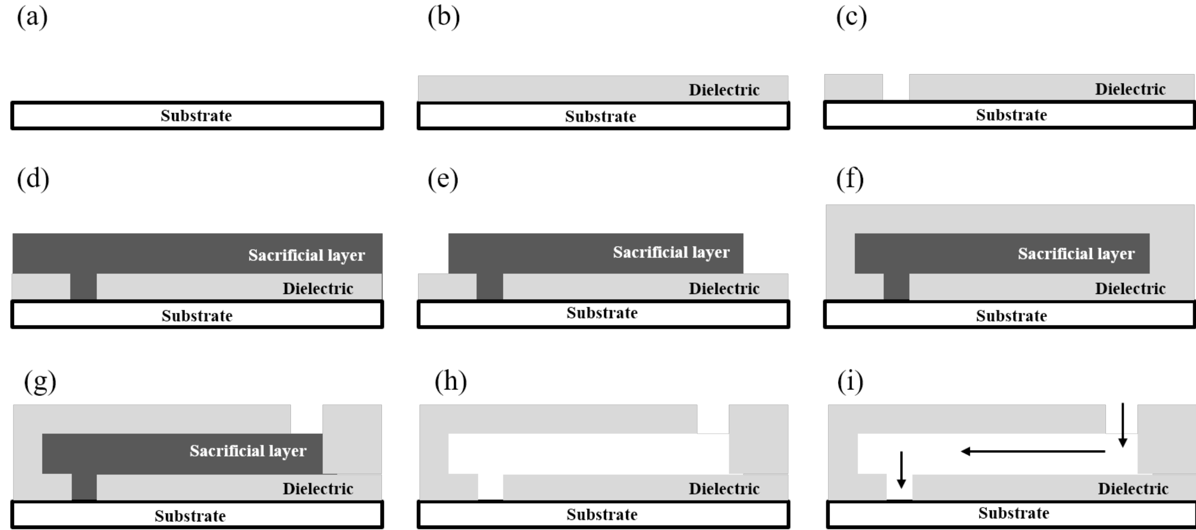


Figure 5.6:

for each precursor. It is as a result that lateral composition gradient inside boxes is observed on ternary material growth.

Growth across the wafer is uniform in a bare wafer material, and the growth front is always parallel to the substrate surface plane. On the contrary, crystal growth in confined boxes favors different growth planes. The higher/lower growth rate for each plane is dependent on confined/growth direction, defects, and growth parameters. Fig. 5.7 shows the TEM analysis of CELO growth on (001) InP substrate at 600°C [100]. Crystal growth under the current growth condition and substrate favor multiple growth planes, as summarized in fig. 5.7 (f). The formation of junctions not perpendicular to the substrate surface would degrade TFET performance due to a lower junction field and reduced electrostatics. Thus, template fabrication and growth parameters need to be optimized for sharp and perpendicular heterojunction interfaces.

The shape of the box could be designed arbitrarily, as long as an exposed seed hole is present inside the cavity. The cavity height, which is determined by the thickness of the sacrificial layer, is ≈ 50 nm to ensure uniform diffusion of metal-organic precursors in

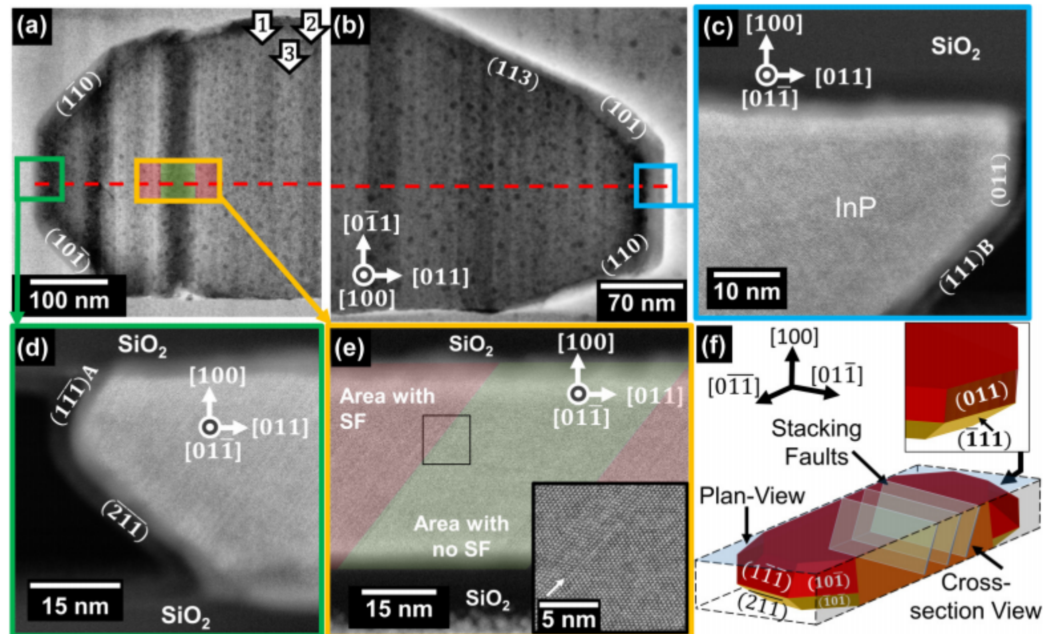


Figure 5.7: InP CELO grown at 600°C on (001) InP substrate with template sitting along [110]. (a) and (b) top view TEM images at two ends of the crystal. (c) and (d) cross-sectional TEM images cutting from red dash lines in (a) and (b). (e) Cross-sectional TEM images focusing on defects. (f) Schematic structure showing the formed growth planes in the box. [100] Copyright © 2020 by American Physical Society. All rights reserved.

boxes in our study. A <30 nm thick sacrificial design might result in the collapse of boxes during the surface wet clean before MOCVD growth because of the large surface tension between the top and bottom dielectrics. In addition, bowing of boxes using resist-based template fabrication is also observed after high-temperature MOCVD growth, as shown as process R in fig. 5.8 [101]. It is therefore important to properly design the cavity considering those restrictions.

The concept of CELO is exciting, and the feasibility of a thin body makes it promising in modern highly-scaled MOS transistors. However, there are two main problems in CELO that are hard to be solved: parasitic nucleation and defect generation. In ideal selective area growth, the nucleation should only happen on the semiconductor surface. In reality, the dielectric surface energy is modified by series of processing, and the

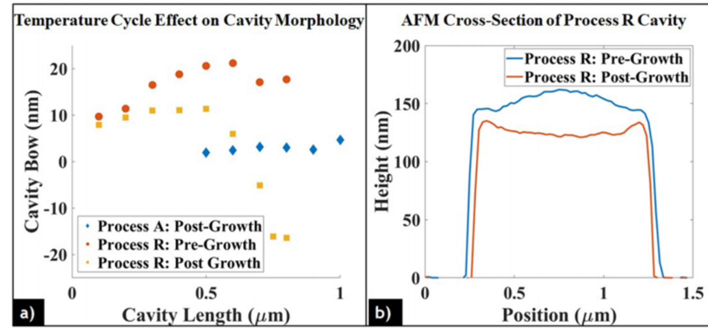


Figure 5.8: (a) Cavity bow vs. cavity length, and (b) cavity morphology before and after growth. Process A and R stand for template fabrication using thin-film deposited and resist-based sacrificial and top dielectric layers, respectively. Reproduced from [101], with the permission of AIP Publishing.

nucleation on dielectrics is always observed. The organic/inorganic residues/polymers generated from resist application, wet and dry etching could also serve as nucleation sites anywhere on the template. Fig. 5.9 shows the (a) SEM and (b) OM images of parasitic nucleation on the CELO template after MOCVD growth. As can be seen, large amounts of parasitic grown crystals are observed both in the field and on top of boxes. Note that the parasitic grown crystals have sizes much greater than crystal grown in boxes because of a higher growth rate in-field with respect to inside boxes. The presence of these giant parasitic crystals makes the further fabrication of FETs very challenging. Furthermore, parasitic nucleation happens not only outside boxes (in-field) but also inside boxes. The residuals from the box preparation process form nucleation sites inside boxes. The unwanted crystal nucleates and grows on those arbitrary distributed cites in the cavity would stop/affect the rest of growth from seed holes. In the MOCVD growth of III-V materials, the growth window (depending on reaction paths) lies roughly between 450 - 650 °C. Some of the doped layers require a low growth temperature for a higher electron/hole concentration, e.g., carbon-doped P-InGaAs and GaAs. However, the growth temperature in CELO is usually kept high ($\approx$ 600°C) to reduce parasitic nucleations. In addition, it is found that a higher flow of DiSi and the incorporation of

As (for InGaAs growth) in CELO would give rise to parasitic nucleations.

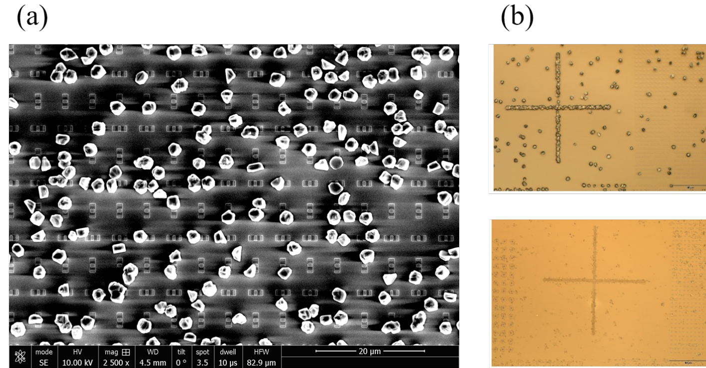


Figure 5.9: Top view (a) SEM and (b) OM images of parasitic nucleation on CELO template after MOCVD growth. Courtesy of Brian Markman, Simone Šuran Brunelli, Aranya Goswami.

Defect formation is another limiting factor of the application of CELO in TFETs. As discussed in chapter 3, trap-assisted tunneling degrades TFET off-state performance, where the trap states are straightforwardly linked with defects in crystals. In conventional bare wafer epitaxy, achieving defect-free high-quality crystals is not very hard. In contrast, many defects could form in confined growth for many reasons: parasitic nucleation inside boxes, not-optimized growth parameters, wafer orientation, and surface roughness of dielectric layers. A nearly defect-free InP CELO growth and a vertical growth front are achieved on a (110) InP substrate with template sitting along $[1\bar{1}0]$. On a (100) InP substrate with template sitting along $[0\bar{1}\bar{1}]$, numbers of twins/stacking faults are generated in CELO growth at the growth temperature of 550 - 630°C, as shown in fig. 5.10. The origin of defect generation is attributed to the roughness of the dielectric surface, and different thermal expansion coefficients of dielectrics and grown materials [100].

Defect-free and sharp vertical heterojunctions (InP/GaAs/InAs) are both successfully achieved in CELO growth at UCSB using (110) InP substrate [102]. Nevertheless, with large amounts of effort and all the great works in these years, it is still challenging to

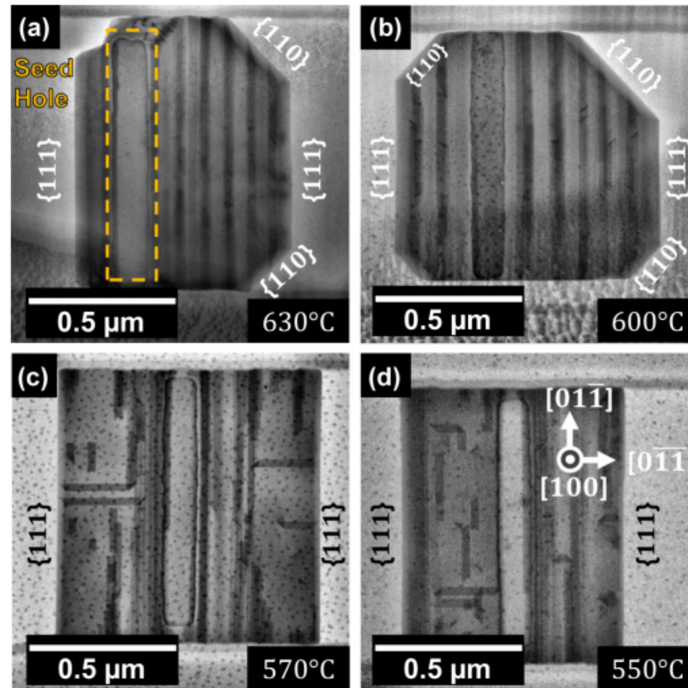


Figure 5.10: Top view TEM images of InP CELO growth on (100) InP substrate at (a) 630°C, (b) 600°C, (c) 570°C, (d) 550°C. Template sits along [$0\overline{11}$]. [100] Copyright © 2020 by American Physical Society. All rights reserved.

deal with all the non-idealities while pursuing heterojunctions and the designated doping concentration in certain materials, especially in an academic cleanroom. Therefore, the pursuit of 3-HJ InGaAs/GaAsSb/InAs/InP TFET is changed to the Gen. 2– anchor-shape top-down vertical structure.

5.2.2 Gen. 2–Anchor-Shape Top Down Vertical Structure

Anchor-shape vertical structure is secondly brought up as an alternative approach for pursuing 3-HJ TFETs, featuring a top-down self-aligned gate and no planarization process. Two metal pads are formed and connected to the fin (active device region) simultaneously during the initial drain contact formation. Fig. 5.11 (a) shows top view schematic anchor-shape vertical structure. The anchors on both sides are believed to

support a thin fin at a thickness < 15 nm, where fin collapse could happen due to surface tension. To isolate between fin (active device) and two pads, wet etch to undercut the semiconductor layers is required. Red circles in fig. 5.11 (a) indicate the region that has semiconductor underneath undercut. Failure on isolation would result in a huge parasitic body leakage under pads, which degrades off-state performance and SS in MOSFETs/TFETs. Furthermore, the electrical signal is coming from two pads at the side. A larger drain resistance is expected with a small fin width and large fin length. A small fin width is critical and essential in MOSFETs and TFETs for good electrostatics. Thus, fin length is kept under 5 μm to ensure the drain resistance would not dominate FETs performances over channel resistance.

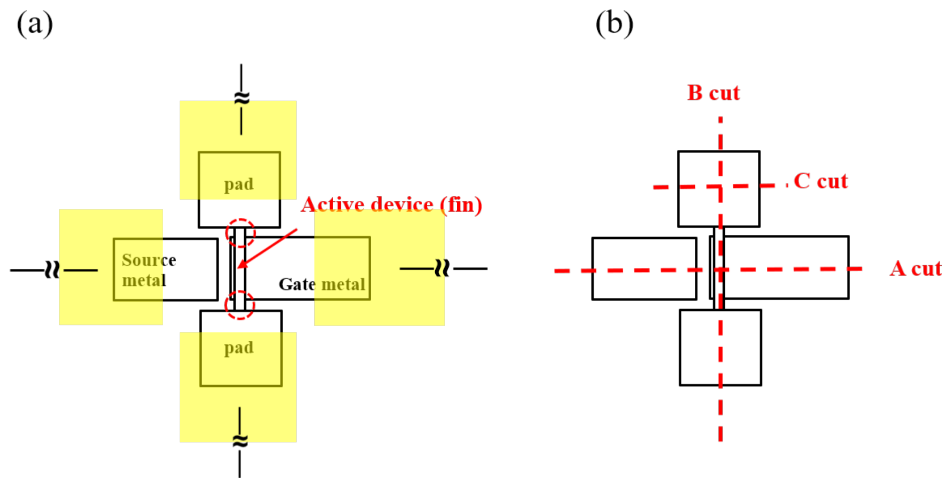


Figure 5.11: (a) Top view schematic anchor-shape vertical structure and (b) 3 different cuts in cross-sections used in explaining the process flow for anchor-shape vertical FETs (that will be shown later).

Fabrication process flow for anchor-shape vertical FETs is divided into three parts: front-end processing, isolation, and back-end wiring. Front-end processing focuses on the active device (fin) region, which will be explained using cross-sectional schematic diagrams cutting along fin, as shown in the A cut in fig. 5.11 (b). Similarly, the goal of isolation is to isolate between pad and fin; thus, it will be shown using a B cut. Large

measurement pads are directly connected to the two pads supporting the fin, so back-end wiring will be illustrated with a C cut. Note that since anchor structure is not the final structure that succeeds, the explanation of process flow would not go detailed.

The fabrication of anchor-shape vertical FETs starts from blanket drain contact Mo/TiW (20/150 nm) deposition. After that, a ≈ 40 nm of HF-resistive SiN_x and 30 nm of Cr is deposited (fig. 5.12 (b)). Drain stack is further patterned and dry etched to expose semiconductor layers (fig. 5.12 (c)). Fist SiN_x sidewalls (≈ 20 nm) are formed (fig. 5.12 (d)) and 10 nm of InGaAs is wet etched (fig. 5.12 (e)). To expose the InP sidewall channel, second SiN_x sidewalls (≈ 15 nm) are formed (fig. 5.12 (f)) and InP channel is wet etched under HCl-based solution (fig. 5.12 (g)). HCl-based InP wet etch has high selectivity over arsenide materials. Thus, the etch of channel would stop on InAs channel well in TFETs, or on P^+ InGaAs source layer in MOSFETs, leading to a self-aligned gate at the channel/source junction. Consiering the thickness of high-k and InAs channel well, a well-aligned gate requires a few cycles of digital etch into epi layers prior to high-k deposition (fig. 5.12 (h)). ALD in-situ nitrogen plasma clean, high-k and TiN/Ru gate deposition, and pre and post metal H_2 annealing are all the same as in the fabrication of planar MOSFETs (fig. 5.12 (i)). Gate metal and high-k are then pattered and dry etched (fig. 5.12 (j)) and source contact is patterend and lift-off (fig. 5.12 (k)).

The cross-sectional schematic structure of device after front end process along B cut (in fig. 5.11 (b)) is shown in fig. 5.13 (a). The isolation process starts from patterning and isotropic wet etch of InGaAs. The InGaAs wet etch not only etch down to expose InP substrate underneath but also undercut all the arsenide layers (InAs, GaAsSb, InGaAs in TFETs, and InGaAs only in MOSFETs) at the unmasked region. After that, InP wet etch is done to etch down into the substrate and also up to remove the InP channel, as shown in fig. 5.13 (c).

The cross-sectional schematic structure of the device after the front-end process and

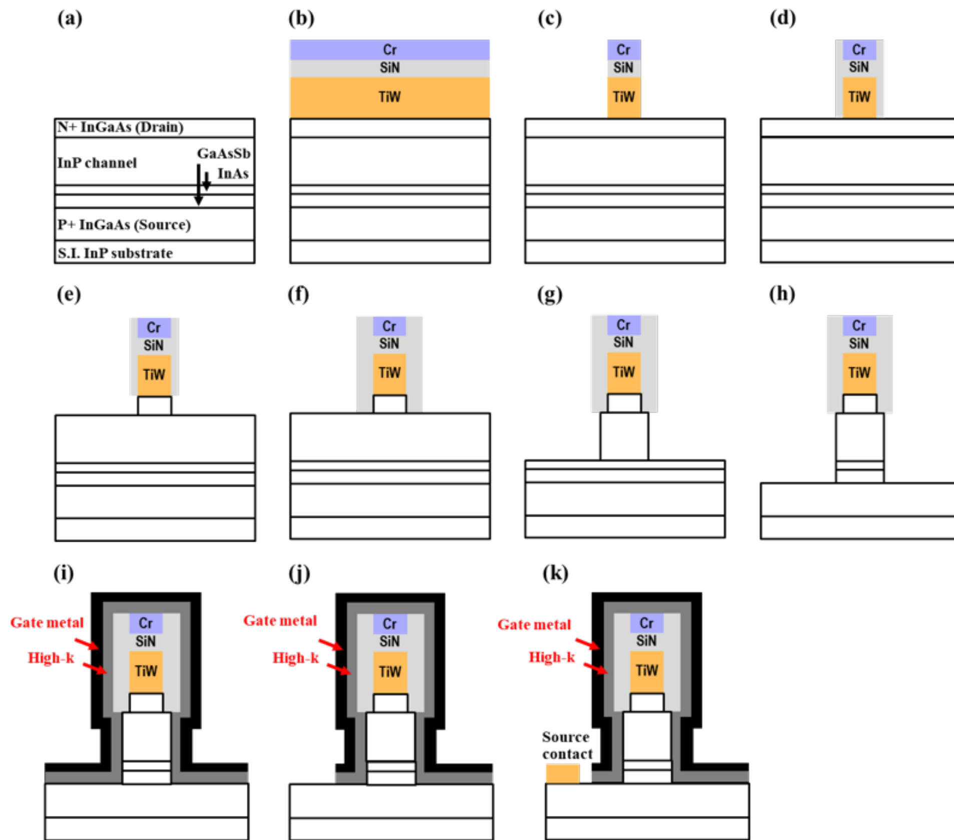


Figure 5.12: Front end process flow of anchor-shape vertical FETs.

isolation along with C cut (in fig. 5.11 (b)) is shown in fig. 5.14 (a). Back-end wiring starts from patterning and dry-etch of Cr layer on top of two metal pads (fig. 5.14 (b)). After that, HF-resistive SiN_x spacer is patterned and dry etched to expose the top of TiW drain contact (fig. 5.14 (c)). The dotted area in fig. 5.14 (b) and (c) represent the undercut region, which is nothing but air after wet etching. To passivate the surface and prevent a short circuit between TiN/Ru, Cr, and TiW, a 3 nm Al_2O_3 and a 40 nm SiN_x spacer are deposited and patterned. SiN_x is dry-etched and Al_2O_3 is wet etch in developer for > 5 minutes (fig. 5.14 (d)). Note that even though it is not specified here, the SiN_x spacer is also used as surface passivation in-field. Thus, SiN_x on top of source and gate metal is also patterned and removed simultaneously. At last, large area measurement

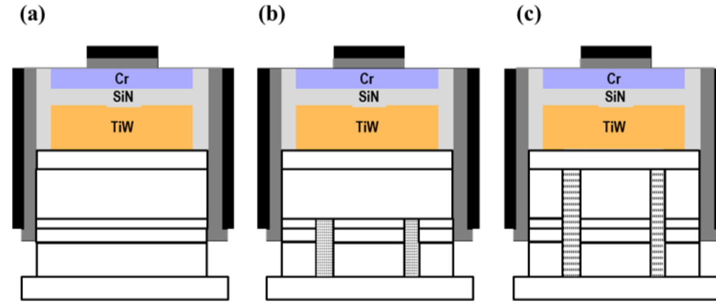


Figure 5.13: Isolation process of anchor-shape vertical FETs.

pads are patterned and lifted off. The schematic top view of the finalized device could be found in fig. 5.11 (a). Note that most of the process steps are shared between anchor-shape vertical MOSFETs and TFETs, except the number of cycles of digital etch prior to high-k growth. This process flow is the improved version after iterations by batches of the MOSFET run.

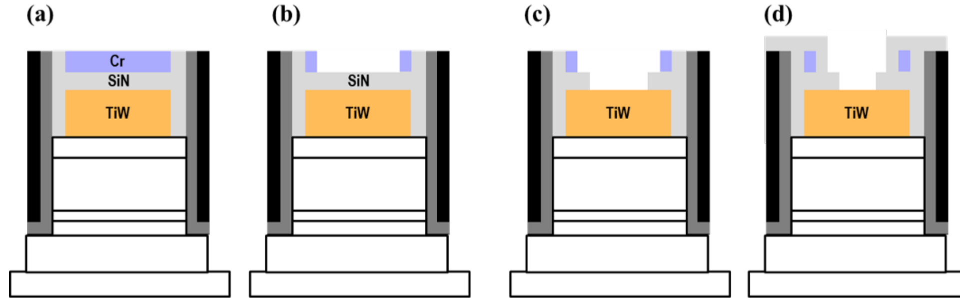


Figure 5.14: Back end wiring process flow of anchor-shape vertical FETs.

The key of anchor-shaped FETs is complete isolation between pads and fin. It is important to make sure InGaAs source and InP channel are clearly undercut. However, it is also impossible to tell if they are cleared from the top view when a thick opaque TiW layer presents. In the development of the isolation process, test structures are fabricated using thin SiN_x ($\approx$ 15 nm) as the etching hard mask. The epi structure for the test is the same as vertical MOSFETs. After SiN_x deposition, patterning (in the same anchor-shaped), and dry etch, InGaAs drain contact layer and InP channel are wet

etched with SiN_x mask and expose the source InGaAs layer. After that, samples are patterned using the same mesa pattern as in device fabrication. Isotropic InGaAs wet etch (H_3PO_4:H_2O_2:DI=1:1:25) following facet-dependent InP etch (HCl:H_3PO_4=1:4) are done. Fig. 5.15 shows the SEM image of the test structure after isolation, having fin width of 80 nm, fin length of 1 μm, and L_{sp} of 100 nm. L_{sp} is the spacing between pads and the edge of the mesa. High contrast white fin means that the semiconductor underneath is still present, while the low contrast region at two ends (indicated by red dotted circles) suggests the successful undercut of InGaAs. The bevels at the corner of the pads (indicated by white arrows) are generated from the InP facet-dependent etch. Additionally, this facet-dependent InP etch results in tapper-out features ($\approx 145°$) in InP along [011] (which is south-north in this figure), and causes the merge of InP between pads and mesa edge.

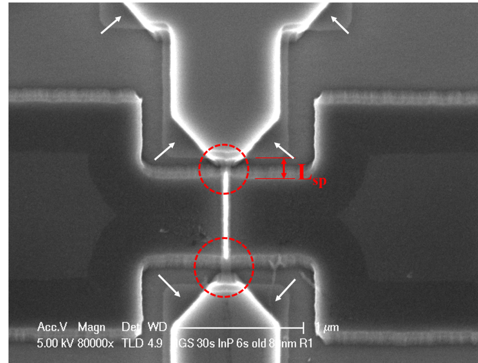


Figure 5.15: SEM image of the test structure after isolation, having fin width of 80 nm, fin length of 1 μm, and L_{sp} of 100 nm. L_{sp} is the spacing between pads and the edge of mesa.

Fig 5.16 shows the SEM images of different test structures with (a) fin length of 2, 3, and 5 μm at L_{sp} of 100, 150, 500 nm, respectively, and (b) the multi-finger design at a fin length of 2 μm and L_{sp} of 200 nm. The undercut in multi-finger structures works the same as in a single fin structure. It is concluded that a >150 nm L_{sp} is needed to separate the InP underpads from InP under mesa effectively.

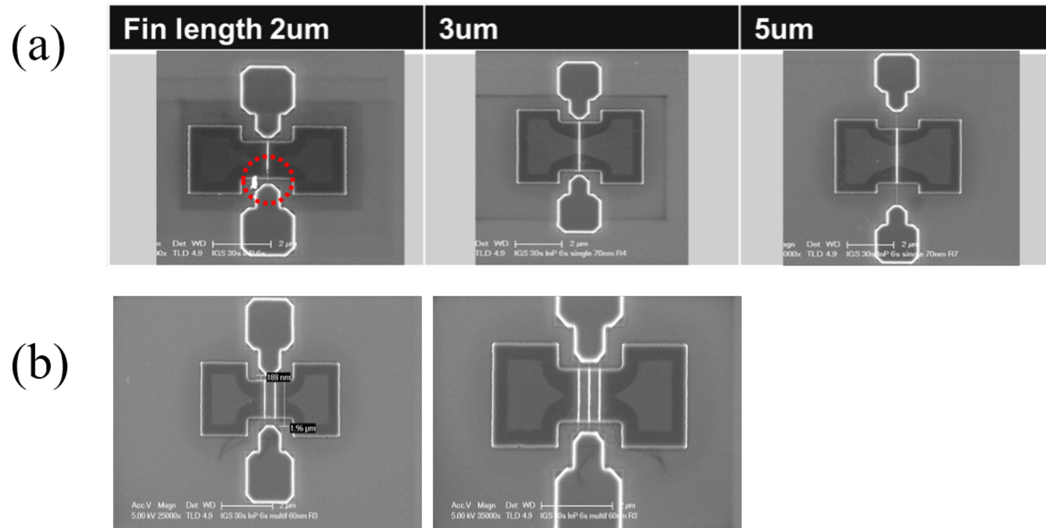


Figure 5.16: SEM images of different test structures with (a) fin length of 2, 3, and 5 μm at L_{sp} of 100, 150, 500 nm, respectively, and (b) the multi-finger design at a fin length of 2 μm and L_{sp} of 200 nm.

Even though fin isolation works well in test structures, the semiconductor etch happens to be problematic in device fabrication due to the presence of metal layers with much different work functions. Fig. 5.17 shows the cross-sectional TEM images of eighth batch anchor-shape vertical MOSFETs along with A cut as in fig. 5.11 (b). There is a void in the middle of the fin. Energy-dispersive X-ray analysis (EDX) in the inner plot shows that Pt fills the region where supposed to be TiW/Mo stack, and leaves with a void at the center. The well standing fin with Ru gate surrounding it indicates that the disappearance of TiW/Mo happens after finishing the front-end process. In principle, none of the chemical etchants/strippers/developers used in the whole process would etch TiW or Mo. After some experimental tests and detailed analysis done step by step on the fabrication process, it is concluded that TiW/Mo is fully etched away in developer during the surface passivation Al_2O_3 layer removal because of electrical chemical reactions (so-called Galvanic corrosion). The galvanic reaction takes place when metals with particular work function differences are electrically connected. In other words, it could

happen when the sample is put in an electrolyte (acid or base solution) with metals (Mo, Tiw, Ru, and Al_2O_3) exposure. In fig. 5.17, the Pt fill of the fin takes place during Pt protection layer deposition in focused ion beam (FIB) in the preparation of TEM.

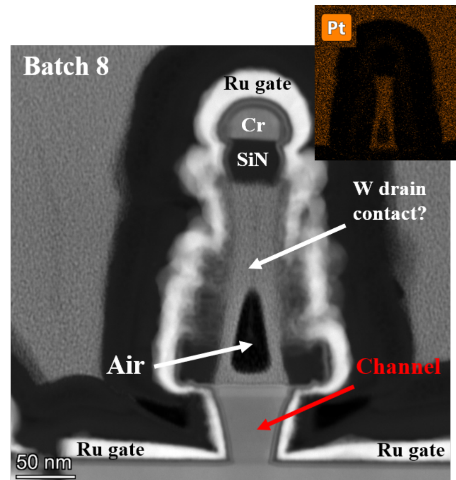


Figure 5.17: Cross-sectional TEM images of 8th batch anchor-shape vertical MOSFETs along A cut (according to fig. 5.11 (b)). The inner plot shows the EDX analysis of elemental Pt.

Fig. 5.18 (a) shows the cross-sectional TEM image of 8th batch anchor-shape vertical MOSFETs along C cut (in fig. 5.11 (b)) focusing on the side. Red arrow points to the void, where it should have been a $\approx$ 200 nm TiW/Mo stack. Because TiW/Mo is etched, Ti/Au (measurement pads) is in direct contact with InGaAs, and causes huge drain resistance. In addition, it is observed that the occurrence of galvanic effect could be in developer or in InGaAs or InP wet etchants. 2 ways are tried to overcome this problem. In the fabrication of ninth and tenth batch anchor-shape MOSFETs, the order of numerous processing steps is changed such that Al_2O_3 (surface passivation layer)is etched before TiW is exposed. Moreover, the developing time of final lithography is kept short to minimize the exposure of TiW and Ru in the developer. Another way to protect metal is by depositing dielectric layers (SiN_x). The top of TiW and Ru surfaces are covered by SiN_x almost all the time until measurement pad lift-off. Both of the potential

solutions are incorporated in the fabrication. Nevertheless, ninth and tenth batch vertical MOSFETs still catch up with galvanic corrosion with those carefully designed processing orders. Fig. 5.18 (c) and (d) shows the cross-sectional SEM images of the ninth and tenth batch of MOSFETs along A and C cut as in fig. 5.11 (b). In 10th batch MOSFETs, galvanic reaction attacks not only TiW, but also InGaAs and InP. As indicated by the red arrow in (c), the etch of TiW and semiconductor layers are originated from the edges of pads, which is different from what is commonly seen in (a) and (b). The etching time of the InP channel is increased by ≈ 8 seconds in tenth batch MOSFETs, which gives rise to the undercut of the InP channel and leads to some exposure of InGaAs drain contact layer. Comparing (c) with (a), it is clear that the InGaAs drain layer (which should be in contact with TiW/Mo) is fully etched away in (c) during isolation. The clear of InGaAs drain exposes the bottom of Mo at the edges of pads. The exposure of Mo and Ru in an electrolyte (InGaAs and InP etchants) again results in galvanic corrosion.

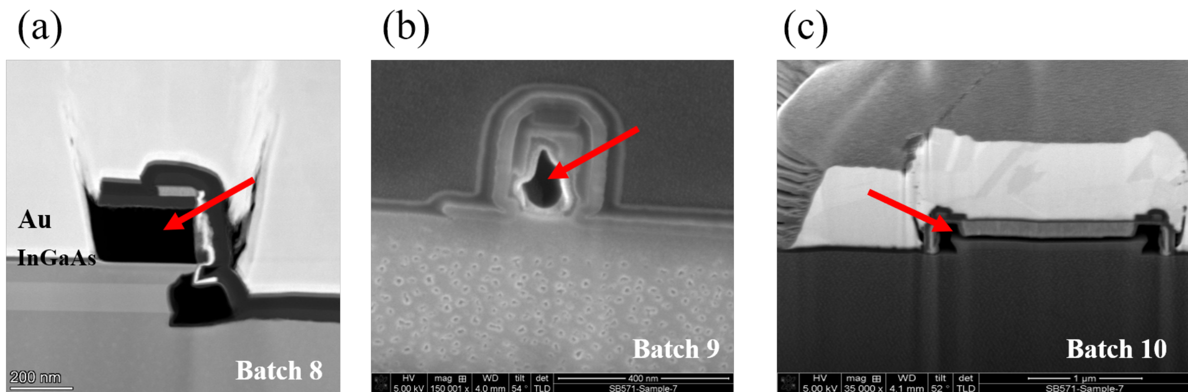


Figure 5.18: Cross-sectional TEM images of (a) 8th, (b) 9th, and (c) 10th batch anchor-shaped vertical MOSFETs along C cut, A cut and C cut, respectively (acoording to fig. 5.11 (b)).

After 3 - 4 batches of trials/iterations of anchor-shaped MOSFETs aiming to avoid the galvanic reaction, it is concluded that the process modules of anchor-shaped FET fabrication are not compatible with each other. The process is far from being robust,

and new issues come up when modules are modified. Therefore, the anchor-shape vertical structure is abandoned, and the top-down planarization process comes up.

5.2.3 Gen. 3–Top Down Planarization Process

Vertical planarization FETs have measurement pads connected to metal contacts using the planarization process, instead of the anchor structure. In most vertical TFET structures (as discussed in 5.1), two times planarization is usually employed. The first planarization is applied to etch the wrapped-around metal gates surrounding fins/nanowires. The second planarization serves as the dielectric spacer that insulates drain contact/measurement pads from gate metal. The vertical planarization process in this study uses one-time planarization for conformal gate etching. The insulation between gate/drain metals is done by thick SiN_x sidewalls (> 80 nm). Fig. 5.19 shows the top view schematic structure of planarization FETs. The yellow area indicates measurement pads. The following discussion on process flow will be cross-sectional cut along the red dash line.

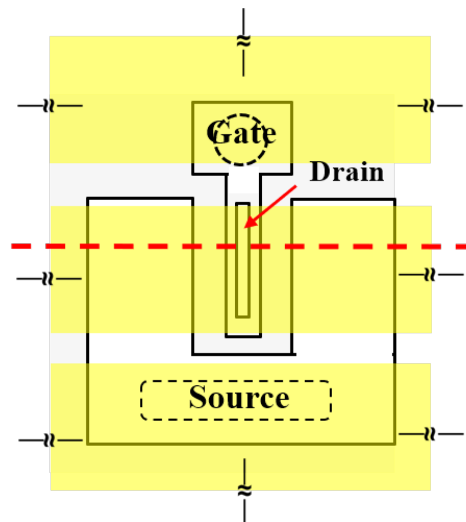


Figure 5.19: Top view schematic structure of vertical planarization FETs. The yellow area indicates measurement pads.

Fig. 5.20 shows front end process flow for planarization FETs. Top down planariza-

tion FETs have a similar front end process with anchor-shape structure. Fabrication starts from blanket depositing drain contact Mo/TiW (20/500 nm), SiO_2/SiN_x (80/40 nm), and Cr (30 nm) (fig. 5.20 (b)). Drain stack is further patterned and dry etched (fig. 5.20 (c)). Cr layer is lifted-off by > 4 minutes dip in BHF (fig. 5.20 (d)). Note that BHF dip not only removes SiO_2, SiN_x, and Cr layers, but also cleans up the redeposited polymer on TiW sidewalls generated in TiW dry etch. SiN_x dual sidewall formation, InGaAs and InP etch, high-k and metal gate deposition are idential as in anchor-shape structure. First $\approx$ 15 - 20 nm SiN_x sidewalls are formed (fig. 5.20 (e)) and 10 nm InGaAs drain contact layer is wet etched (fig. 5.20 (f)). To expose InP vertical channel, second $\approx$ 10 - 15 nm SiN_x sidewalls are formed (fig. 5.20 (g)), and InP is wet etched (fig. 5.20 (h)). Prior to high-k deposition, a 3 - 10 cycles of digital etch is done to align the following TiN/Ru gate at source/channel junction (fig. 5.20 (h)), where digital etch includes a 10 minutes UV ozone exposure following by a > 3 minutes dip in HCl:DI=1:10. Gate formation is identical to the fabrication of TiN/Ru gate planar MOSFETs (could be found in chapter 4). ALD in-situ nitrogen plasma clean, ZrO_2 and TiN/Ru gate deposition, and pre and post metal H_2 annealing are done (fig. 5.20 (j)), and are patterend and etched (fig. 5.20 (k)). After that, Source metal contact Ti/Pd/Au (15/200/65 nm for MOSFETs and Pd/Ti/Pd/Au (5/10/15/70 nm) for TFETs are lifted-off (fig. 5.20 (l)).

Before starting the back-end wiring process, gate and source metal posts Ti/Pd/Au (15/20/510 nm) are deposited in one go on their metal contacts. The feature and location of metal posts are indicated as black dash circle/rectangle in fig. 5.19. The thickness of the Au layer may change with different epi designs, and it depends on the height difference between the top of Ru and TiW.

Back-end wiring process flow for vertical planarization FETs is shown in fig. 5.21. A 30 nm SiN_x passivation layer is blanket deposited to protect the delicate fins and metal posts prior to planarization with spin-on dielectric benzocyclobutene (BCB). BCB

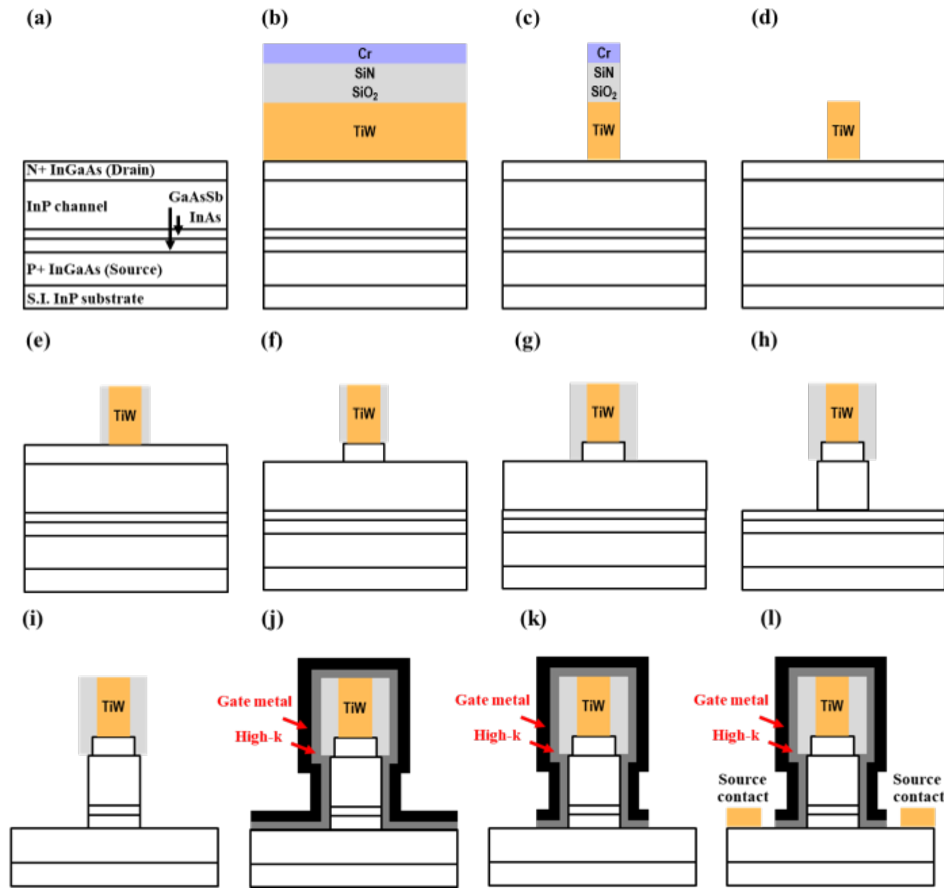


Figure 5.20: Front-end process flow for vertical top-down planarization FETs.

is spinned and baked (fig. 5.21 (a)), and ashed back to expose fins and metal posts (fig. 5.21 (b)). Note that the SiN_x passivation layer that surrounds the features is also removed by BCB ash. After that, the exposed Ru layer is wet/dry-etched. Thin TiN layer is further oxidized in O_2 plasma exposure and the TiN/high-k is removed in BHF dip (fig. 5.21 (c)). A thick SiN_x layer ($>$ 120 nm in-field and $>$ 60 nm sidewalls) is then deposited, patterned and dry etched (fig. 5.21 (d)). At last, Ti/Au (15/10000 nm) measurement pads is lifted-off to complete the fabrication (fig. 5.21 (e)). The SiN_x layer improves the adhesion between BCB and the following metal pads and is employed as spacers on TiW sidewalls. With thick enough SiN_x sidewalls (with thickness $>$ the thickness of TiN/Ru gates + dual SiN_x sidewalls that form in front end process), the further deposited Ti/Au

pads would not directly contact on sidewall TiN/Ru. By employing this method, two times planarization process is successfully reduced to 1 time in the fabrication of vertical FETs.

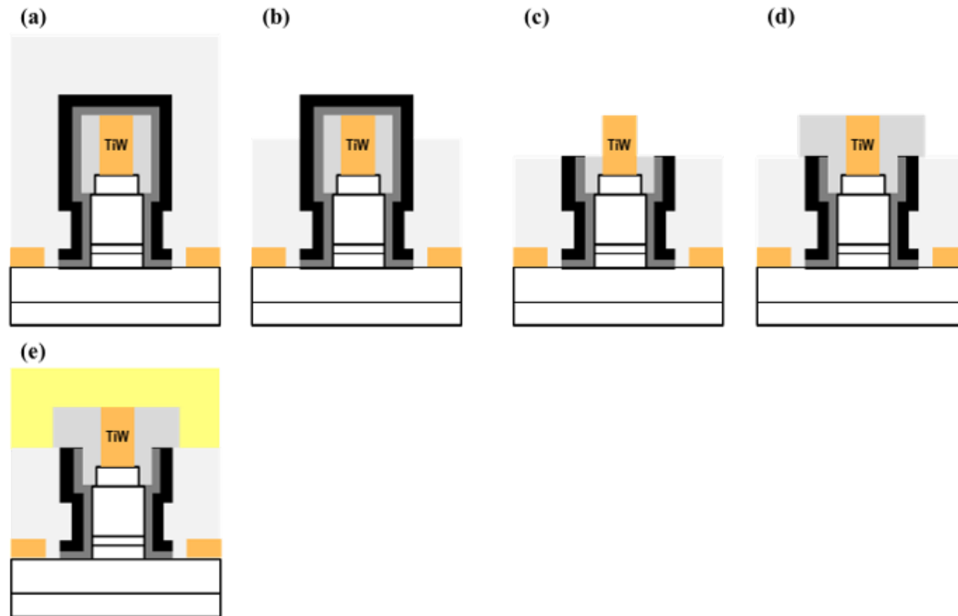


Figure 5.21: Back-end wiring process flow for vertical top-down planarization FETs.

5.3 Fabrication of Vertical Planarization FETs

This section discusses the design of each process module. Process developments, limitations, and issues will be specified, and potential solutions will be proposed. Detailed process parameters for the fabrication of vertical planarization FETs could be found in the appendix.

5.3.1 Drain Metel Formation

TiW/Mo drain stack formation is inherited from the HBT emitter process, which Yihao Fang develops at UCSB. Detailed discussion on the development of high aspect

ratio sub 100 nm TiW dry etch could be found in his thesis.

E-beam deposited Mo has low contact resistivity to N^+ InGaAs, while DC co-sputtered TiW (with a low pressure of 5 mtorr and high power of 300 watts) has low sheet resistance. Refractory TiW/Mo drain stack enables a low drain resistance in highly scaled MOSFETs and TFETs. A 500/20 nm TiW/Mo stack is used in the vertical planarization process. The thick TiW is designed to facilitate contacting the drain in back end wiring, but the high aspect ratio TiW dry etch limits the scaling of body thickness.

A 90-degree vertical TiW contact is preferred for the scaling of the body. However, the heat accumulation on the substrate in high power (source/bias power: 600/200 watt) SF_6/Ar plasma for drain stack dry etch results in TiW undercut, and forms a thin neck. With the lowest chuck/chiller temperature of $\approx$ 10 - 15°C in the ICP system at UCSB, pause in TiW dry etching is needed to prevent heat accumulation on the substrate. TiW etching in the recent works at UCSB is paused every 40 seconds, and samples are taken out for around 10 minutes after each 40 second etch. Fig. 5.22 shows the 30°tilted SEM images of TiW/Mo stack after dry etch and Cr lift-off. The sample size is roughly 1.5"×1.5". In the center die of a 60 nm body design (a), TiW width turns out to be $\approx$ 40 nm on the head, $\approx$ 25 nm at the neck, and $\approx$ 80 nm at foot. Additionally, in the edge die of an 80 nm body design (b), TiW width turns out to be $\approx$ 50 nm on the head, $\approx$ 15 nm at the neck, and $\approx$ 60 nm at foot. It is clear that the edge of samples has an undercut much stronger than at the center, owing to a strong loading effect in TiW etch. Indeed, a higher etch rate is always observed at the edge of a $>$ 1"×1" sample, and a greater undercut is expected. With that in mind, the body width is usually designed/ranged in between 60 - 130 nm. A designed feature with $<$ 60 nm would either have TiW fall/break or be fully undercut. This phenomenon reduces the yield in vertical planarization FETs.

It is also observed that the amount of undercut depends on the substrate. Table 5.2 shows the dimension of TiW foot after TiW etch using different substrates. TFET epi

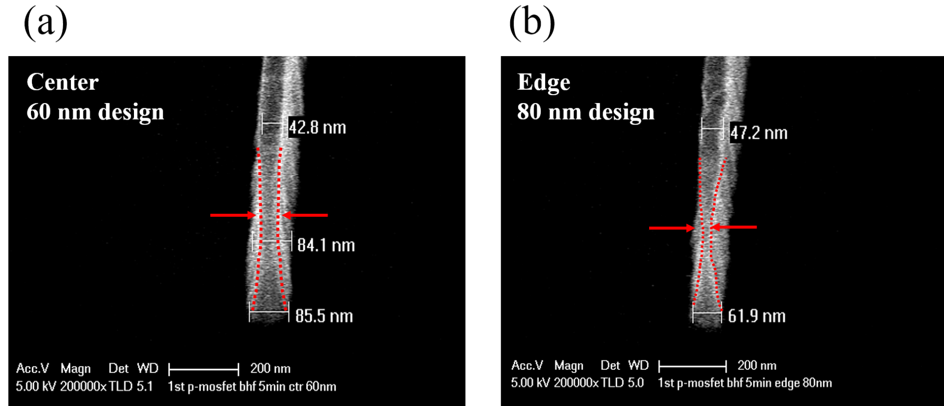


Figure 5.22: 30°tilted SEM images of TiW/Mo stack after dry etch and hard mask lifting-off. (a) 60 nm fin width design at the center die, and (b) 80nm fin width design at the edge die of a 1.5"×1.5" sample.

on a thicker substrate has increased feature sizes. This suggests TiW dry etch might require to be optimized on different epi or substrates before device fabrication, among which substrate thickness and thermal conductivity of layers could both affect TiW dry etch. A lower chuck temperature or a cryogenic ICP is recommended to achieve a high aspect ratio vertical TiW etch.

	60 nm design	130 nm design
MOSFET epi (substrate 300 μm)	80 nm	150 nm
TFET epi (substrate 500 μm)	100 nm	190 nm

Table 5.2: The dimension of TiW foot after TiW dry etch using different substrates. The measurements are done at the center die of sample pieces.

After dry etching of drain stack, Cr hard mask is lifted-off by the underneath SiO_2/SiN_x layers in a long (> 4 minutes) BHF dip. The BHF etch time should be long enough, such that SiO_2/SiN_x could be entirely undercut on devices with any fin width. During BHF etch, a vigorous stir of the wafer holder laterally and vertically is needed to prevent the Cr film from attaching back onto the sample surface.

5.3.2 Dual SiN_x Sidewall Process

SiN_x sidewalls are formed by blanket deposition of SiN_x by PECVD, following a low power anisotropic SiN_x dry etch by ICP. 1st SiN_x sidewalls are formed before InGaAs drain contact layer wet etch, and 2nd SiN_x sidewalls are formed before InP channel wet etch.

SiN_x sidewalls have three purposes in this process. First, it was found that galvanic corrosion could sometimes happen if the corner of the Mo/InGaAs interface is exposed during N^+ InGaAs wet etch and InP channel wet etch. Thus, the thickness of 1st sidewalls should be at least thicker than the lateral undercut of InGaAs in wet isotropic InGaAs etch. Second, if there is no 2nd SiN_x sidewall covering the edge of the thin InGaAs drain contact layer, high-k and metal gates would directly contact the sides of InGaAs, which may cause some degree of gate induced drain leakage (GIDL). Note that the 5 nm higher bandgap N^+ InP layer beneath N^+ InGaAs could also suppress GIDL. In addition, gate metal overlapping on drain semiconductor layers could deplete the electrons near to the surface in the drain. The phenomenon would only be seen when body thickness is highly scaled. With the SiN_x sidewall spacers, the wet etching process is more robust, and fewer non-idealities need to be considered in analyzing electrical data of devices.

Most importantly, high-k and metal gate deposition directly on the sidewalls of TiW result in the non-negligible gate to drain leakage. It is experimentally confirmed that the nucleation of high-k on metals (TiW/Mo or Au) is problematic. Compared to the high-k grown on semiconductors, a thinner/discontinuous high-k on TiW/Mo sidewalls would result from bad nucleation. Therefore, SiN_x sidewall spacers need to be inserted between the gate stack and TiW/Mo to prevent the gate/drain leakage current.

5.3.3 Semiconductor Wet Etching

InGaAs drain contact layer and InP channel layer are selectively wet etched. InGaAs etch is done using H_3PO_4:H_2O_2:DI=1:1:25 at an etch rate of $\approx$ 4 nm/sec at R.T., and at minimum time of 7 seconds. This etch is fully isotropic, so a lateral undercut is expected. Typically, a 20 nm undercut (10 nm from each side) in InGaAs is assumed in a 10 nm InGaAs layer etching. The etching of InGaAs stops at the InP surface. InP etching is done using HCl:H_3PO_4=1:4 at an etch rate of $\approx$ 4 nm/sec at R.T., and for a minimum of 9 seconds. This etching is facet-dependent, and the resulted facets/sidewall angles are shown in fig. 5.23 [103, 104]. (100) InP substrate is used and devices (fins) orient along $[01\bar{1}]$. This InP wet etch gives almost vertical sidewalls along $[01\bar{1}]$ (b), and tapper-out sidewalls ($\approx$ 35°) along [011] following (211) group (c), respectively [103]. Note that there exists tapper-out ($\approx$ 55°) portion at the downside of vertical sidewalls along $[01\bar{1}]$ following $(11\bar{1})$ group [103]. These feet could be totally removed by a longer time over etch.

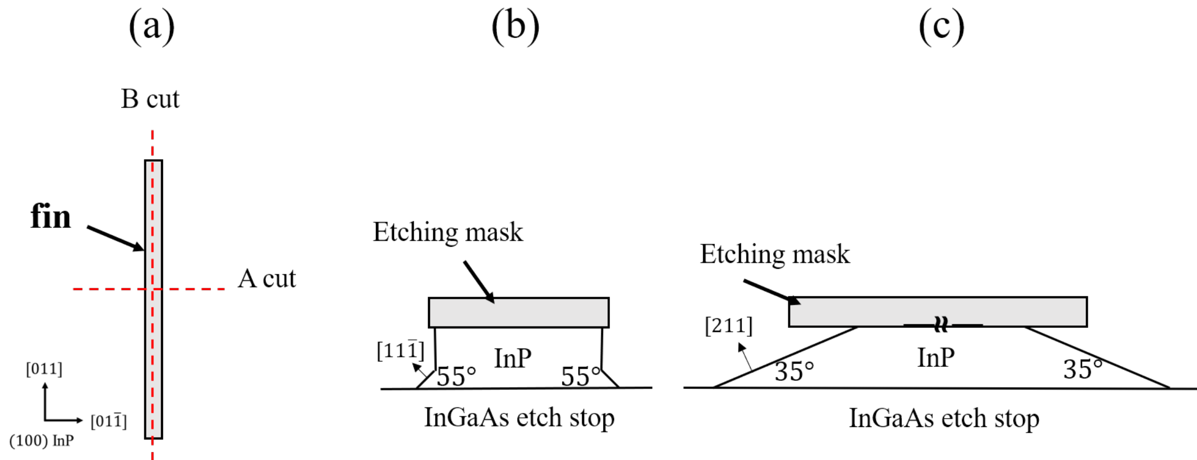


Figure 5.23: (a) Schematic diagram of device (fin) orientation and the resulted facets and sidewall angles after HCl:H_3PO_4=1:4 InP wet etch along A cut ($[01\bar{1}]$) (b), and along B cut ([011])

A low temperature (at ≈ 6°C) InP etch using the same etchant is also investigated. Fig. 5.24 (a) shows the cross-sectional SEM image of fins after 10 nm InGaAs etch following a 2 minutes of 30 nm InP low temperature etch. Note that the etch mask for (a) and (b) is HSQ at a designed fin width of 60 nm. InP foot region is ≈ 20 nm after InGaAs and InP etch. After 3 cycles of digital etch, it is reduced to ≈ 12 nm, as shown in (b). The further shrinkage of the dimension results in the collapse of the fin due to large surface tension. In addition, (c) shows the cross-sectional SEM image of fins with identical semiconductor layer design and wet etches, but a Cr/SiO_2/SiN_x/TiW/Mo (30/80/40/130/20 nm) hard mask. Note that no SiN_x sidewall is added, and Cr/SiO_2/SiN_x is left on the sample without lifted-off in the test. Fin width in (c) is designed at 60 nm. After drain stack dry etch, the bottom of the stack has a dimension of ≈ 70nm. After InGaAs and InP wet etch, the dimension of the foot of the InP layer is ≈ 30 nm. Further shrink of dimension in (c) below 25 nm results in collapse of fin again. This suggests that the edge roughness of TiW after drain stack dry etching has roughly 7 nm thickness variation. The amount of edge roughness caused by TiW dry etch is affected by TiW film quality, TiW total thickness, heat accumulation, and loading effect during etching. Rougher (> 20 nm thickness variation) TiW edge is usually observed in a thick TiW drain. With a good enough control of semiconductor wet etches, as shown above, the limiting factor of minimum body thickness in vertical FETs becomes TiW etch undercut and its edge roughness.

5.3.4 Gate and Source Metallization

Cycles of digital etch (10 minutes of UZ ozone exposure following >3 minutes of HCl: DI=1:10 dip) are done prior to high-k deposition to etch through tunneling junctions (InAs/GaAsSb). The number of the cycle depends on the digital etch rate on InAs and

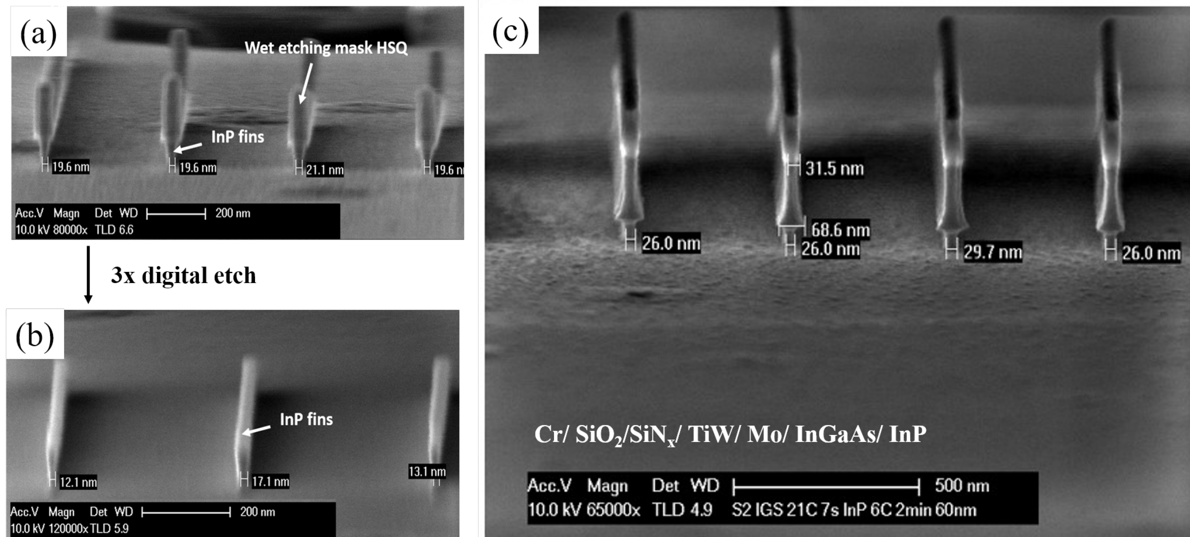


Figure 5.24: Cross-sectional SEM image of fins after 10 nm InGaAs etch following a 2 minutes of 30 nm InP low temperature etch using (a) 30 nm HSQ mask, and (c) Cr/SiO_2/SiN_x/TiW/Mo (30/80/40/130/20 nm) drain stack as hard masks. (b) shows the structures in (a) with a reduced dimension after 3 cycles of digital etch.

GaAsSb and high-k thickness. To ensure a good gate/source alignment in TFETs, the bottom of the gate metal should align to the interface between InAs and GaAsSb. The best electrostatics and highest tunneling need to be optimized by changing the number of digital etch cycles. This digital etch also serves as the surface clean before high-k deposition.

High-k process starts from 9 cycles of alternating N_2-plasma and TMAl dosing ($\approx$ 1 nm AlO_xN_y) at 300°C. This step passivates the surface, and the deposited AlO_xN_y enhances the adhesion of ZrO_2 on semiconductor layers. Next, 40 cycles of ZrO_2 ($\approx$ 2.5 nm) is deposited using H_2O and TEMAZ dosing at 300°C. 30 minutes of ALD in-situ H_2 annealing is then carried out at 350°C at a H_2 flow of 30 sccm and pressure of 220 mtorr to passivate interface dangling bonds. Prior to Ru growth, a 35 cycles TiN ($\approx$ 2 nm) is grown at 300°C using TDMAT together with an N_2- and H_2-plasma (400 watts remote inductively coupled plasma (ICP) power). 675 cycles of ALD Ru ($\approx$ 25 nm) is then deposited using EBCHDRu, O_2 and H_2 cycles at 250°C as discussed in section 4.4.

30 minutes of post gate metal H_2 annealing is done to recover the plasma damage on the high-k/InP interface that is generated during TiN deposition. The annealing condiction is identical to pre-gate metal anneal.

Gate is further patterned and etched. Ru is dry-etched in Cl_2/O_2 plasma at source/bias power of 500/50 watt at 2.5 pa, at a 1.33 nm/sec rate, using photoresist as hard mask. After Ru etch, 1 minute of O_2 plasma exposure at a power of 100 watts at 300 mtorr is done to oxidize the thin TiN. A 1-minute BHF dip is then carried out to etch the oxidized TiN and high-k dielectrics. Resist is stripped after finishing gate etch. 2 cycles of digital etch is used to clean up the InGaAs surface for further source metallization.

Prior to source contact deposition, InGaAs surface is cleaned by a 1-minute dip in HCl: DI=1:10. Source metallization lifts-off E-beam deposited Ti/Pd/Au (15/20/65 nm) for MOSFETs and Pd/Ti/Pd/Au (15/10/15/70 nm) for TFETs. These metal stacks have been shown decent low contact resistivity to n- and p- InGaAs and diffuse < 5 nm into InGaAs [105, 106].

After gate and source metallization, E-beam deposited Ti/Pd/Au is lifted off on gate and source contacts in one go. As shown before, the thickness of Ti and Pd are 15 and 20 nm, while the thickness of Au would vary depending on the height difference between the top of TiW, and the top of gate and source contacts. Note that there is no special surface treatment before metal post-deposition. Fig. 5.25 shows a tilted SEM image of vertical FET after finishing the front-end process.

5.3.5 Back End– Planarization and Wiring

The back-end process starts from a 30 nm SiN_x surface passivation. Planarization uses a spin-on dielectric cyclotene 3022-46 (BCB). A 3 μm BCB is spun coated at 2000 rpm for 45 seconds. Right after that, a 1-hour bake under N_2 at 250°C is done. BCB is ashed

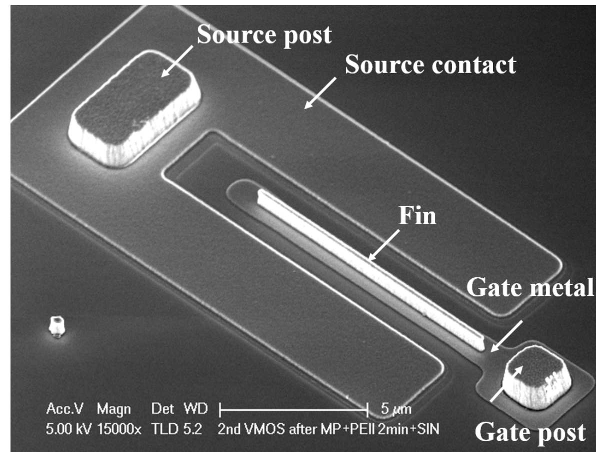


Figure 5.25: Tilted SEM image of vertical FET after finishing front end process

back using CF_4/O_2 plasma at a power of 1000 watts at 40 pa. BCB ashing shows the best uniformity on round/square samples, with center and edge BCB thickness difference of ≈ 200 nm. On triangular sample pieces, the thickness difference in the BCB layer after ashing could be as large as ≈ 400 nm, which dramatically reduces yield. Fig. 5.26 shows the 30°tilted SEM images of the device having the top of fin and gate/source posts exposed after BCB ashing. Red arrows indicate electron charging on non-conductive BCB, and white arrows specify the micro-trenching at the edge of fin and metal posts that are generated due to a higher etch rate of BCB around features. A clear texture of Au could be seen on the gate post in (b), suggesting it is exposed. Nevertheless, it is hard to tell how tall fin/metal posts are exposed by SEM images because of the charging effect, micro-trenching, and penetration of electrons through the BCB layer. Step profilometer and/or atomic force microscope (AFM) is most commonly used to verify their exposure.

The sticking out of fin from planarized BCB is targeted to be > 75 nm. Once the exposure of TiW, gate, and source posts at the center of the sample is verified, Ru that surrounds TiW is then dry or wet etched. Ru wet etching is done using Transcene RU-44 (ceric ammonium nitrate/ nitric acid) at a rate of ≈ 1 nm/sec at R.T. This solution

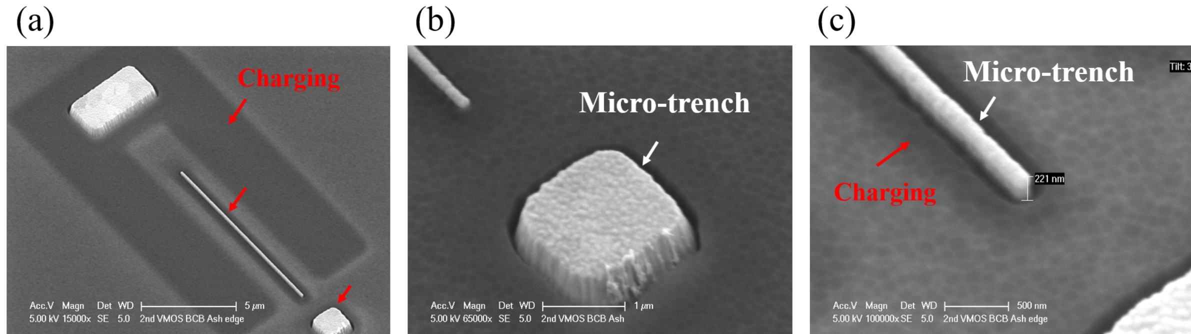


Figure 5.26: The 30°tilted SEM images of the device having the top of fin and gate/source posts expose after BCB ashing. (a) shows the panoramic view, while (b) focuses on the gate post and (c) zooms into TiW fin.

etches InGaAs at speed much faster than Ru but does not attack SiN_x. This suggests Ru wet etch could only be employed in the back end Ru etch process of vertical FETs. In addition, it is observed that the etch rate varies with environmental temperature variation between daytime and nighttime (± 5°C). Moreover, the etch rate would decrease if thin RuO_2 is present on the Ru surface. In the back-end planarization process, BCB is ashed back by CF_4/O_2 plasma. This plasma does not attack Ru, but the oxygen component oxides Ru surface and forms a thin RuO_2 layer on top of Ru. A longer wet etch time is needed to clear 25 nm Ru in the back end Ru etch, comparing to the etch on an un-processed Ru film. Therefore, the etch rate variation given by temperature and/or oxidized Ru surface needs to be considered, and it is better to test the etch every time before back end Ru wet etch, including those variables. In the back end wiring process of vertical MOSFETs, a $\approx$ 15 seconds dip in BHF is done before Ru wet etch to clean up the SiN_x passivation layer and also the surface RuO_2.

Ru dry etching uses Cl_2/O_2 (5/49.5 sccm) plasma at source/bias power of 500/50 watt at 2.5 pa, and at a rate of 1.33 nm/sec. Ru dry etching is vertical, stable, and robust. Because of its anisotropic characteristic, a longer etch time is needed to clear the Ru on the sidewalls. However, this Cl_2/O_2 plasma would attack the BCB around

features, and result in a higher micro-trenching effect. Fig. 5.27 shows the cross-sectional SEM images showing micro-trenching effect in BCB (a) without, and (b) with 1 minutes of Cl_2/O_2 plasma exposure. Note that (a) is an HBT, having a different layer structure and front-end process from vertical FETs. It is found that Cl_2/O_2 plasma barely etches BCB in-field, but the result in (b) indicates it does attack it near to devices and increase micro-trenching. Full removal of BCB around devices would lead to drain to gate short circuit once measurement pads are deposited. Therefore, the back-end Ru dry etching requires to be precisely controlled.

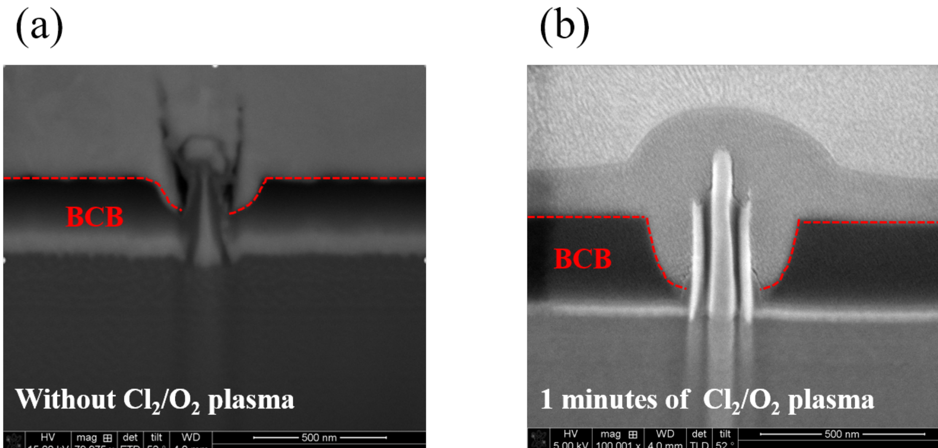


Figure 5.27: Cross-sectional SEM images showing micro-trenching effect in BCB (a) without, and (b) with 1 minutes of Cl_2/O_2 plasma exposure.

After Ru etch, TiN and high-k layers are removed using the same method as shown in gate metallization. Note that the 1-minute BHF dip for TiN/high-k removal also etches the sidewall SiN_x spacers. In fabricating transistors, vertical MOSFETs and TFETs use back end Ru wet etch and dry etch processes, respectively. Fig. 5.28 shows the tilted SEM images of vertical FETs (a) after BCB planarization, (b) after Ru dry etch following TiN/high-k/SiN_x sidewall removal, and (c) after Ru wet etch following TiN/high-k/SiN_x sidewall removal. The back-end Ru wet and dry etching processes show structural problems and cause a poor yield in vertical MOSFETs and TFETs. A detailed explanation

will be given out later.

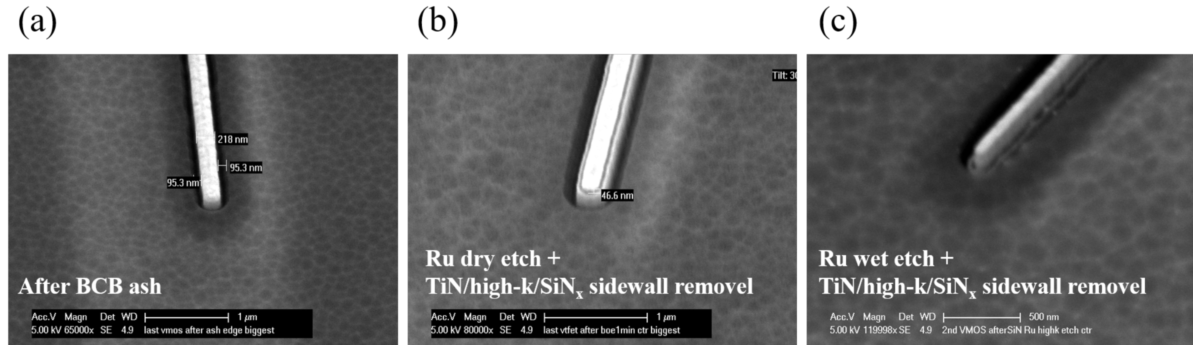


Figure 5.28: 30°tilted SEM images of vertical FETs (a) after BCB planarization, (b) after Ru dry etch following TiN/high-k/SiN_x sidewall removal, and (c) after Ru wet etch following TiN/high-k/SiN_x sidewall removal.

The solution to the current problematic back end Ru wet and dry etching is a conformal but dry Ru etching. Nakahara *et al.* reported a high Ru etching rate in ozone, and a comparably lower etch rate in radical oxygen at 100 - 150°C [107]. The volatile product that forms in Ru etch is RuO_4. Lee *et al.* pointed out oxygen radical is the main an effective reactant for RuO_2 etch [108]. RuO_4 and RuO_3 are the product in the reaction. Indeed, Ru is found to be etched in a plasma clean system (YES EcoClean) at above 100°C at UCSB. YES EcoClean is meant to remove organic residues/resist chemically. This system excites oxygen remotely, generates primarily neutral oxygen radicals. Oxygen radicals diffuse and are scattered across the sample surface. It is found that a 25 nm Ru layer is fully cleared/etched within 1 minute using a power of 0.7 kW at 100°C. Ru etch in EcoClean system is chemically (without ion damage), conformal and dry, and thus is promising as a replacement to the current problematic back-end Ru dry and wet etch.

5.4 Vertical MOSFETs and 3-HJ InGaAs/GaAsSb/InAs/InP TFETs

This section discusses the experimental results of InP channel MOSFETs and 3-HJ InGaAs/GaAsSb/InAs/InP TFETs using the top-down vertical planarization process. A high SS and a low on-current of this first demonstrated 3-HJ TFETs will be shown comparing to the simulated results. The inferior TFET performance will be explained, and the potential solution will be proposed.

5.4.1 Transfer and Output Characteristics of Vertical MOSFETs

Vertical MOSFETs have the epitaxial stack consisting of, from bottom to top, a Fe-doped S.I. (100) InP substrate, a 90 nm thick Si-doped N^+ $In_{0.53}Ga_{0.47}As$ (4×10^{19} cm^{-3}) source layer, 5 nm thick Si-doped N^+-InP (2×10^{19} cm^{-3}) layer, a 50 nm thick Zn-doped P-InP (1×10^{18} cm^{-3}) channel layer, a 5 nm thick Si-doped N^+-InP (2×10^{19} cm^{-3}) drain layer, and a 10 nm thick Si-doped N^+-$In_{0.53}Ga_{0.47}As$ (4×10^{19} cm^{-3}) contact layer. All layers were grown by MOCVD at 600°C.

The output and transfer characteristics of vertical MOSFETs with TiN/Ru gate are shown in fig. 5.29 (a) and (b). As can be seen, strong short-channel characteristics are observed. The low aspect ratio between gate length and fin width (50: 90 nm) results in poor gate electrostatics. Despite the higher bulk potential barrier in the P-InP channel with a doping concentration of 1×10^{18} cm^{-3}, high leakage current density in the thick InP is present. To overcome this issue, a higher P channel doping or a higher aspect ratio between L_g and body thickness (T_{body}) is needed. Nevertheless, peak g_m at V_{DS} = 0.6 V of the vertical MOSFETs measures 0.42 mS/μm, comparable with the L_g =

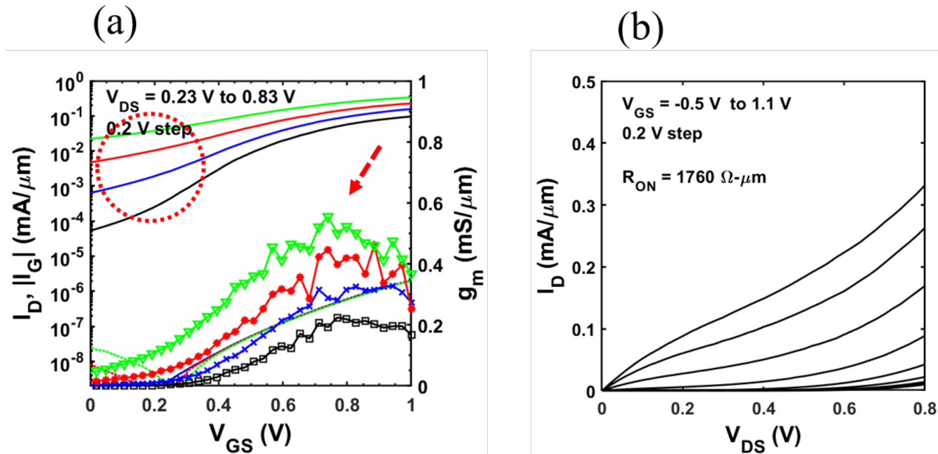


Figure 5.29: Transfer (a) and output (b) characteristics of vertical MOSFETs with TiN/Ru gate. Low ratio between gate length and fin width (50 nm: 90 nm) results in a poor gate control and shows a strong short channel behavior.

50 nm TiN/Ru gate planar MOSFET (as shown in chapter 4). This confirms ALD TiN/Ru/high-k gate's low surface trap density on non-planar InP device structures and proves the top-down vertical planarization process.

Fig. 5.30 (a) and (b) show the cross-sectional and top-view of a vertical MOSFET structure. The TEM image, as shown in (c), is the cross-sectional cut along with A and A' as indicated in fig. 5.30 (b). The image shows that the InP channel is covered by a uniform TiN/Ru film. The 250°C Ru film shows conformally underfilling of notches near to InP channel and almost constant thickness on the sidewalls and in the field, on top of N^+ InGaAs source. The TEM image shows an air gap between the BCB planarization material and the Ti/Au pad metal; this may be due to BCB contraction during the relatively high-temperature Ti/Au deposition [109]. Additionally, a slight recess in TiW at the top is observed in fig. 5.30 (c). This could be explained by a weak galvanic reaction that happens during Ru wet etching or resists development.

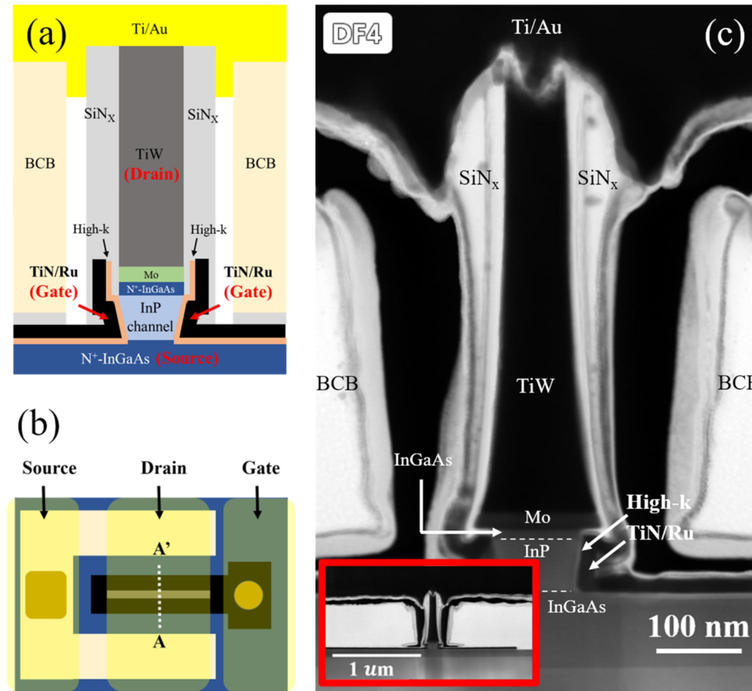


Figure 5.30: The schematic (a) cross-sectional and (b) top view vertical MOSFET structure, and (c) the cross-sectional STEM image of the vertical MOSFET with TiN/Ru gate cutting along A and A', as indicated in (b). The P-InP channel length is 50 nm in this study, while the TiW/Mo drain metal stack is as thick as 520 nm to facilitate contacting the drain. The inner plot in (c) is the image in a large field.

5.4.2 Bad Yield– Problematic Back End Wet Process

Vertical MOSFETs use the top-down planarization process with back end Ru wet etch. Devices show a bad yield across the half 2" sample piece. Most of them have gate open circuits, while some show high drain resistance or even a drain open circuit. The opened gate is due to a wet etch undercut that removes the SiN_x spacers, Ru, TiN, and high-k layers on the channel. Fig. 5.31 shows (a) to (d) shows the schematic cross-section diagrams explaining what happens during the back end wet processes. In the 15 seconds BHF dip prior to Ru wet etch (for SiN_x spacer and surface RuO_2 cleanup), BHF leaks into the BCB/SiN_x interface due to a bad adhesion of BCB (fig. 5.31 (a)). The BHF leak undercuts the SiN_x spacer surrounding fin, and exposes Ru surface without protection

(fig. 5.31 (b)). The further Ru wet etch removes all the uncovered Ru (fig. 5.31 (c)) and same to TiN and high-k etching (fig. 5.31 (d)). Fig. 5.31 (e) shows the optical microscope (OM) images of vertical MOSFETs before (left) and after (right) back end Ru wet etch. As indicated by the red arrow, the color of gate metal (closed to fin) fades after Ru wet etch, indicating Ru undercut is extended to the planar regime.

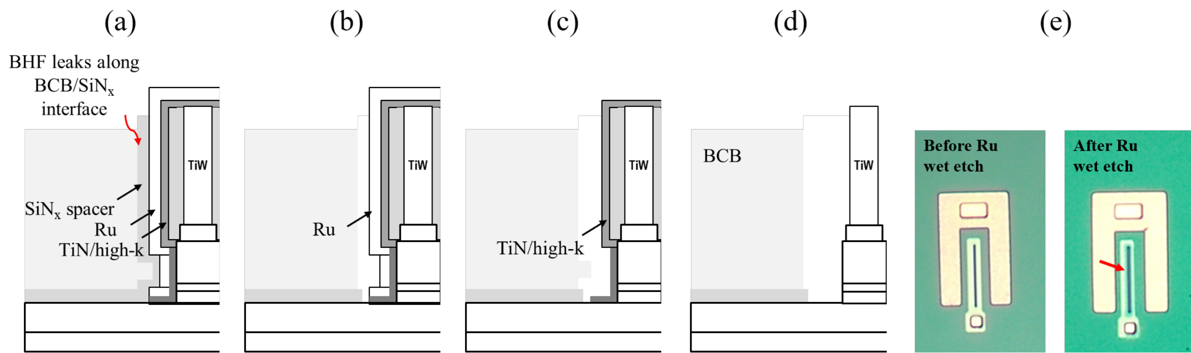


Figure 5.31: (a)-(d) Schematic cross-section diagrams explaining what happens during the back end wet etches. (e) OM images of vertical MOSFETs before (left) and after (right) back end Ru wet etch. As indicated by the red arrow, the color of gate metal (closed to fin) fades after Ru wet etch, indicating Ru undercut is extended to the planar regime.

The reasons for opened gate and/or high drain resistance/opened drain are also examined by STEM-EDX analysis. Fig. 5.32 and 5.33 shows the cross-sectional STEM-EDX mapping of elemental Ru, Ti, Si, In, P, W, Ga, and As on working MOSFETs (that measures the data in fig. 5.29) and opened gate MOSFETs, respectively. It is clear that elemental Ru and Ti signals on the channel are completely gone in opened gate MOSFETs. It is also observed that part of InGaAs drain contact layer in fig. 5.33 is etched. This could be explained by the leak of Ru etchant through SiN$_x$/InGaAs or SiN$_x$/InP interfaces and further attack InGaAs at a high etch rate.

Aside from the bad adhesion of BCB on SiN$_x$, the high pinhole density in BCB is another issue when integrating wet etch with the BCB planarization process. During the first back end BHF dip, BHF seeps through the pinholes in BCB in-field and reaches down

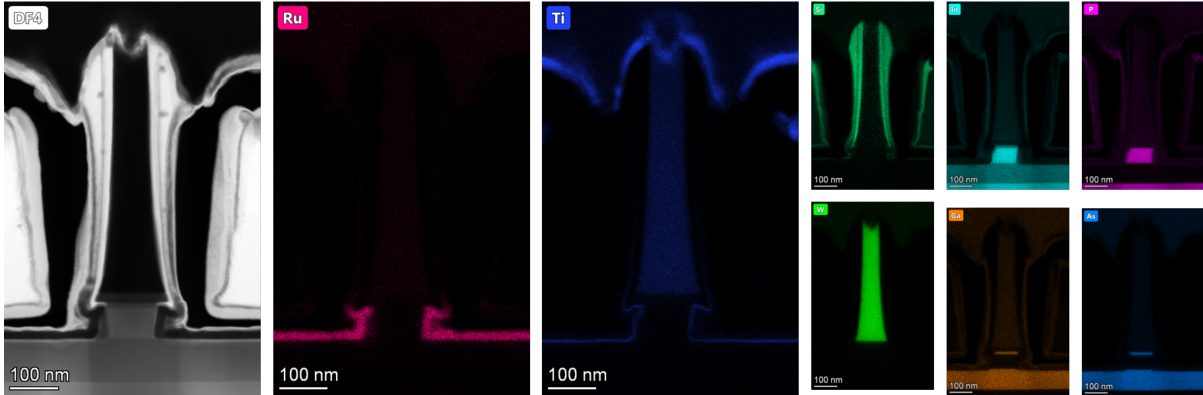


Figure 5.32: Cross-sectional STEM-EDX mapping of elemental Ru, Ti, Si, In, P, W, Ga, and As on working MOSFETs (that measures the data in fig. 5.29).

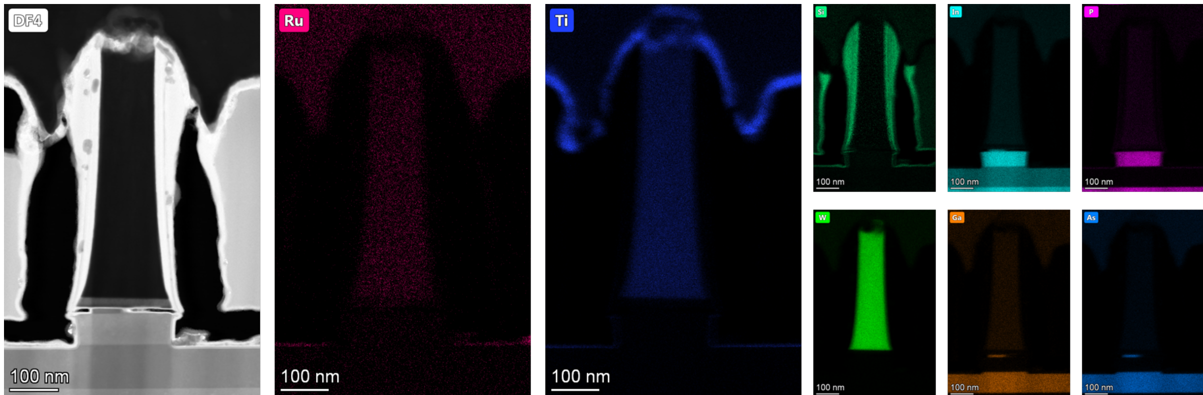


Figure 5.33: Cross-sectional STEM-EDX mapping of elemental Ru, Ti, Si, In, P, W, Ga, and As on opened gate MOSFETs with high drain resistance at the same time.

to the passivation SiN_x layer on the substrate. BHF removes the SiN_x spacer layer, and the following Ru etch seeps through the thin high-k, and attacks the underlying InGaAs source layer. The amount of attack on substrate across the sample piece depends on the pinhole density in BCB dielectric.

5.4.3 First Demonstration of 3-HJ InGaAs/GaAsSb/InAs/InP TFETs

Vertical 3-HJ TFETs have the epitaxial stack consisting of, from bottom to top, a Fe-doped S.I. (100) InP substrate, a 100 nm U.I.D. $In_{0.52}Al_{0.48}As$ back barrier layer, a 3 nm U.I.D. InP etch stop layer, a 75 nm C-doped P^{+}-$In_{0.53}Ga_{0.47}As$ (2×10^{19} cm^{-3}) source layer, a 25 nm C-doped P^{+}-$In_{0.53}Ga_{0.47}As$ (5×10^{19} cm^{-3}) source barrier, a 3.6 nm C-doped P^{+}-$GaAs_{0.51}Sb_{0.49}$ (5×10^{19} cm^{-3}) source well, a 2.4 nm Si-doped N^{+}-InAs (5×10^{19} cm^{-3}) channel well, a 15 nm Be-doped P-InP (best effort $\approx 5\times10^{18}$ cm^{-3}) channel barrier, a 12.6 nm U.I.D. InP, a 10 nm Si-doped N^{+}-InP (2×10^{19} cm^{-3}) drain, a 10 nm Si-doped N^{+}-$In_{0.53}Ga_{0.47}As$ (3×10^{19} cm^{-3}) drain contact layer. Epi is bought from Intelliepi and the high doping of P-InP channel barrier is on best-effort basis (which has hole concentration of $\approx 5\times10^{18}$ cm^{-3} according to Intelliepi). Note that the InAlAs back barrier and InP etch stop layers are designed for anchor-shape vertical structure, thus are not used/taken in part in vertical planarization TFETs.

The body thickness of vertical 3-HJ TFETs is designed to range between 65 nm to 125 nm, and it is estimated to come out as 100 nm to 170 nm at the tunneling junction. Fig. 5.34 and 5.35 shows the transfer and output characteristics of 3-HJ TFETs having different body thickness. As can be seen, transfer curves are strongly affected by drain bias. This short channel effect, similar to the DIBL in MOSFETs, causes drain-induced barrier tunneling in TFETs. At the subthreshold regime, the center of the body is controlled by both gate and drain bias. By applying a high drain bias, the potential in the channel is pushed down, and electrons in the source tunnel through the barrier and get to the drain terminal. This short channel behavior is given by the low aspect ratio between gate length and fin width (30: 100 nm) despite a high P-InP channel doping of (5×10^{18} cm^{-3}), and is a result of poor gate electrostatics in TFETs. High SS of above

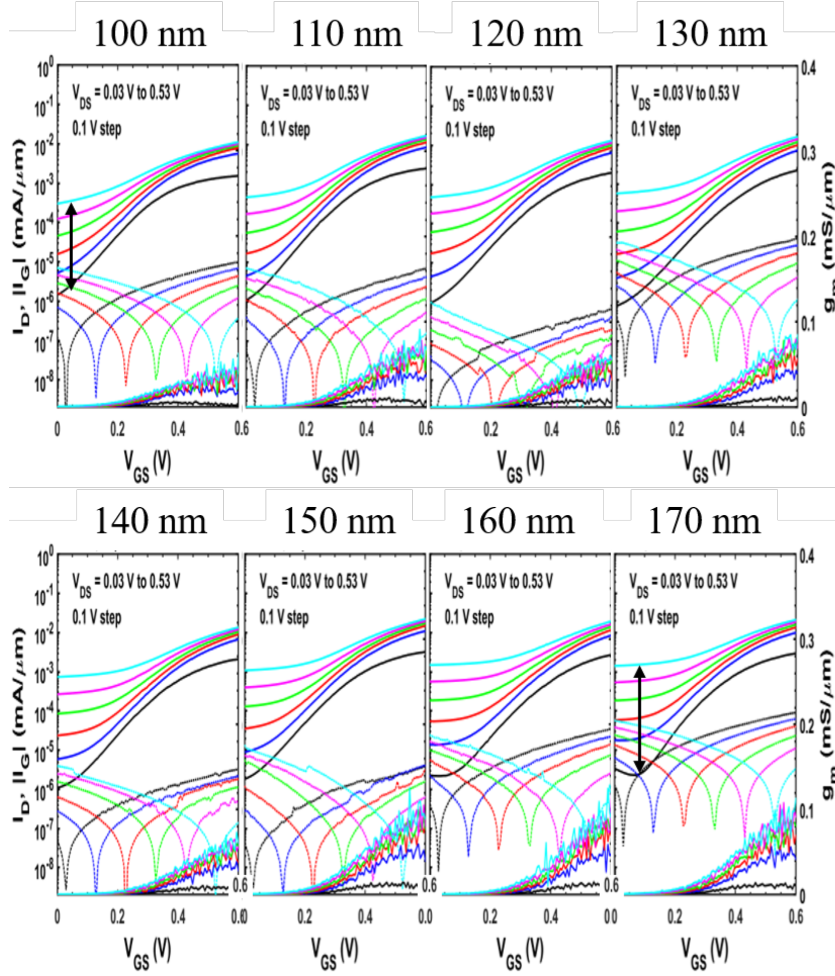


Figure 5.34: Transfer characteristics of 3-HJ TFETs having estimated body thickness of 100 nm to 170 nm at tunneling junction.

80 mV/dec at any V_{DS} and low on/off ratio are obtained. This short channel effect in TFETs is quantified using the same method as DIBL in MOSFETs for better consistency in the thesis. The DIBL characteristic is roughly 0.67 V/V for a T_{body} = 100 nm TFET and 1 V/V for a T_{body} = 170 nm TFET at V_{DS} = 0.3 V.

Peak transconductances do not show a clear trend with varied body thickness, as shown in fig. 5.34. Peak transconductance increases with drain bias at V_{DS} between 0 to 0.3 V, and roughly keeps at the same as drain bias gets higher. Output characteristics, as shown in fig. 5.35, are affected by the short channel effect and transconductance. Red

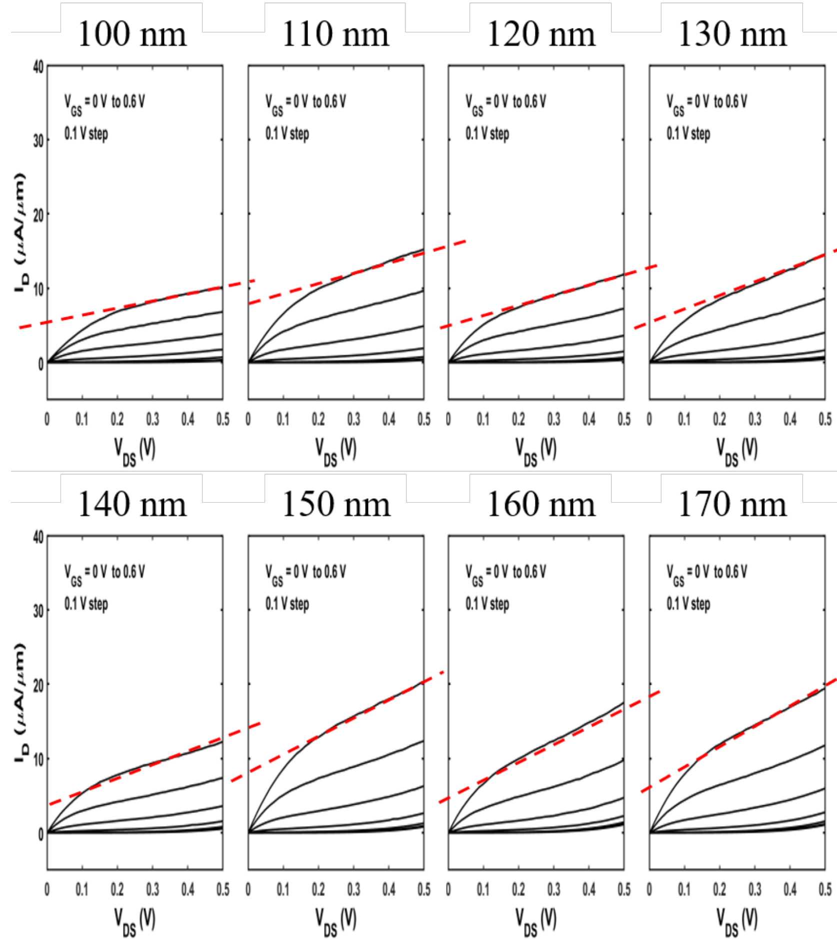


Figure 5.35: Output characteristics of 3-HJ TFETs having estimated body thickness of 100 nm to 170 nm at tunneling junction. Red dash lines are for the guidance of the degree of short channel effect.

dash lines are for the guidance of the degree of short channel behavior. A higher slope is observed in TFETs with a thicker body, which is consistent with the higher DIBL, as discussed before. At V_{DS} = V_{GS} - V_{TH} = 0.3 V, the averaged on current is roughly 2 $\mu A/\mu m$.

The experimental on current is much lower than the simulated I_{on} of $\approx$ 100 $\mu A/\mu m$ at P-InP doping of 5×10^{18} cm^{-3}. The low on-current might result from gate misalignment, fringing field coupling, high-k/InP interface traps, and the problematic epi (having high surface roughness at tunneling junction). As discussed in chapter 3, gate alignment

makes a significant impact on the performance of 3-HJ TFETs. A $\pm$ 4 nm gate/source underlap/overlap could lower the on-current by 2 orders of magnitude, while overlap by 2 nm could reduce it by 2 - 3 times. Even with the self-aligned gate in these top-down vertical TFETs, the number of digital etch before high-k deposition still needs to be optimized. Moreover, our vertical FET structure has some degree of fringing field that couples gate to the source's potential. This is not considered in simulation and might potentially be lower on-current. Additionally, D_{it} at high-k/InP interface degrades gate control over the channel. The loss of gate electrostatics could slightly reduce I_{on}. Most importantly, high surface roughness on semiconductor layers closed to InAs/GaAsSb/InGaAs junctions is observed during fabrication. Fig. 5.36 shows the SEM images (a) after InP wet etching, and (b) after finishing cycles of digital etch prior to high-k deposition in the fabrication of 3-HJ TFETs. The exposed surface is InAs and GaAsSb in (a) and (b), respectively. As can be seen, the high roughness/problematic epi could result in InAs strain relaxation, layer thickness variation, and defects/traps generation closed to tunneling junctions. All of them could lead to the decrease of on current.

Fig. 5.37 compared the performances of InP channel (a) MOSFET and (b) 3-HJ TFET at similar bias condition. The short channel effect in TFET with L_g = 30 nm and P-InP of 5×10^{18} cm^{-3} is better than in MOSFETs with L_g = 50 nm and P-InP of 1×10^{18} cm^{-3} as expected. The highest transconductance in TFET is ≈ 0.1 mS/μm, which is 1/4 of that in MOSFET. The on-current of MOSFET is dominant by the short channel effect, thus make the comparison with TFET meaningless.

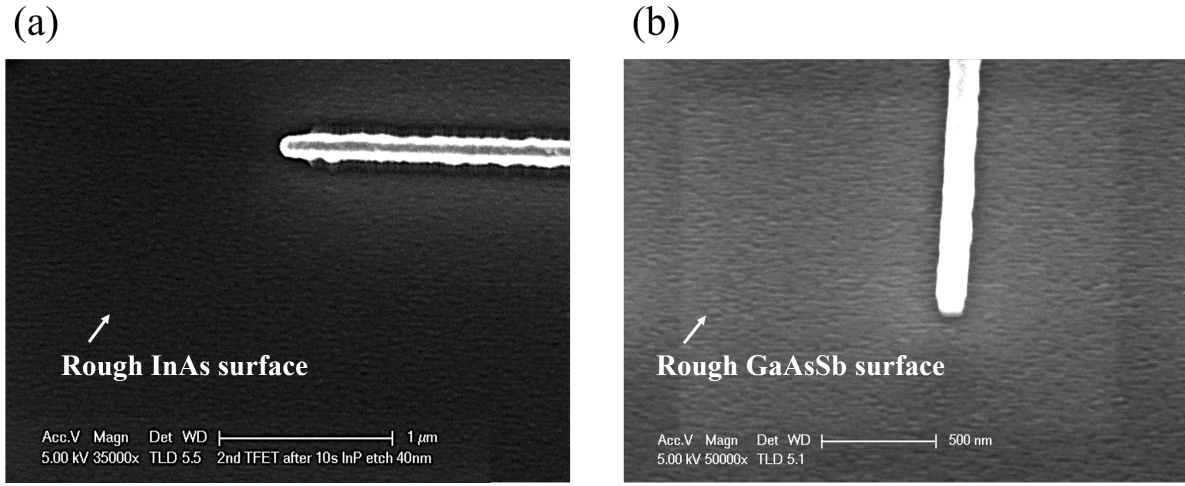


Figure 5.36: SEM images in the fabrication of 3-HJ TFETs (a) after InP wet etching, and (b) after finishing cycles of digital etch prior to high-k deposition. The exposed surface is InAs and GaAsSb in (a) and (b), respectively.

5.4.4 Negative Differential Resistance and Trap-assisted Tunneling in 3-HJ TFETs

NDR characteristic in tunneling diodes/TFETs is a way to confirm if the current transport mechanism is dominant by tunneling. Typically, NDR happens at drain bias between 0 to -0.3 V [39, 95, 96]. Fig. 5.38 (a) shows source current vs. V_{DS} with a floating gate. It is clear that there is NDR lying at $V_{DS} \approx 1.2$ V. Fig. 5.38 shows the $I_D - V_D$ characteristics of 3-HJ TFETs under negative drain bias and V_{GS} of 0 - 0.3 V in (a) linear and (b) log scale. Comparing to the on-current in output characteristics (fig. 5.35), the absolute I_D under negative drain bias is roughly 10 times higher than under positive drain bias. Moreover, the NDR in 3-HJ TFETs roughly lies at V_{DS} of -0.9 V, which is much smaller than in common TFETs. The drain current shows dispersive behavior under different V_{GS} before reaching the NDR region, joins to a similar drain current at NDR point, and further disperses again at V_{DS} below -1 V.

At V_{DS} below -1 V, drain current is dominant by the P-i-N (PNPN in this case) on-

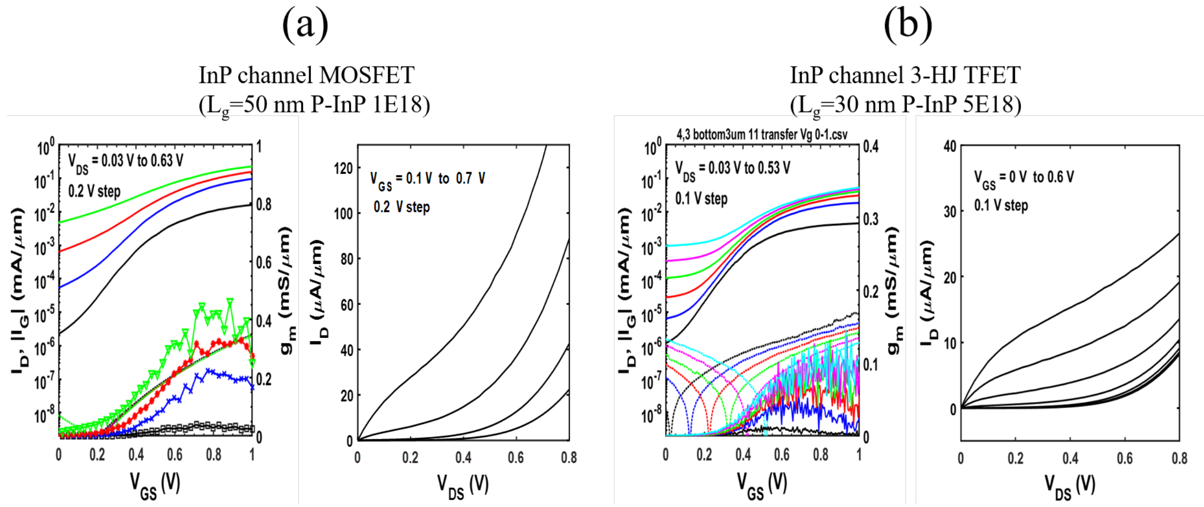


Figure 5.37: Comparison of the transfer and output characteristics of InP channel (a) MOSFET and (b) 3-HJ TFET at similar bias condition.

state diffusion current. At V_{DS} between 0 to -0.9 V, this high drain current (which is $\approx$ 10 times higher than the on-current of 3-Hj TFETs) could be explained by heterojunction interface traps assisted tunneling in 3-HJ TFETs. Considering the thick body in this demonstration of 3-HJ TFET, the following discussion on the transport characteristics is divided into two parts: 1. in the channel that is closed to the high-k/semiconductor interface, and 2. in the body. Fig. 5.39 shows the schematic band diagram of thick body 3-HJ TFETs near to high-k/semiconductor interface without (top) and with (bottom) heterojunction interface traps under different bias conditions. The bound states in the source well and channel well align at V_{GS} = 0.3 V. By applying positive drain bias, electrons tunnel from the source into the channel, as shown in fig. 5.39 (b). If a negative V_{DS} is employed, electrons diffuse from the drain into the channel, and tunnel from the channel into the source, as shown in fig. 5.39 (c). This indicates the drain current under negative V_{DS} should be smaller than that under positive V_{DS} in principle.

With the rough InAs and GaAsSb surfaces as shown in fig. 5.36, it is straightforward to think of the presence of defects and heterojunction interface traps in our 3-HJ TFETs.

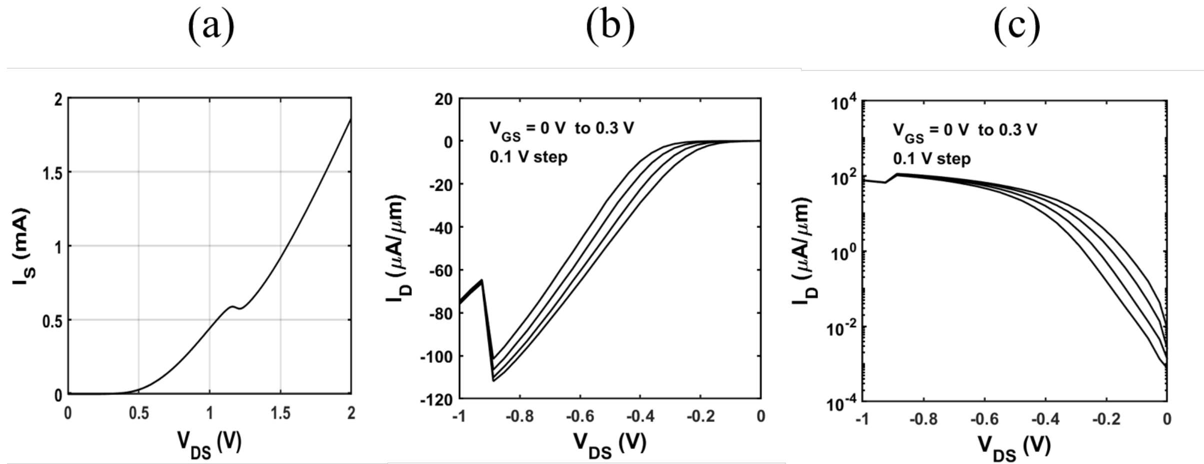


Figure 5.38: Source current vs. V_{DS} with a floating gate (a), and the $I_D - V_D$ characteristics of 3-HJ TFETs under negative drain bias and V_{GS} of 0 - 0.3 V in (a) linear and (b) log scale.

Assuming there is an InAs/InP interface trap state lying close to the conduction band of InAs in 3-HJ TFETs, as shown in fig. 5.39 (d), the transport behavior at certain bias conditions would be different. When the gate is on and V_{DS} is positively biased (fig. 5.39 (b)), electrons could tunnel from the source into InAs well through TAT, but would be further blocked by the InP channel barrier. Some electrons are excited to higher states by phonon-scattering, but the number of that is comparably small with respect to the electrons tunneling in the designed window. Thus, this TAT does not contribute to the on-current in 3-HJ TFETs under normal operating conditions. On the contrary, when 3-HJ TFET is under negative V_{DS} as shown in fig. 5.39 (f), the electrons that diffuse from the drain into the channel would fill out the lower energy states in the InAs well. TAT enhances the tunneling current under this condition.

Fig. 5.40 shows the schematic band diagram of thick body 3-HJ TFETs in the body without (top) and with (bottom) heterojunction interface traps under different bias conditions. Without heterojunction interface traps, there is no tunneling current under any V_{DS} because of the misalignment in bound state energies in the channel and

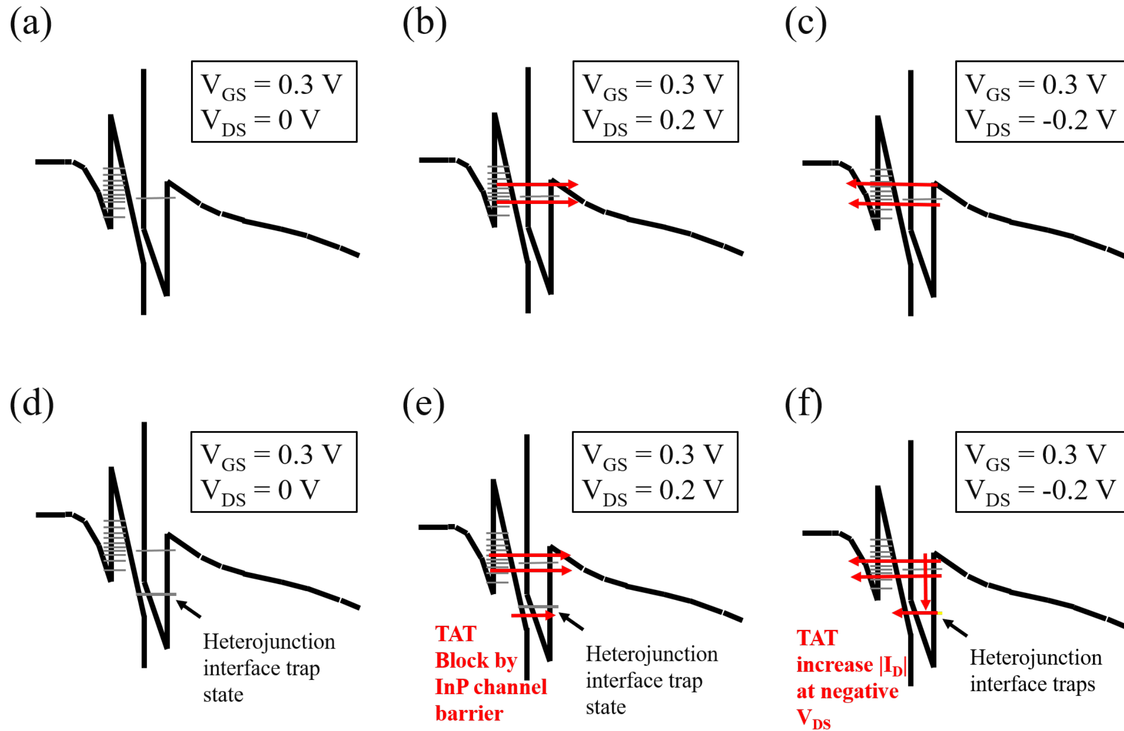


Figure 5.39: Schematic band diagram of thick body 3-HJ TFETs near to high-k/semiconductor interface without (top) and with (bottom) heterojunction interface traps under different bias conditions.

source wells, as shown in fig. 5.40 (a), (b), and (c). Similar to the transport behavior near the high-k/semiconductor interface, heterojunction interface traps introduce TAT current only at negative V_{DS} in the body, as shown in fig. 5.40 (d), (e), and (f). This suggests that TAT with trap states deep in channel well does not give rise to on current in 3-HJ TFETs, but would significantly enhance the drain current at negative V_{DS}. That drain current (at negative V_{DS}) is contributed from not only channel near to high-k/semiconductor interface, but also the whole body. This suggests that a thick body could result in a higher drain current under negative V_{DS}. Therefore, TAT well explains the much higher drain current under negative V_{DS} than under positive V_{DS} in our 3-HJ TFETs.

In common thin body P-i-N TFETs, electrons diffuse from the drain into the channel,

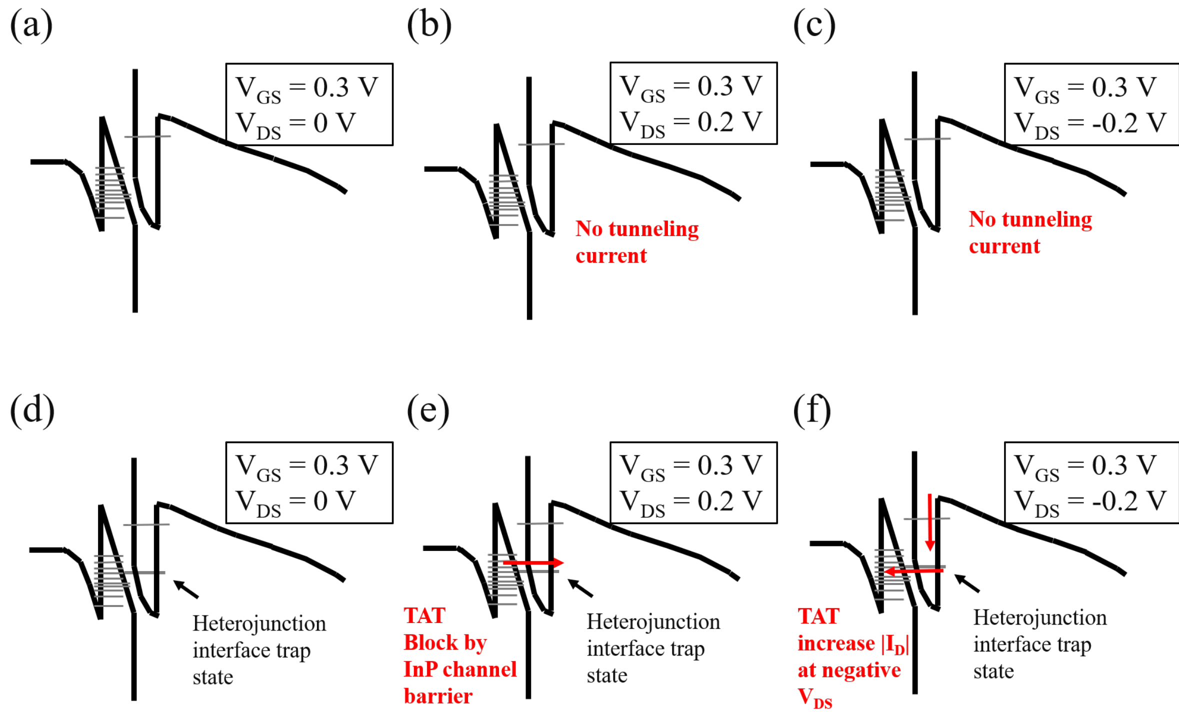


Figure 5.40: Schematic band diagram of thick body 3-HJ TFETs in the body without (top) and with (bottom) heterojunction interface traps under different bias conditions.

and tunnel from the channel into the source at small negative V$_{DS}$. The further decrease in the drain bias (similar to high injection in forward bias PN diode) results in the upward bending of channel potential. As V$_{DS}$ decreases more, the tunneling current would finally be cut off; thus, NDR appears. In the thick body 3-HJ TFETs with heterojunction interface traps lying deep in channel well, tunneling could only be cut off at a much lower V$_{DS}$ because of the presence of resonant bound states. This results in the occurrence of NDR at a low V$_{DS}$ = -0.9 V as in fig. 5.38 (b).

5.4.5 Summary

InP channel vertical MOSFET and 3-HJ TFETs are fabricated using a top-down planarization process. Vertical MOSFETs with TiN/Ru gate exhibit a comparable high peak g$_m$ of 0.42 mS/μm at V$_{DS}$ = 0.6 V and L$_G$ = 50nm with planar MOSFETs, indicating

successful proof of process. The first demonstrated vertical InGaAs/GaAsSb/InAs/InP TFETs show SS of > 80 mV/dec at V_{DS} = 0.1 V, and I_{on} of 2 $\mu m/\mu m$ at V_{DS} = V_{GS} - V_{TH} = 0.3 V with strong short channel effect. Despite SS > 60 mV/dec, tunneling is the main transport mechanism in 3-HJ TFETs confirmed by the measured NDR characteristics at negative V_{DS}. The high SS is due to the short channel effect and high-k/InP interface traps. The low on-current could result from gate misalignment, reduced electric field at tunneling junction due to fringing field, and problematic epi (having high surface roughness at tunneling junction. The NDR behavior in 3-HJ TFETs is analyzed with TAT to explain the experimental measured $I_D - V_D$ characteristics.

Chapter 6

Conclusion

6.1 Summary

Toward the ultimate goal of replacing Si finfets and nanowire FETs in the low power logic component, TFETs must be proved with a reduced subthreshold leakage (SS $<$ 60 mV/dec) as well as a high on-current at a lower supply voltage ($V_{DD} < 0.5$ V). Until now, interconnect scaling is emerging to be the next bottleneck over transistor scaling in terms of RC time delay in Si CMOS [X]. I_{on} greater than 100 $\mu A/\mu m$ at $V_{DD} = 0.3$ V is an absolute bottom line for TFETs to be considered better than MOSFETs. Therefore, it is crucial to pursue a high on-current in TFET rather than simply reducing SS.

In order to increase the on-current in TFET, tunneling probability needs to be maximized, and a thin tunneling distance is required. A summary of methods for tunneling enhancement, including material design, geometry optimization, heterojunction incorporation, and doping profile engineering, are discussed. 3-HJ resonant enhanced TFET is of particular interest due to its potential of achieving 100% transmission. However, previous works on 3-HJ TFET design failed to consider physical limitations in TFET fabrication,

such as Matthew Blakeslee's limit of strain layer epitaxy, the restricted doping concentration in phosphide materials, and body thickness. A realistic InGaAs/GaAsSb/InAs/InP TFET is proposed. The optimization of this 3-HJ TFET is done by varying source and channel well thickness based on the highest on current targeting at a supply voltage of 0.3 V. By employing PNPN doping design in a comparably thick body (T_{body} = 12 nm) 3-HJ TFET, the simulated I_{on} reaches $\approx$ 300 $\mu A/\mu m$ at a P-InP body doping of 2×10^{19} cm^{-3}, and $\approx$ 100 $\mu A/\mu m$ at a P-InP body doping of 5×10^{18} cm^{-3}.

Different gate metallization techniques are investigated. A high-quality ALD TiN/Ru gate process is developed, and the conformity of ALD deposited Ru is improved by a low-temperature ABC-type growth. Planar InP channel MOSFETs with ALD TiN/Ru gate accomplish a record low averaged SS of 68 mV/dec at V_{DS} = 0.1 V among long gate length devices, and a record high peak g_m of 0.75 mS/μm at V_{DS} = 0.6 V for a L_g = 80 nm device. The ALD TiN/Ru gate is integrated with the inherited source and drain metallization on different device structures. 3 generations of device physical structures are tried. InP channel vertical MOSFETs and 3-HJ InGaAs/GaAsSb/InAs/InP TFETs are demonstrated using top-down planarization process. InP channel vertical MOSFETs exhibit a peak g_m of 0.42 mS/μm at V_{DS} = 0.6 V. This peak g_m is comparable to that in planar MOSFETs with L_g = 50 nm, indicating proof of process. The first demonstrated 3-HJ TFETs show a SS > 80 mV/dec at V_{DS} = 0.1 V, and I_{on} of 2 $\mu A/\mu m$ at $V_{DS} = V_{GS} - V_{TH}$ = 0.3 V with strong short channel effect. Despite SS > 60 mV/dec, tunneling is the main transport mechanism in 3-HJ TFETs confirmed by the NDR characteristics. The low experimental performances comparing to simulation are owing to short channel effect, gate misalignment, reduced electric field at tunneling junction due to fringing field, and problematic epi.

6.2 Future work

Even though the first demonstrated 3-HJ TFET shows low performance, it is still too early to determine if this design is problematic. Fabrication and process modifications/refinements, including increasing gate length to eliminate short channel effect, an epi with smooth heterojunction interfaces to minimize interface trap density, and the optimization of gate alignment, need to be done. Back end Ru dry/wet etching is the process that causes a bad yield in vertical planarization FETs. A dry and chemical etching by oxygen radicals (i.e., YES EcoClean system) is a solution to that. Furthermore, a higher P-doped InP channel is a certain way to improve electrostatics, thus increase I_{on}.

In fact, demonstrating higher capabilities of TFETs over MOSFETs at a low power regime is just the first step. To justify the cost of developing/employing III-V technology in VLSI, much higher performance and the ability/feasibility of transistor scaling is required. Additionally, the integration of TFETs on the current Si CMOS platform might also be a vital topic in the future.

After all, developing a new technology that competes with the rapidly changing VLSI Si CMOS is never easy. It is crucial to follow the state-of-the-art progression of Si finfets and GAA FETs while keeping in mind the advantages and disadvantages of TFETs.

Appendix A

Planar and Vertical FETs Process Flow

A.1 Overview

The MOSCAP, planar MOSFET, and vertical MOSFET structures are grown on semi-insulating InP substrate by MOCVD at UCSB. The InP substrates with 3-HJ TFET epitaxial structure are acquired from Intelliepi Epitaxy Technology, Inc., where P-InP channel with doping concentration of 5E18 is on a best effort basis. Epitaxial structure is grown on a 3 inch semi-insulating InP substrate in MBE system. Device fabrication is done in the UCSB nanofabrication facility. In this appendix, detailed recipes, parameters, and step-by-step process flow of the fabrication of planar and vertical FETs are described.

A.1.1 MOCVD growth

The system used in this study is Thomas Swan 2" horizontal laminar flow reactor with H_2 as carrier gas. Real time metal-organic flows are monitored and controlled by

	a	b	Bath Temp (°C)	PMO (Torr)
TMGa	8.1	1703.0	-3.9	55.1
TMIn	10.5	3014.0	21.9	2.0
TBAs	7.2	1509.0	1.9	56.7
TBP	7.6	1539.0	17.9	197.5
TMAl	8.2	2134.8	2.4	3.0
DeZn			0.0	5.0

Table A.1: MO sources used in this study

	TMGa (sccm)	TMIn (sccm)	TBA (sccm)	TBP (sccm)	V/III	DiSi/DeZn (sccm)	n (cm^{-3})	p (cm^{-3})	G.R. (A/s)
InGaAs	2	113	150		25.4	30	9E+19		1.89
InP		397		190	47.4	4.5		1E+18	3.65
InP		198.5		190	94.8	5.5	1E+19		1.83

Table A.2: MOCVD growth parameters

Epison acoustic impedence monitors. Table A.1 summarizes the metal-organic source parameters and the calculated partial pressure. Molar flow (F), which determines the amount of precursors that flows into the liner, is given by

$$F(mol/min) = P_{MO}/P_{bubbler} \times \text{ (MO flow in sccm) / (22400 mol/min)}$$

where $P_{bubbler}$ is 760 torr. In order to incorporating sufficient silicon and zinc dopants into the grown film, the growth temperature is fixed at 600°C. The metal-organic source flows, V/III ratio, growth rate, and electron/hole concentration are summarized in table A.2

A.1.2 Drain contact formation

- Surface preclean by HCl:DI (25:250 mL) for 1 minute prior to loading into E-beam evaporator.
- Load samples and Mo source into E-beam#1. Wait until pressure pumped below 1E-6 torr. Start degasing Mo source for 15 minutes.
- Deposit Mo at a rate of 0.3 A/sec

- Cooling down for 30 minutes. Unload the sample from E-beam#1 and load into sputter#5 system immediately.
- Deposit TiW using dual beam. Power: 280/300 W (Ti/W). Pressure: 5 mtorr. Rotation: 20 rpm. Height: 1.5. Thickness: 500 nm.
- Deposit SiO_2 and SiN (80/40 nm) at 250°C by PECVD#1.
- Load into Thermal#1. Wait until pressure below 2E-6 torr.
- Degas Cr and deposit 30 nm Cr as hard mask for TiW etch.
- After finish, cool down for 15 minutes minimum then unload the sample.
- Deice 2% HSQ for 15 minutes. Spin HSQ at 4000 rpm. Bake at 200°C for 2 minutes.
- Load into EBL system. Write with 4th lens. Beam current: 10 nA (for >60 nm features). Base dose: 5500 $\mu C/cm^2$.
- Mix NaOH:NaCl:DI (2g/8g/200mL) until fully dissolve. Develop for 1 minute. Do not stir. Flowing DI for 3 minutes minimum. Check with OM under dark field.
- O_2 clean ICP#1 for 10 minutes before use. Change gas from N_2 to He. Season the chamber for 2 minutes. Etch 30 nm Cr using recipe as in table A.3.
- CF_4 clean following by O_2 clean the chamber for >7 minutes minimum individually. Change gas from CF_4 and CHF_3 to SF_6 and Ar. Season the chamber for 2 minutes.
- Wait until chamber temperature goes back to 50°C (chuck temp should always be $\approx$ 10°C.)
- Etch SiN/SiO_2/TiW using recipe as in table A.3. TiW etch time no more than 40 sec each run to avoid heat accumulation on substrates, that will result in TiW

Cr etch					
Cl_2 (sccm)	O_2 (sccm)	Pressure	Power	Bias	Etch time
24	6	1.33 (Pa)	600W	50W	1 min 30 s
TiW & SiO_2 & SiN_x etch					
SF_6 (sccm)	Ar (sccm)	Pressure	Power	Bias	Etch rate
20	5	1 (Pa)	600W	200W	4 nm/s

Table A.3: ICP#1 Cr and TiW etch recipe

undercut. Check with SEM, Detek and substrate colors between run. After seeing substrate colors gradient (which happenes when TiW is almost clean), add ≈15sec to clean up TiW as well as the Mo underneath.

- Acetone, IPA, DI clean 3 minutes individually.
- Mix Tergitol NP-10 ≈ 2 cc with DI ≈ 20 cc thoroughly in lift-off only BHF beaker. Then add > 250cc BHF to the beaker. Stir for 15 minutes until foam appears at liquid surface.
- Put samples in wafer holder that sits vertically. Lift-off Cr in BHF for 5 minutes. Stir both horizontally and vertically vigorously during lifting-off to ensure the Cr film is not attaching back to sample surface.
- Flowing DI for 5 minutes. Completely blow dry before loading into PECVD#1 system for SiN_x formation. (If wafer is not fully dried, it would break in high temperature chamber due to the thermal stress)

A.1.3 SiN_X sidewall Formation

- O_2 clean PECVD#1 for 10 minutes and season the chamber for 2 minutes.
- Deposit selected thickness of SiN_x. Add silicon witness together with device samples.

Low power SiN_x etch					
CF_4 (sccm)	O_2 (sccm)	Pressure	Power	Bias	Etch rate
20	2	0.3 (Pa)	25W	19W	12 nm/min

Table A.4: ICP#1 low power SiN_x etch recipe

- Measure the deposited thickness on witness using ellipsometry.
- O_2 clean the ICP#1 chamber for 10 minutes and season the chamber for 2 minutes.
- Wait until chamber temperature cool down to 50°C.
- Etch SiN_x on witness silicon for 1 minute using recipe as shown in table A.4. Check the etched thickness by ellipsometry and calculate the etch rate.
- Etch SiN_x on device samples. Always do a ¿10% over etch to ensure it is cleaned.
- Before the upcoming semiconductor wet etch, 2 cycles of digital etch is needed to clean up the surface. This step is to prevent the polymer reposition on samples from SiN_x etch that might affect the wet etch. Digital etch: 10 minutes UV ozone following by 1 minute HCl:DI (25:250 mL) and 2 minutes flowing DI.

A.1.4 Semiconductor wet etch

InGaAs etch

- Mix H_3PO_4:H_2O_2:DI (10:10:250 mL) and stir at 200 rpm for at least 10 minutes.
- Light off. Etch with 7 sec minimum. Etch rate: 4nm/sec.

InP etch

- Mix H_3PO_4:HCl (200:50 mL) and stir at 200 rpm for at least 15 minutes.
- Light off. Etch with 9 sec minimum. Etch rate: 10nm/sec.

InAlAs etch

- Mix H_3PO_4:H_2O_2:DI (10:10:250 mL) and stir at 200 rpm for at least 10 minutes.
- Light off. Etch with 7 sec minimum. Etch rate: 3nm/sec. Must stir vigorously during etch to ensure the etching uniformity.

After any wet etch, flowing DI water for 3 minutes minimum is needed. Check the thickness by Detek.

A.1.5 ALD high-k and TiN/Ru

Oxford FlexAL plasma-enhanced ALD system is used in this study. Remote ICP is designed to minimize the plasma damage on channel. Three seperate chambers are connected by a vacuumed transfer chamber, which is attached to load-lock. Chamber 1 and 3 are designated to deposit metals and dielectrics, respectively, while chamber 2 is dedicated to ion-milling. The TiN/Ru gates are done by in-siu transfering samples from chamber 3 (for TiN growth) to chamber 1 (for Ru growth) without breaking the vacuum.

- For all strucures, at least 1 cycle of digital etch is done to clean up the native oxide on surface. Digital etch: UV ozone for 10 minutes, HCl:DI (25:250 mL) dip for 1 minute minimum, then flowing DI for 2 minutes. For MOSCAP or planar-MOSFET structure, samples are dippe in BHF for 1 minutes follwoing by flowing DI for 2 minutes. All samples need to completely blow dry. (If wafer is not fully dried, it would break in high temperature ALD chamber due to the thermal stress)
- Load into ALD chamber#3 right after digital etch. Add a silicon witness sample together with device samples to check the deposited thickness and etch rates for the later etching test.

- Deposit 9 cycles of AlO_xN_y, following selected cycles of ZrO_2. Ramp up chamber#3 to 350°C and anneal with H_2 flow 30 sccm, pressure 220 mtorr for 30 minutes.
- Unload the samples and cool the chamber back down to 300°C.
- Check the high-k thickness on silicon witness using ellipsometry.
- Load the samples into ALD chamber#3. Attach the silicon witness that has high-k on it. Deposit 35 cycles of TiN.
- After done, transfer the samples to chamber#1 without breaking vacuum. Deposit selected cycles of Ru. This step will take few hours.
- Unload samples right after it's done.
- Load samples to chamber#3 to perform 30 minutes H_2 anneal at 350°C to recover the plasma damage from TiN growth. Parameters are the same as above.

Table A.5, A.6 and A.7 summarizes the growth parameters and growth rates of interfacial AlO_XN_Y, ZrO_2, TiN and Ru.

TiN 300°C (0.57 A/cycle)					
TDMAT			N2/H2 Plasma (12/4 sccm)		
Dose time	Hold time	Pressure	Plasma power	Plasma time	Pressure
0.25s	0.75s	80mTorr	400W	20s	2.5mTorr

ZrO_2 300°C (0.75 A/cycle)					
TEMAZ			H2O		
Dose time	Hold time	Pressure	Dose time	Hold time	Pressure
1s	0.01s	220mTorr	0.05s	—	80mTorr

AlO_xN_y 300°C (1 nm/9 cycles)					
N2 Plasma (20 sccm)			TMA		
Plasma power	Plasma time	Pressure	Dose time	Hold time	Pressure
100	2s	20mTorr	0.04s	—	80mTorr

Table A.5: ALD high-k and TiN deposition parameters

Ru 250°C with H_2 (0.37 A/cycle)						
Ex03-Ru			O2 (100 sccm)		H2 (50 sccm)	
Dose time	Hold time	Pressure	Dose time	Pressure	Dose time	Pressure
2.5s	2.5s	1500mTorr	5s	500mTorr	5s	500mTorr

Table A.6: ALD Ru (with H_2) deposition parameters

Ru 300°C without H_2 (0.6 A/cycle)				
Ex03-Ru			O_2 (100 sccm)	
Dose time	Hold time	Pressure	Dose time	Pressure
2.5s	2.5s	1500mTorr	5s	500mTorr

Table A.7: ALD Ru (without H_2) deposition parameters

A.1.6 TiN/Ru etch

- CF_4 clean ICP#2 system for 15 minutes minimum. Season the chamber for 2 minutes.

- Test the Ru etch on witness sample (witness sample contains Ru/TiN/ZrO_2/AlO_xN_y on silicon substrate). If chamber condition is bad (i.e. chamber coated by byproduct polymers that are generated from other users' etch), the polymers would redeposit on our samples during Ru etch. This will affect the upcoming TiN/high-k wet etch. Ru etch recipe is shown in table A.8.

- PE II to oxidize the test witness. Power: Pressure: 300 mtorr. Power: 100 watt under O_2 environment for 1 minutes.

- BHF dip for 1 minutes on witness sample. Flowing DI for 3 minutes. If the silicon witness comes out hydrophobic, the upper layers are all cleaned. Double check if its clean using ellipsometry.

- If the test on wintess goes well, etch device samples with the same method as shown above. If the test witness comes out to be hydrophilic or there are ¿2 nm layers remain on samples (check by ellipsometry), contact staff. ICP#2 chamber needs to be wet cleaned.

Cl_2 (sccm)	O_2 (sccm)	Pressure	Power	Bias	Etch rate
49.5	5.5	2.5Pa	500W	50W	1.33nm/sec

Table A.8: ICP#2 Ru dry etch recipe

A.1.7 BCB Planarization

Planarization is done using BCB curing technique. Because there will be large strain generating by this process, fins need to be mechanically protected/supported by SiN film to prevent being torn off.

BCB ash				
CF_4 (sccm)	O_2 (sccm)	Pressure	Power	Etch rate
50	200	40 (Pa)	1000W	variant

Table A.9: ICP#1 BCB ash recipe

- Turn on blue oven and set the N_2 flow to 100 scfh to purge the chamber.
- Apply cyclotene 3022-46 (BCB) on the sample. Soak for 1 minute and spin-coat at 2000 rpm for 45 sec.
- Immediately transfer the sample into the Blue Oven. Set the N_2 flow to 60 scfh. seconds.
- Ramp to 50°C over 5 minutes and hold temperature for another 5 minutes.
- Ramp to 100°C over 15 minutes and hold temperature for another 15 minutes.
- Ramp to 150°C over 15 minutes and hold temperature for another 15 minutes.
- Ramp to 250°C over 1 hour and hold temperature for another 1 hour.
- Wait >8 hours for the chamber to cool down with the chamber door closed
- Measure the BCB thickness by ellipsometry.
- Season ICP#1 ash chamber for 2 minutes. Make sure the temperature is at 50°C.
- Ash samples using the recipe shown in table A.9 and check BCB thickness in between runs.
- Check if posts/TiW are revealed by Detek, AFM and SEM. If not, add 10-60 sec ash until they're all >50 nm above BCB surface.
- Be sure not over ash due to micro-trenching effect.

A.2 Planar MOSFET Process Flow

1	Dummy gate		Solvent clean	Acetone/IPA/DI 3min		
2		Litho(EBL)	HSQ	4000 rpm bake 200C 2mins Develop 1min + DI flowing 3min NaOH:NaCl:DI = 2g:8g:200ml	4th lens-10nA base:5500 Do not stir during develop/rinse	OM check (dark field)
3			Surface clean	HCl:H2O 1:10 1min	Do it right before loading into MOCVD	
4		MOCVD	S/D regrowth	U.I.D. InP spacer ~8nm N+ InP spacer ~7nm N+ InGaAs ~130 nm	N+ InGaAs 9E19 by Hall measurement	check 4-point Rsh ~20ohm/square
5	Mesa litho		Solvent clean + Dehydration (opt.)	Acetone 3min/IPA 3min/DI 3min 110C >3min		
6		Litho(Autostepper)	SPR 955 (0.9um)	3000rpm 90C 90s Ex = 0.27s F = 0 PEB 110C 90s; 300MIF 1min		
7	Mesa isolation etch	PE II	O2 descum	O2 300mT 100W 30s	To clean up resist residue	
8			Hard bake	>110C >8min	To improve resist adhesion on semiconductor –>prevent etch undercut	
9			InGaAs etch H3PO4/H2O2/H2O 1:1:25 10:10:250	Light-off; stir w/ stirrer for 10 min before; use squirt bottles ER: ~3nm/s (50%OE) + DI 3min		OM & Detek check
10			InP etch H3PO4/HCl 4:1 R.T. 200:50	Light-off; stir w/ stirrer for 10 min before; use squirt bottles Etch 8s up (ER:~10nm/s) +DI 3min	Etch into InP substrate ~100 nm	OM & Detek check
11			Strip	NMP >1 hour + IPA 3min + DI 3min		
12	High-k		BHF	BHF 2min	remove HSQ & native oxide	
13			Digital etch x1 (opt.)	UV ozone 10 mins HCl:H2O=1:10 1min		
14		ALD chamber 3	seasoning	ALD seasoning N* 60 cycles with carrier wafer		
15			N*+ZrO2	N* 9 cycles + ZrO2 40 cycles		check with dummy
16			H2 Anneal	350C 30min H2 flow=30 sccm	In-situ after ZrO2. Set time to 40min (10min heat up & 30min anneal)	
17	Metal gate		TiN	300C 35 cycles (~2nm)		
18		ALD chamber1	Ru	300C 500 cycles (~30nm)	Dep Ru right after TiN	
19		ALD chamber 3	H2 Anneal	350C 30min H2 flow=30 sccm	Post gate metal H2 anneal to recover plasma damage from TiN deposition	
20	Gate litho	PECVD #1	SiN etch (opt.)	30nm		
21		Autostepper	SPR 955 (0.9um)	3000rpm 90C 90s Ex = 0.27s F = 0 PEB 110C 90s; 300MIF 1min		
22	Gate metal etch	ICP #1 (opt.)	SiN etch	low power etch OE<20%	12nm/min	check with dummy
23		ICP #2	Chamber clean	15 min CF4 + 15 min O2 clean		
24			Ru etch	Ru: Cl2/O2 = 5/49.5 sccm 500W/50W ER: 1.333 nm/s OE>20%		check with dummy
25	High-k & TiN etch	PE II	TiN oxidation	O2 300mT 100W 1min		
26			BHF TiN & high-k etch	1min		check with dummy
27			Strip	NMP >1 hour + IPA 3min + DI 3min		
28	S/D litho		Digital etch x2	**UV Ozone 10min** **HCl:H2O 1:10 1min**		
29			Solvent clean + Dehydration (opt.)	Acetone 3min/IPA 3min/DI 3min 110C >3min		
30		Litho(Autostepper)	nLof 2020 (2um)	3000rpm 110C 60s Ex = 0.17s F = -6 PEB 110C 90s; 300MIF 80s		
31	S/D metal	PE II	O2 descum	O2 300mT 100W 30s	To clean up resist residue	
32			Surface clean	HCl:H2O 1:10 1min		
33		Thermal evaporator	Pd/Ni/Au	20/50/200 nm		
34			Lift-off	NMP overnight + IPA 3min + DI 3min		OM check
35	Passivation	ALD chamber 3	Al2O3	30 cycles	To passivate semiconductor surface dangling bonds	

A.3 Vertical FET Process Flow

1	Drain metal stack		Solvent clean	Acetone/IPA/DI 3min		
2			Surface clean	HCl:H2O 1:10 1min		
3		E beam #1	Mo dep	20nm (0.3A/sec)	Degas for 15 minutes @1.3	
4		Sputter #5	TiW dep	280W/300W; tilt 10/10 height:1.5; rotate:20 66min ~500nm TiW		
5		PECVD	SiO2/SiN	80nm/40nm		
6		Thermal evaporator	Cr mask	30nm		
7	Drain litho	PE II	O2 descum	300mT/100W 30s		
8		Litho(EBL)	HSQ	4000 rpm bake 200C 2mins develop=1min DI 3min NaOH:NaCl:DI = 2g:8g:200ml	4th lens-10nA; base dose: 5500 Do not stir during develop/rinse	OM check
9	Drain metal etch	ICP #1	Cr etch	75-90s	change descum N2–>He	
10		ICP #1	Gas change	SF6/Ar (He descum)		
11			TiW etch	1. O2 clean 15min 2. change gas 3. chamber seasoning #162 (HP) 2min	ER = 4.3nm/sec Higher T results in TiW undercut –>no more than 40s each run (prevent T ramp up during etch)	Loading effect –> 1. Need to surround sample with TiW spacers 2. Sample + spacers size no biger than 2' wafer OM/SEM check
12			Cr mask removal BHF	less than 4-5min		OM/SEM check
13					DI 3min	
14	1st sidewall	PECVD	SiN sidewall	25-30nm		
15			SiN etch	low power etch OE: 10%	12nm/min	Check with dummy
16			Digital etch x3	UV Ozone 10min HCl:H2O 1:10 1min		
17	InGaAs etch		InGaAs etch H3PO4/H2O2/H2O 1:1:25 10:10:250	Light-off; stir w/ stirrer for 10 min before; use squirt bottles Etch 7s up (ER:~3nm/s) + DI 3min		OM check
18			Solvent clean	Acetone/IPA/DI 3min		
19	2nd sidewall	PECVD	2nd sidewall	25-30nm		
20		ICP #1	SiN etch	low power etch OE: 10%	12nm/min	Check with dummy
21			Digital etch x2	UV Ozone 10min HCl:H2O 1:10 1min		
22	InP etch		InP etch H3PO4/HCl 4:1 R.T. 200:50	Light-off; stir w/ stirrer for 10 min before; use squirt bottles Etch 8s up (ER:~10nm/s) +DI 3min	Deposit high-k right after InP etch + digital etch	OM check
23			Surface clean (opt.)	HCl:H2O 1:10 30s		
24	Junction digital etch		Digital etch x N	UV Ozone 10min HCl:H2O 1:10 5min	N depends on high-k and junction thickness	
25	High-k	ALD chamber 3	chamber season	ALD seasoning 60 cycles ZrO2 with carrier wafer		
26			N*+ZrO2	N* 9 cycles + ZrO2 40 cycles		check with dummy
27			H2 Anneal	350C 30min with H2 30 sccm		
28	Gate metal	ALD chamber 3&1	TiN/Ru	**TiN: 300C 35c (~2nm)** **Ru: 300C 500c (30nm)** **or 250C 675c (25nm)**		
29		ALD chamber 3	H2 Anneal	350C 30min with H2 30 sccm	Post gate metal H2 anneal to recover plasma damage from TiN deposition	
30		PECVD #1 (opt.)	SiN hard mask for Ru etch	30nm	Optional. Can use SPR as hard mask for micron size features	
31		Litho (Autostepper)	SPR 955 (0.9um)	3000rpm 90C 90s Ex = 0.27s F = 0 PEB 110C 90s; 300MIF 1min		
32		ICP #1 (opt.)	SiN etch	low power etch OE<20%	12nm/min	Check with dummy
33		ICP #2	Chamber clean	15 min CF4 + 15 min O2 clean		
34			Ru etch	Ru: Cl2/O2 = 5/49.5 sccm 500W/50W ER: 1.333 nm/s OE>20%		Check with dummy
35	High-k & TiN etch	PE II	TiN oxidation	O2 300mT 100W 1min		
36			TiN & high-k etch BHF	1min		Check with dummy
37			Strip	NMP (80C bath) >1 hour + IPA 3min + DI 3min		
38	Source litho		Digital etch x3	UV Ozone 10min HCl:H2O 1:10 1min		
39			Solvent clean + Dehydration (opt.)	Acetone 3min/IPA 3min/DI 3min 110C >3min		
40		Litho (Autostepper)	nLof 2020 (2um)	3000rpm 110C 60s Ex = 0.17s F = -6 PEB 110C 90s; 300MIF 80s		

41	Source metal	PE II	O2 descum	O2 300mT 100W 30s	To clean up resist residue	
42			Surface clean	HCl:H2O 1:10 1min		
43		E-beam #1 or #4	N contact for MOSFETs P contact for TFETs	20/50/100nm Ti/Pd/Au n metal 5/10/5/100nm Pd/Ti/Pd/Au p metal		
44			Lift-off	NMP >1 hour + IPA 3min + DI 3min		OM check
45	Post litho		Solvent clean + Dehydration	Acetone 3min/IPA 3min/DI 3min 110C >3min		
46		Litho (Autostepper)	nLof 2020 (2um)	3000rpm 110C 60s Ex = 0.17s F = -6 PEB 110C 90s; 300MIF 80s		
47	Post metal	PE II	O2 descum	O2 300mT 100W 30s	To clean up resist residue	
48		E-beam #1 or #4	Ti/Pd/Au		Post thickness depends on the height difference between the top of fins and SM/GM surface	
49			Lift-off	NMP (80C bath) >1 hour + IPA 3min + DI 3min		OM check
50	Back end planarization		Surface clean	HCl:H2O 1:10 1min		
51		PECVD #1	SiN	30nm	Protect fins from being destroyed by high strain BCB	
52			BCB spin	Soak for 1min 2000rpm	*Any bubble will result in huge height difference after bake *Keep pippette head at the same place and close to sample surface during drop to prevent bubbles formation	
53		Blue oven	BCB bake	250C 1hr in N2 Whole process takes ∼8 hours	* Use specific holder to prevent sample sticking to chamber floor *Surround the sample with Al foil to prevent sample drop	
54		ICP #1	BCB ash			1. Check BCB thickness by ellipsometry (BCB>900 nm) 2. Check the exposed between detek & ellipsometry (700<BCB<900 nm) 3. Check the exposed by AFM at last
55						
56	Drain exposure	ICP #2	Ru etch	**Details please check Chapter 5		
57		PE II	TiN oxidation	O2 300mT 100W 1min		
58			TiN & high-k etch BHF	1min		check with dummy
59	D/G isolation	PECVD #1	SiN	SiN >150nm	Prevent drain-gate short after putting on PAD	
60		Litho (Autostepper)	SPR 955 (0.9um)	3000rpm 90C 90s Ex = 0.27s F = 0 PEB 110C 90s; 300MIF 1min		
61		ICP #1	SiN etch	10%<OE<20%	12nm/min	check with dummy
62			Strip	NMP (80C bath) 1 hour + IPA 3min + DI 3min	Don't strip for more than 1 hour to prevent BCB tear off	
63	PAD litho		Solvent clean + Dehydration (opt.)	Acetone 3min/IPA 3min/DI 3min 110C >3min		
64		Litho (Autostepper)	nLof 2020 (2um)	3000rpm 110C 60s Ex = 0.17s F = -6 PEB 110C 90s; 300MIF 80s		
65	PAD metal	PE II	O2 descum	O2 300mT 100W 30s		
66		Ebeam #1 or #4		Ti/Au 200/10000A		
67			Lift-off	NMP (80C bath) 1 hour + IPA 3min + DI 3min	Don't strip for more than 1 hour to prevent BCB tear off	OM check

Bibliography

[1] P. Kushwaha, *Modeling of Substrate Depletion, Self-heating, Noise and High Frequency Effects in Fully Depleted SOI MOSFETs.* PhD thesis, Indian Institute of Technology Kanpur, 2016.

[2] K. Mistry, C. Allen, C. Auth, B. Beattie, D. Bergstrom, M. Bost, M. Brazier, M. Buehler, A. Cappellani, R. Chau, *et. al.*, *A 45nm logic technology with high-k+ metal gate transistors, strained silicon, 9 Cu interconnect layers, 193nm dry patterning, and 100% Pb-free packaging*, in *2007 IEEE International Electron Devices Meeting*, pp. 247–250, IEEE, 2007.

[3] M. Bohr and K. Mistry, *Intel's revolutionary 22 nm transistor technology*, Intel website (2011).

[4] H. Bu, *Nanometer Transistors Inching Their Way into Chips*, IBM Research, Innovation (2017).

[5] E. Sicard, *Introducing 7-nm FinFET technology in Microwind*, .

[6] M. L. Rieger, *Retrospective on VLSI value scaling and lithography*, Journal of Micro/Nanolithography, MEMS, and MOEMS **18** (2019), no. 4 040902.

[7] K.-S. Li, P.-G. Chen, T.-Y. Lai, C.-H. Lin, C.-C. Cheng, C.-C. Chen, Y.-J. Wei, Y.-F. Hou, M.-H. Liao, M.-H. Lee, *et. al.*, *Sub-60mV-swing negative-capacitance FinFET without hysteresis*, in *2015 IEEE International Electron Devices Meeting (IEDM)*, pp. 22–6, IEEE, 2015.

[8] M. Kobayashi, N. Ueyama, K. Jang, and T. Hiramoto, *Experimental study on polarization-limited operation speed of negative capacitance FET with ferroelectric HfO 2*, in *2016 IEEE International Electron Devices Meeting (IEDM)*, pp. 12–3, IEEE, 2016.

[9] Takaakitsurumi, S.-M. Num, Y.-B. Kil, and S. Wada, *High frequency measurements of PE hysteresis curves of PZT thin films*, Ferroelectrics **259** (2001), no. 1 43–48.

[10] D. Kwon, K. Chatterjee, A. J. Tan, A. K. Yadav, H. Zhou, A. B. Sachid, R. Dos Reis, C. Hu, and S. Salahuddin, *Improved subthreshold swing and short channel effect in FDSOI n-channel negative capacitance field effect transistors*, IEEE Electron Device Letters **39** (2017), no. 2 300–303.

[11] K. Ni, A. Saha, W. Chakraborty, H. Ye, B. Grisafe, J. Smith, G. B. Rayner, S. Gupta, and S. Datta, *Equivalent oxide thickness (EOT) scaling with hafnium zirconium oxide high-κ dielectric near morphotropic phase boundary*, in *2019 IEEE International Electron Devices Meeting (IEDM)*, pp. 7–4, IEEE, 2019.

[12] J. L. Moll, *Physical and Quantum Electronics*. New York (N.Y.) : McGraw-Hill, 1964.

[13] L. D. Landau and E. M. Lifshitz, *Quantum mechanics: non-relativistic theory*, vol. 3. Elsevier, 2013.

[14] A. C. Seabaugh and Q. Zhang, *Low-voltage tunnel transistors for beyond CMOS logic*, Proceedings of the IEEE **98** (2010), no. 12 2095–2110.

[15] J. M. Stork and R. D. Isaac, *Tunneling in base-emitter junctions*, IEEE transactions on electron devices **30** (1983), no. 11 1527–1534.

[16] R. B. Fair and H. W. Wivell, *Zener and avalanche breakdown in As-implanted low-voltage Si np junctions*, IEEE Transactions on Electron Devices **23** (1976), no. 5 512–518.

[17] F. Kesamanly, I. MALTSEV, D. Nasledov, L. Nikolaeva, M. Pivovarov, V. Skripkin, and I. UKHANOV, *Structure of the conduction band in indium arsenide(Indium arsenide crystals conduction band structure with various electron concentrations, determining electron mass dependence on concentration and temperature)*, Fizika i Tekhnika Poluprovodnikov **3** (1969) 1182–1187.

[18] M. Heller and R. Hamerly, *Hole transport in gallium antimonide*, Journal of applied physics **57** (1985), no. 10 4626–4632.

[19] Y. A. Goldberg and N. Schmidt, *Handbook series on semiconductor parameters*, vol **1** (1999) 191–213.

[20] V. Bougrov, M. Levinshtein, S. L. Rumyantsev, and A. Zubrilov, *Properties of Advanced Semiconductor Materials GaN, AlN, InN, BN, SiC, SiGe*, Bougrov, ME Levinshtein, SL Rumyantsev, A. Zubrilov.- NY: Wiley (2001) 1–50.

[21] G. Harbeke, *Optical properties of GaN*, RCA Review: A Technical Journal... **36** (1975) 163.

[22] M. A. Green, *Intrinsic concentration, effective densities of states, and effective mass in silicon*, Journal of Applied Physics **67** (1990), no. 6 2944–2954.

[23] P. Long, J. Z. Huang, M. Povolotskyi, G. Klimeck, and M. J. Rodwell, *High-Current Tunneling FETs With ($1\bar{1}0$) Orientation and a Channel Heterojunction*, IEEE Electron Device Letters **37** (2016), no. 3 345–348.

[24] J. Z. Huang, P. Long, M. Povolotskyi, G. Klimeck, and M. J. Rodwell, *P-type tunnel FETs with triple heterojunctions*, IEEE Journal of the Electron Devices Society **4** (2016), no. 6 410–415.

[25] P. Long, J. Z. Huang, M. Povolotskyi, G. Klimeck, and M. J. W. Rodwell, *High-Current Tunneling FETs With (¡inline-formula¿ ¡tex-math notation="LaTeX"¿$1\bar{1}0$ ¡/tex-math¿¡/inline-formula¿) Orientation and a Channel Heterojunction*, IEEE Electron Device Letters **37** (2016), no. 3 345–348.

[26] M. S. Ram and D. B. Abdi, *Dopingless PNPN tunnel FET with improved performance: Design and analysis*, Superlattices and Microstructures **82** (2015) 430–437.

[27] T. Kumari, P. Saha, D. K. Dash, and S. K. Sarkar, *Modeling of dual gate material hetero-dielectric strained PNPN TFET for improved ON current*, Journal of Materials Engineering and Performance **27** (2018), no. 6 2747–2753.

[28] D. B. Abdi and M. J. Kumar, *In-built N+ pocket pnpn tunnel field-effect transistor*, IEEE Electron Device Letters **35** (2014), no. 12 1170–1172.

[29] M. Rahimian and M. Fathipour, *Improvement of electrical performance in junctionless nanowire TFET using hetero-gate-dielectric*, Materials Science in Semiconductor Processing **63** (2017) 142–152.

[30] S. Yadav, D. Sharma, B. V. Chandan, M. Aslam, D. Soni, and N. Sharma, *A novel hetero-material gate-underlap electrically doped TFET for improving DC/RF and ambipolar behaviour*, Superlattices and Microstructures **117** (2018) 9–17.

[31] R. Narang, M. Saxena, R. Gupta, and M. Gupta, *Assessment of ambipolar behavior of a tunnel FET and influence of structural modifications*, JSTS: Journal of Semiconductor Technology and Science **12** (2012), no. 4 482–491.

[32] C. K. Pandey, A. Singh, and S. Chaudhury, *Effect of asymmetric gate–drain overlap on ambipolar behavior of double-gate TFET and its impact on HF performances*, Applied Physics A **126** (2020), no. 3 1–12.

[33] S. Sant, E. Memisevic, L.-E. Wernersson, and A. Schenk, *Impact of Non-idealities on the performance of InAs/(In) GaAsSb/GaSb tunnel FETs*, Composants nanoélectroniques (2018).

[34] A. Schenk, S. Sant, K. Moselund, H. Riel, E. Memisevic, and L.-E. Wernersson, *The impact of hetero-junction and oxide-interface traps on the performance of InAs/Si and InAs/GaAsSb nanowire tunnel FETs*, in *2017 International Conference on Simulation of Semiconductor Processes and Devices (SISPAD)*, pp. 273–276, IEEE, 2017.

[35] R. N. Sajjad, W. Chern, J. L. Hoyt, and D. A. Antoniadis, *Trap assisted tunneling and its effect on subthreshold swing of tunnel FETs*, IEEE Transactions on Electron Devices **63** (2016), no. 11 4380–4387.

[36] U. E. Avci, B. Chu-Kung, A. Agrawal, G. Dewey, V. Le, R. Rios, D. H. Morris, S. Hasan, R. Kotlyar, J. Kavalieros, *et. al.*, *Study of TFET non-ideality effects for determination of geometry and defect density requirements for sub-60mV/dec Ge TFET*, in *2015 IEEE International Electron Devices Meeting (IEDM)*, pp. 34–5, IEEE, 2015.

[37] A. Vandooren, D. Leonelli, R. Rooyackers, A. Hikavyy, K. Devriendt, M. Demand, R. Loo, G. Groeseneken, and C. Huyghebaert, *Analysis of trap-assisted tunneling in vertical Si homo-junction and SiGe hetero-junction tunnel-FETs*, Solid-State Electronics **83** (2013) 50–55.

[38] S. Sant, A. Schenk, K. Moselund, and H. Riel, *Impact of trap-assisted tunneling and channel quantization on InAs/Si hetero tunnel FETs*, in *2016 74th Annual Device Research Conference (DRC)*, pp. 1–2, IEEE, 2016.

[39] E. Memisevic, J. Svensson, M. Hellenbrand, E. Lind, and L.-E. Wernersson, *Vertical InAs/GaAsSb/GaSb tunneling field-effect transistor on Si with* $S = 48$ *mV/decade and* $I_{on} = 10\ \mu A/\mu m$ *for* $I_{off} = 1\ nA/\mu m$ *at* $V_{ds} = 0.3\ V$, in *2016 International Electron Devices Meeting (IEDM)*, pp. 19.1.1–19.1.4, IEEE, 2016.

[40] A. Alian, S. E. Kazzi, A. Verhulst, A. Milenin, N. Pinna, T. Ivanov, D. Lin, D. Mocuta, and N. Collaert, *Record 47 mV/dec top-down vertical nanowire InGaAs/GaAsSb tunnel FETs*, in *2018 IEEE Symposium on VLSI Technology*, pp. 133–134, IEEE, 2018.

[41] V. Nagavarapu, R. Jhaveri, and J. C. S. Woo, *The Tunnel Source (PNPN) n-MOSFET: A Novel High Performance Transistor*, IEEE Transactions on Electron Devices **55** (2008), no. 4 1013–1019.

[42] S. Koester, I. Lauer, A. Majumdar, J. Cai, J. Sleight, S. Bedell, P. Solomon, S. Laux, L. Chang, S. Koswatta, W. Haensch, P. Tomasini, and S. Thomas, *(Invited) Are Si/SiGe Tunneling Field-Effect Transistors a Good Idea?*, ECS Transactions **33** (2019), no. 6 357–361.

[43] G. Dewey, B. Chu-Kung, J. Boardman, J. M. Fastenau, J. Kavalieros, R. Kotlyar, W. K. Liu, D. Lubyshev, M. Metz, N. Mukherjee, P. Oakey, R. Pillarisetty, M. Radosavljevic, H. W. Then, and R. Chau, *Fabrication, characterization, and physics of III–V heterojunction tunneling Field Effect Transistors (H-TFET) for steep sub-threshold swing*, in *2011 International Electron Devices Meeting*, pp. 33.6.1–33.6.4, IEEE, 2011.

[44] W. Li, H. Liu, S. Wang, S. Chen, and Z. Yang, *Design of High Performance Si/SiGe Heterojunction Tunneling FETs with a T-Shaped Gate*, Nanoscale Research Letters **12** (2017), no. 1 1–8.

[45] P. Long, J. Z. Huang, M. Povolotskyi, D. Verreck, J. Charles, T. Kubis, G. Klimeck, M. J. Rodwell, and B. H. Calhoun, *A tunnel FET design for high-current, 120 mV operation*, in *2016 IEEE International Electron Devices Meeting (IEDM)*, pp. 30.2.1–30.2.4, IEEE, 2016.

[46] L. De Michielis, L. Lattanzio, K. E. Moselund, H. Riel, and A. M. Ionescu, *Tunneling and occupancy probabilities: How do they affect tunnel-FET behavior?*, IEEE electron device letters **34** (2013), no. 6 726–728.

[47] J. Z. Huang, P. Long, M. Povolotskyi, G. Klimeck, and M. J. Rodwell, *Sb-and Al-Free ultra-high-current tunnel FET designs*, in *2017 Fifth Berkeley Symposium on Energy Efficient Electronic Systems & Steep Transistors Workshop (E3S)*, pp. 1–3, IEEE, 2017.

[48] D. Todd, *Lois: "Save me, Dr. HBT – from those evil HFETs"*, in *Slideplayer*, p. 5, 2015.

[49] s. C. Abernathy, *Compound semiconductor growth by metallorganic molecular beam epitaxy (MOMBE)*, Materials Science and Engineering: R: Reports **14** (1995), no. 5 203–253.

[50] C. Abernathy, P. Wisk, S. Pearton, and F. Ren, *Mg doping of InP and InGaAs grown by metalorganic molecular beam epitaxy using bis-cyclopentadienyl magnesium*, Applied physics letters **62** (1993), no. 3 258–260.

[51] N. Nordell, P. Ojala, W. Van Berlo, G. Landgren, and M. K. Linnarsson, *Diffusion of Zn and Mg in AlGaAs/GaAs structures grown by metalorganic vapor-phase epitaxy*, Journal of applied physics **67** (1990), no. 2 778–786.

[52] Y. Gao, S. Godefroy, J. Benchimol, F. Alaoui, F. Alexandre, and K. Rao, *SIMS analysis of InP, GaAs and InGaAs layers grown by chemical beam epitaxy*, Surface and Interface Analysis **16** (1990), no. 1-12 36–40.

[53] J. Benchimol, J. Mba, A. Duchenois, B. Sermage, P. Launay, D. Caffin, M. Meghelli, and M. Juhel, *CBE growth of carbon doped InGaAs/InP HBTs for 25 Gbit/s circuits*, Journal of crystal growth **188** (1998), no. 1-4 349–354.

[54] A. Ougazzaden, J. Holavanahalli, M. Geva, and L. E. Smith, *Carbon doping of InAlAs in LP-MOVPE using CBr4*, Journal of crystal growth **221** (2000), no. 1-4 66–69.

[55] J. Yan, G. Ru, Y. Gong, and F.-S. Choa, *Study of p-type carbon doping on In0. 53Ga0. 47As, In0. 52Al0. 2Ga0. 28As & In0. 52Al0. 48As*, in *Proceedings of SPIE-The international Society for Optical Engineering*, pp. 446–449.

[56] V. Chobpattana, T. E. Mates, W. J. Mitchell, J. Y. Zhang, and S. Stemmer, *Influence of plasma-based in-situ surface cleaning procedures on HfO2/In0. 53Ga0. 47As gate stack properties*, Journal of Applied Physics **114** (2013), no. 15 154108.

[57] V. Chobpattana, T. E. Mates, J. Y. Zhang, and S. Stemmer, *Scaled ZrO2 dielectrics for In0. 53Ga0. 47As gate stacks with low interface trap densities*, Applied Physics Letters **104** (2014), no. 18 182912.

[58] S. Lee, V. Chobpattana, C.-Y. Huang, B. J. Thibeault, W. Mitchell, S. Stemmer, A. Gossard, and M. Rodwell, *Record I on (0.50 mA/μm at V DD= 0.5 V and I off= 100 nA/μm) 25 nm-gate-length ZrO 2/InAs/InAlAs MOSFETs*, in *2014 Symposium on VLSI Technology (VLSI-Technology): Digest of Technical Papers*, pp. 1–2, IEEE, 2014.

[59] J. Matthews and A. Blakeslee, *Defects in epitaxial multilayers: I. Misfit dislocations*, Journal of Crystal growth **27** (1974) 118–125.

[60] J. Matthews and A. Blakeslee, *Defects in epitaxial multilayers: II. Dislocation pile-ups, threading dislocations, slip lines and cracks*, Journal of Crystal Growth **29** (1975), no. 3 273–280.

[61] S. Steiger, M. Povolotskyi, H.-H. Park, T. Kubis, and G. Klimeck, *NEMO5: A parallel multiscale nanoelectronics modeling tool*, IEEE Transactions on Nanotechnology **10** (2011), no. 6 1464–1474.

[62] Y. Tan, M. Povolotskyi, T. Kubis, T. B. Boykin, and G. Klimeck, *Transferable tight-binding model for strained group IV and III-V materials and heterostructures*, Physical Review B **94** (2016), no. 4 045311.

[63] T. A. Ameen, H. Ilatikhameneh, J. Z. Huang, M. Povolotskyi, R. Rahman, and G. Klimeck, *Combination of equilibrium and nonequilibrium carrier statistics into an atomistic quantum transport model for tunneling heterojunctions*, IEEE Transactions on Electron Devices **64** (2017), no. 6 2512–2518.

[64] C.-Y. Chen, H. Ilatikhameneh, J. Z. Huang, G. Klimeck, and M. Povolotskyi, *Impact of body thickness and scattering on III–V triple heterojunction TFET modeled with atomistic mode-space approximation*, IEEE Transactions on Electron Devices **67** (2020), no. 8 3478–3485.

[65] H. Ilatikhameneh, T. A. Ameen, C. Chen, G. Klimeck, and R. Rahman, *Sensitivity challenge of steep transistors*, IEEE Transactions on Electron Devices **65** (2018), no. 4 1633–1639.

[66] R.-H. Yan, A. Ourmazd, and K. F. Lee, *Scaling the Si MOSFET: From bulk to SOI to bulk*, IEEE Transactions on Electron Devices **39** (1992), no. 7 1704–1710.

[67] D. Hisamoto, W.-C. Lee, J. Kedzierski, H. Takeuchi, K. Asano, C. Kuo, E. Anderson, T.-J. King, J. Bokor, and C. Hu, *FinFET-a self-aligned double-gate MOSFET scalable to 20 nm*, IEEE transactions on electron devices **47** (2000), no. 12 2320–2325.

[68] G. Bae, D.-I. Bae, M. Kang, S. Hwang, S. Kim, B. Seo, T. Kwon, T. Lee, C. Moon, Y. Choi, *et. al.*, *3nm GAA technology featuring multi-bridge-channel FET for low power and high performance applications*, in *2018 IEEE International Electron Devices Meeting (IEDM)*, pp. 28–7, IEEE, 2018.

[69] J. E. Mahan, *Physical vapor deposition of thin films.* 2000.

[70] P. O. Oviroh, R. Akbarzadeh, D. Pan, R. A. M. Coetzee, and T.-C. Jen, *New development of atomic layer deposition: processes, methods and applications*, Science and technology of advanced materials **20** (2019), no. 1 465–496.

[71] R. W. Johnson, A. Hultqvist, and S. F. Bent, *A brief review of atomic layer deposition: from fundamentals to applications*, Materials today **17** (2014), no. 5 236–246.

[72] J. Musschoot, Q. Xie, D. Deduytsche, S. Van den Berghe, R. Van Meirhaeghe, and C. Detavernier, *Atomic layer deposition of titanium nitride from TDMAT precursor*, Microelectronic Engineering **86** (2009), no. 1 72–77.

[73] W. Lee, C. H. An, S. Yoo, W. Jeon, M. J. Chung, S. H. Kim, and C. S. Hwang, *Electrical Properties of ZrO2/Al2O3/ZrO2-Based Capacitors with TiN, Ru, and TiN/Ru Top Electrode Materials*, physica status solidi (RRL)–Rapid Research Letters **12** (2018), no. 10 1800356.

[74] S.-S. Yim, D.-J. Lee, K.-S. Kim, S.-H. Kim, T.-S. Yoon, and K.-B. Kim, *Nucleation kinetics of Ru on silicon oxide and silicon nitride surfaces deposited by atomic layer deposition*, Journal of Applied Physics **103** (2008), no. 11 113509.

[75] J. Heo, S. Y. Lee, D. Eom, C. S. Hwang, and H. J. Kim, *Enhanced Nucleation Behavior of Atomic-Layer-Deposited Ru Film on Low-k Dielectrics Afforded by UV-O3 Treatment*, Electrochemical and Solid State Letters **11** (2007), no. 2 G5.

[76] Z. Mi, C. Tao, T. Jing-Jing, R. Guo-Ping, J. Yu-Long, L. Ran, and Q. Xin-Ping, *Effect of pretreatment of TaN substrates on atomic layer deposition growth of Ru thin films*, Chinese Physics Letters **24** (2007), no. 5 1400.

[77] T.-L. Wu, D. Marcon, B. Bakeroot, B. De Jaeger, H. Lin, J. Franco, S. Stoffels, M. Van Hove, R. Roelofs, G. Groeseneken, *et. al.*, *Correlation of interface states/border traps and threshold voltage shift on AlGaN/GaN metal-insulator-semiconductor high-electron-mobility transistors*, Applied Physics Letters **107** (2015), no. 9 093507.

[78] S.-H. Kim, *Highly-conformal Ru Thin Films by Atomic Layer Deposition Using Novel Zero-valent Ru Metallorganic Precursors and O_2 for Nano-scale Devices*, Electrical & Electronic Materials **28** (2015), no. 2 25–33.

[79] D. S. Kwon, C. H. An, S. H. Kim, D. G. Kim, J. Lim, W. Jeon, and C. S. Hwang, *Atomic layer deposition of Ru thin films using (2, 4-dimethyloxopentadienyl)(ethylcyclopentadienyl) Ru and the effect of ammonia treatment during the deposition*, Journal of Materials Chemistry C **8** (2020), no. 21 6993–7004.

[80] R. Engel-Herbert, Y. Hwang, and S. Stemmer, *Comparison of methods to quantify interface trap densities at dielectric/III-V semiconductor interfaces*, Journal of applied physics **108** (2010), no. 12 124101.

[81] M. Zhang, W. Chen, S.-J. Ding, X.-P. Wang, D. W. Zhang, and L.-K. Wang, *Investigation of atomic-layer-deposited ruthenium nanocrystal growth on SiO_2 and Al_2O_3 films*, Journal of Vacuum Science & Technology A: Vacuum, Surfaces, and Films **25** (2007), no. 4 775–780.

[82] J. Soethoudt, F. Grillo, E. A. Marques, J. R. van Ommen, Y. Tomczak, L. Nyns, S. Van Elshocht, and A. Delabie, *Diffusion-Mediated Growth and Size-Dependent Nanoparticle Reactivity during Ruthenium Atomic Layer Deposition on Dielectric Substrates*, Advanced Materials Interfaces **5** (2018), no. 24 1800870.

[83] L. G. Wen, P. Roussel, O. V. Pedreira, B. Briggs, B. Groven, S. Dutta, M. I. Popovici, N. Heylen, I. Ciofi, K. Vanstreels, *et. al.*, *Atomic layer deposition of ruthenium with TiN interface for sub-10 nm advanced interconnects beyond copper*, ACS applied materials & interfaces **8** (2016), no. 39 26119–26125.

[84] C.-Y. Chen, H.-Y. Tseng, H. Ilatikhameneh, T. A. Ameen, G. Klimeck, M. J. Rodwell, and M. Povolotskyi, *Doping Profile Engineered Triple Heterojunction*

TFETs With 12-nm Body Thickness, IEEE Transactions on Electron Devices **68** (2021), no. 6 3104–3111.

[85] Z. Lu, C. Lagarde, E. Sacher, J. Currie, and A. Yelon, *A surface analytical study of GaAs (100) cleaning procedures*, Journal of Vacuum Science & Technology A: Vacuum, Surfaces, and Films **7** (1989), no. 3 646–650.

[86] D. Aspnes and A. Studna, *Chemical etching and cleaning procedures for Si, Ge, and some III-V compound semiconductors*, Applied Physics Letters **39** (1981), no. 4 316–318.

[87] Y. Sun, Z. Liu, F. Machuca, P. Pianetta, and W. E. Spicer, *Optimized cleaning method for producing device quality InP (100) surfaces*, Journal of Applied Physics **97** (2005), no. 12 124902.

[88] W. J. Mitchell, B. J. Thibeault, D. D. John, and T. E. Reynolds, *Highly selective and vertical etch of silicon dioxide using ruthenium films as an etch mask*, Journal of Vacuum Science & Technology A: Vacuum, Surfaces, and Films **39** (2021), no. 4 043204.

[89] C. Huang, S. Lee, V. Chobpattana, S. Stemmer, A. Gossard, B. Thibeault, W. Mitchell, and M. Rodwell, *Low Power III–V InGaAs MOSFETs featuring InP recessed source/drain spacers with I on= 120 μA/μm at I off= 1 nA/μm and V DS= 0.5 V*, in *2014 IEEE International Electron Devices Meeting*, pp. 25–4, IEEE, 2014.

[90] S. Cwik, K. N. Woods, M. J. Saly, T. J. Knisley, and C. H. Winter, *Thermal atomic layer deposition of ruthenium metal thin films using nonoxidative coreactants*, Journal of Vacuum Science & Technology A: Vacuum, Surfaces, and Films **38** (2020), no. 1 012402.

[91] S. Yeo, S.-H. Choi, J.-Y. Park, S.-H. Kim, T. Cheon, B.-Y. Lim, and S. Kim, *Atomic layer deposition of ruthenium (Ru) thin films using ethylbenzen-cyclohexadiene Ru (0) as a seed layer for copper metallization*, Thin Solid Films **546** (2013) 2–8.

[92] J. Lu and J. W. Elam, *Low temperature ABC-type Ru atomic layer deposition through consecutive dissociative chemisorption, combustion, and reduction steps*, Chemistry of Materials **27** (2015), no. 14 4950–4956.

[93] S. N. Chopra, M. F. Vos, M. A. Verheijen, J. G. Ekerdt, W. M. Kessels, and A. J. Mackus, *Atomic layer deposition of ruthenium using an ABC-type process: Role of oxygen exposure during nucleation*, Journal of Vacuum Science & Technology A: Vacuum, Surfaces, and Films **38** (2020), no. 6 062402.

[94] M. Fujimatsu, H. Saito, and Y. Miyamoto, *71 mV/dec of sub-threshold slope in vertical tunnel field-effect transistors with GaAsSb/InGaAs heterostructure*, in *2012 International Conference on Indium Phosphide and Related Materials*, pp. 25–28, IEEE, 2012.

[95] A. Alian, S. El Kazzi, A. Verhulst, A. Milenin, N. Pinna, T. Ivanov, D. Lin, D. Mocuta, and N. Collaert, *Record 47 mV/dec top-down vertical nanowire InGaAs/GaAsSb tunnel FETs*, in *2018 IEEE Symposium on VLSI Technology*, pp. 133–134, IEEE, 2018.

[96] K. Tomioka, H. Gamo, J. Motohisa, and T. Fukui, *Vertical Gate-All-Around Tunnel FETs Using InGaAs Nanowire/Si with Core-Multishell Structure*, in *2020 IEEE International Electron Devices Meeting (IEDM)*, pp. 21–1, IEEE, 2020.

[97] C. Convertino, C. B. Zota, H. Schmid, D. Caimi, L. Czornomaz, A. M. Ionescu, and K. E. Moselund, *A hybrid III–V tunnel FET and MOSFET technology platform integrated on silicon*, Nature Electronics **4** (2021), no. 2 162–170.

[98] D. Cutaia, K. E. Moselund, H. Schmid, M. Borg, A. Olziersky, and H. Riel, *Complementary III–V heterojunction lateral NW tunnel FET technology on Si*, in *2016 IEEE Symposium on VlSI Technology*, pp. 1–2, IEEE, 2016.

[99] S. Mookerjea, D. Mohata, R. Krishnan, J. Singh, A. Vallett, A. Ali, T. Mayer, V. Narayanan, D. Schlom, A. Liu, *et. al.*, *Experimental demonstration of 100nm channel length In 0.53 Ga 0.47 As-based vertical inter-band tunnel field effect transistors (TFETs) for ultra low-power logic and SRAM applications*, in *2009 IEEE international electron devices meeting (IEDM)*, pp. 1–3, IEEE, 2009.

[100] A. Goswami, S. T. Š. Brunelli, B. Markman, A. A. Taylor, H.-Y. Tseng, K. Mukherjee, M. Rodwell, J. Klamkin, and C. J. Palmstrøm, *Controlling facets and defects of InP nanostructures in confined epitaxial lateral overgrowth*, Physical Review Materials **4** (2020), no. 12 123403.

[101] S. T. Šuran Brunelli, B. Markman, A. Goswami, H.-Y. Tseng, S. Choi, C. Palmstrøm, M. Rodwell, and J. Klamkin, *Selective and confined epitaxial growth development for novel nano-scale electronic and photonic device structures*, Journal of Applied Physics **126** (2019), no. 1 015703.

[102] S. T. Suran Brunelli, A. Goswami, B. Markman, H.-Y. Tseng, M. Rodwell, C. Palmstrøm, and J. Klamkin, *Horizontal heterojunction integration via template-Assisted selective epitaxy*, Crystal Growth & Design **19** (2019), no. 12 7030–7035.

[103] R. Klockenbrink, E. Peiner, H.-H. Wehmann, and A. Schlachetzki, *Wet Chemical Etching of Alignment V-Grooves in (100) InP through Titanium or In0. 53Ga0. 47As Masks*, Journal of the Electrochemical Society **141** (1994), no. 6 1594.

[104] P. Eliáš, I. Kostič, J. Šoltỳs, and S. Hasenöhrl, *Wet-etch bulk micromachining of (100) InP substrates*, Journal of Micromechanics and Microengineering **14** (2004), no. 8 1205.

[105] V. Jain, *Indium Phosphide DHBTs in a Refractory Emitter Process for THz Electronics*. University of California, Santa Barbara, 2011.

[106] A. Baraskar, *Development of Ultra-Low Resistance Ohmic Contacts for indium gallium arsenide/indium phosphide HBTs*. University of California, Santa Barbara, 2011.

[107] M. Nakahara, S. Tsunekawa, K. Watanabe, T. Arai, T. Yunogami, and K. Kuroki, *Etching technique for ruthenium with a high etch rate and high selectivity using ozone gas*, Journal of Vacuum Science & Technology B: Microelectronics and Nanometer Structures Processing, Measurement, and Phenomena **19** (2001), no. 6 2133–2136.

[108] E.-J. Lee, J.-W. Kim, and W.-J. Lee, *Reactive Ion Etching Mechanism of RuO2 Thin Films in Oxygen Plasma with the Addition of CF4, Cl2, and N2*, Japanese journal of applied physics **37** (1998), no. 5R 2634.

[109] S. Seok, *Fabrication and modeling of nitride thin-film encapsulation based on anti-adhesion-assisted transfer technique and nitride/BCB bilayer wrinkling*, IEEE Transactions on Components, Packaging and Manufacturing Technology **6** (2016), no. 9 1301–1307.