Antimonide-based III-V Multigate Transistors

by

Wenjie Lu

B.S. Electrical Engineering,
S.M., Electrical Engineering and Computer Science

Submitted to the Department of Electrical Engineering and Computer Science in partial fulfillment of the requirements for the degree of

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

June 2018

Signature of author: _____

Department of Electrical Engineering and Computer Science

May 23, 2018

Certified by: _____

Jesús A. del Alamo

Professor of Electrical Engineering and Computer Science
Thesis Supervisor

Accepted by: _____

Leslie A. Kolodziejski

Professor of Electrical Engineering and Computer Science

Chair, Committee for Graduate Students

To Chenxi and My Parents

Antimonide-based III-V Multi-gate Transistors

by Wenjie Lu

Submitted to the Department of Electrical Engineering and Computer Science
in partial fulfillment of the requirements for the degree of
Doctor of Philosophy

ABSTRACT

As Si CMOS technology advances, alternative channel materials are under extensive investigation to replace or augment Si in future generations of nanoelectronics. III-V compound semiconductors, such as InGaAs and InGaSb are promising candidates as channel materials for MOSFETs as a result of their extraordinary transport properties. In the past few years, rapid growth in the research of InGaAs n-channel multi-gate MOSFETs has taken place. However, progress in the InGaSb p-channel device research has remained stagnant. In this thesis, InGaSb multi-gate transistor technology has been pioneered and the first InGaSb FinFET has been demonstrated.

Critical technological challenges for realizing InGaSb FinFETs have been overcome. First, a dry etching technique of heterostructures containing antimonide-based compounds has been developed. Etched fins and vertical nanowires show smooth, vertical sidewalls, high aspect ratio, and compatibility with the InGaAs system. Second, a novel antimonide-compatible digital etch technique has been developed which can improve fin sidewall quality and device performance. Lastly, ohmic contacts have been investigated to reduce the parasitic source-drain resistance. The developed contact system delivers a record low contact resistivity.

With the integration of the newly developed technologies, InGaSb p-channel FinFETs are demonstrated for the first time. Three generations of InGaSb FinFETs are fabricated following an optimization path for device design and process technology. The most aggressively scaled InGaSb FinFETs, with a minimum fin width of 10 nm and channel height of 23 nm, have achieved a maximum transconductance per device footprint of 704 μS/μm, a record value for any existing antimonide-based p-channel FETs. In addition, the fabricated FinFETs have been electrically characterized, and device properties such as scaling behavior, impact of channel strain, and OFF-state leakage current have been studied. The work in this thesis has pushed significantly the state-of-the-art of antimonide-based electronic device technology.

Thesis supervisor: Jesús A. del Alamo
Professor of Electrical Engineering and Computer Science

Acknowledgments

This thesis is not possible without the support, guidance, and love from many people.

First, I want to thank my thesis advisor, Prof. Jesús A. del Alamo, for his six years of supervision. Jesús has spent a great amount of time to guide me during the course of my master and doctoral study. He is always willing to spend time with me to dig into semiconductor physics and experimental details, whenever I have questions. He is patient and kind, supporting me even at the most difficult (and seemingly impossible) time during my study at MIT. He taught me how to do research, and his high standard and thoroughness towards research is a priceless gift I have received from him. Jesús has shown me how to be a man of humility and integrity. He has shown me an example of research career, not to be driven by data or paper, but to pursue scientific truth and enjoy discovery. I am very privileged to be his student.

My gratitude also goes to my thesis committee members, Prof. Tomás Palacios and Prof. Max Shulaker. I am always inspired by their passion towards electronic research. They have given me many insightful suggestions that are critical to the completion of this thesis.

I want to thank other professors and researchers with whom I have worked together: Dr. Brian Bennett and Dr. Shawn Mack at Naval Research Laboratory, Dr. Jin Dong Song at Korean Institute of Science and Technology, Prof. Steven George and Dr. Younghee Lee at University of Colorado – Boulder, Dr. Nerissa Draeger at Lam Reseasrch, Dr. Charles Sullivan and Dr. Jin Kim at Sandia National Laboratories, Dr. Salim El Kazzi and Dr. Ali Alian at imec, Dr. Xiabing Lou at Harvard University, and Dr. Ryan Iutzi and Prof. Eugene Fitzgerald at MIT. Their knowledge and hard work have made this thesis possible. Without their expertise and help, I would not even have any material to work with!

Thanks to all the friends at the del Alamo group members, Dr. Alon Vardi (my long-term officemate), Xiaowei Cai, Xin Zhao, Yufei Wu, Shireen Warnock, Alex Guo, Jianqiang Lin, Luke Guo, Donghyun Jin, Ethan Lee, Alex Lednev, and Lisa Kong. I am especially grateful to Alon. He has generously taught me so much about semiconductor physics, and stayed with me countless

nights in the cleanroom, struggling together on our processes. Also, I want to thanks my friends on the sixth floor of building 39, Tao Yu, Winston Chern, Sameer Jogelaker, Min Sun, Jie Hu, Ahmad Zubair, Xu Zhang, Yuhao Zhang, Lili Yu, Bin Lu, Nadim Chowdbhury, Ujwal Radhakrishna, Ivan Chan, Rushabh Shah, and Cheng Wang. We have spent great time together.

I also want to thank many staff at MTL. Thanks to Debroah Hodges-Pabon, Mary O'Neil, Elizabeth Kubicki, Joseph Baylon, Valerie DiNardo, Sherene Aram, Mara Karapetian, Ludmila Leoparde, and William Maloney. Thank you for making MTL a wonderful community. I am very fortunate to work with many of the staff during the organization of MARC.

I am grateful to the staff in the cleanrooms, too. Thanks to Vicky Diadiuk, Dennis Ward, Dave Terry, Eric Lim, Robert Bicchieri, Donal Jamieson, Bernard Alamariu, Paul Tierney, Mark Mondol, Paul McGrath and Whitney Hess. From you, I learned much about fabrication, from the very basic to the black magic.

I want to thank some people who are intimate and very special in my life. I am forever grateful to Dr. Ce Liu and Irene Liaw for their six years of love and guidance. Because of them, I can come to know God and know the truth. I want to thank my best friends at MIT, Donglai Wei, Wengong Jin and Peilun Dai, for their love and encouragement.

This thesis is impossible without the love of my parents and grandparents. I can be here simply because of their sacrifices. Last, I want to thank my wife, Chenxi, for her unconditional love, understanding and support over the years.

This work was sponsored by the SRC, DTRA, KIST, Lam Research, NSF, and Samsung.

Contents

List of Figures

List of Tables

CHAPTER 1. INTRODUCTION

1.1 Introduction to CMOS scaling

Since 1965, Moore's law has charted the advancement of semiconductor integrated circuits. The scaling of silicon transistor dimensions into the nanometer size drives the exponential increase in the density and performance of complementary metal-oxide-semiconductor (CMOS) technology. For more than fifty years, the microelectronic industry has followed extremely successfully the motto ‘smaller is better’. The transistor count of integrated circuit chips doubles approximately every two years, and the improvement in switching energy, switching speed and cost per transistor improves exponentially as MOSFETs have decreased in size [1]–[4]. Fig. 1-1 depicts such geometric growth of transistor count in the past fifty years [5].

However, it has been predicted that conventional transistor scaling would slow down significantly and eventually end around 2025 [6], [7], when the physical limits of silicon-based metal-oxide-semiconductor field-effect transistors (MOSFETs) would be reached [8]. In fact, as early as 2005, it has been argued by Gordon Moore himself that Moore’s law would end by then because of the exponential nature of the growth. In recent years, there have been increasing expressions of concern and arguments on whether Moore’s law will die soon. Nevertheless, facing unprecedented

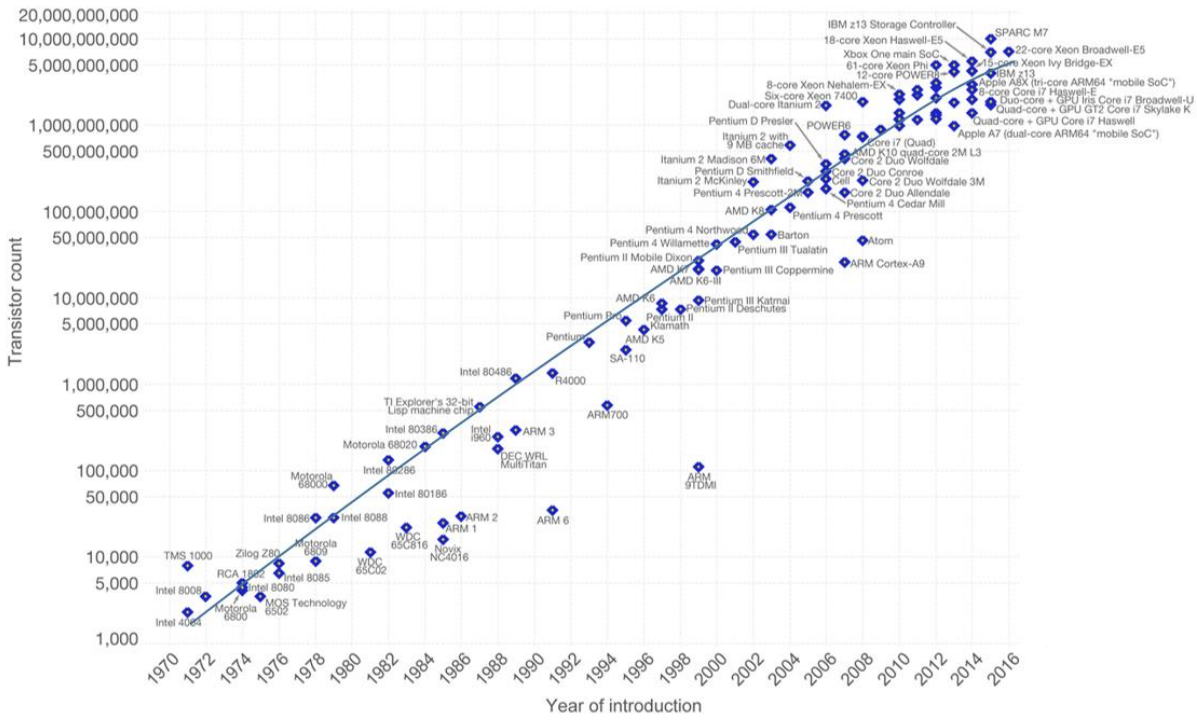


Fig. 1-1: Transistor count in integrated processors from 1971-2016, illustrating Moore's Law. Figure adapted from [5].

challenges, researchers worldwide have been pushed to be evermore creative. Extraordinary innovations have been made in CMOS research in recent years and Moore's law is still marching forward.

To understand CMOS scaling and the associated challenges, one needs to identify that in logic chips, transistors are used as switches, and the goal is to switch between the ON state and OFF state with maximum speed but minimum energy. The switching delay of a MOSFET is approximately:

$$\tau_d \sim \frac{C_G V_{DD}}{I_{ON}} \tag{1-1}$$

And the switching energy is approximately:

$$E_d \sim C_G V_{DD}^2 \tag{1-2}$$

τ_d, the delay time, is proportional to C_G, the total gate capacitance, and V_{DD}, the supply voltage, and inversely proportional to I_{ON}, the ON-state current. Fig. 1-2 illustrates the historic scaling behavior of V_{DD} and ON-state current density (I_{ON}/W) [9]. In the early years of CMOS scaling, V_{DD} remained constant and I_{ON} increased at a rate of roughly 37% per generation, which can be referred to as a constant-voltage scaling path. Beyond 2000, the electric field in the transistor became intolerable so that V_{DD} (along with threshold voltage V_T) had to decrease This heralded the era of constant-field scaling. As a result, the increase in current density slowed down to ~20% per generation and eventually reached a plateau (and is even dropping more lately). As V_{DD} continues to drop, obtaining increased performance becomes harder and harder.

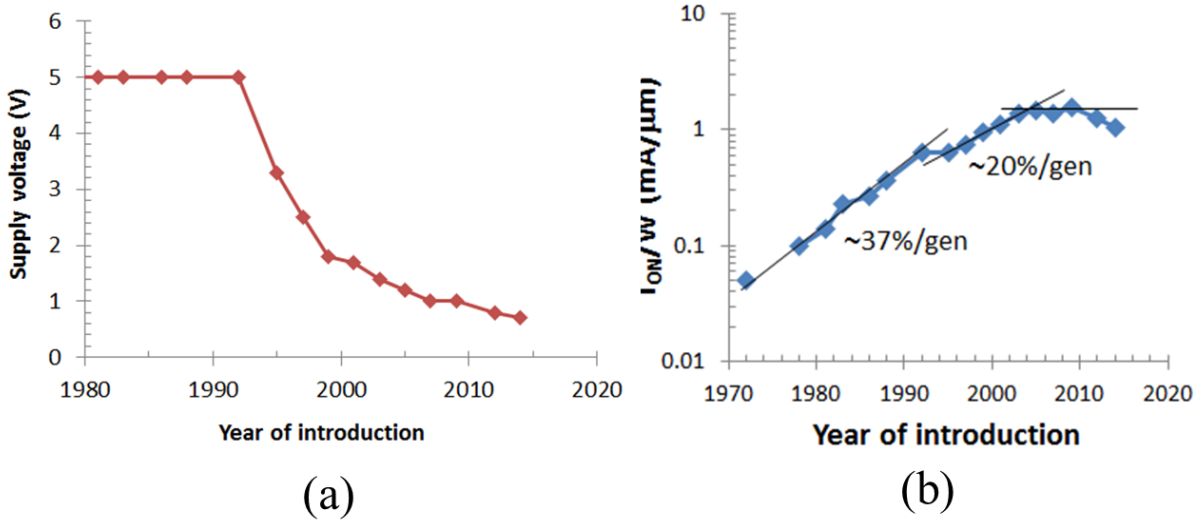


Fig. 1-2: (a) Supply voltage scaling of CMOS technologies. (b) ON-state current density scaling for n-channel Si MOSFETs at nominal operation voltage.

The above discussion is a simplistic view of CMOS scaling, and a comprehensive and insightful study of the scaling properties can be found in [9]. Currently, MOSFET scaling can be best described as power-constrained scaling, as the dynamic power density has become as high as 100 W/cm^2 [10]. Further transistor scaling requires more reduction in V_{DD} and hence greater compromises in device performance. In addition, as the channel length of Si MOSFETs has entered the nanometer-scale, short-channel effects have become severe, undermining the scaling benefits. Therefore, future CMOS technology requires innovations in both materials and device structures.

1.2 Multi-gate MOSFETs

As discussed earlier, the driving force behind transistor evolution is the scalability of Si CMOS. Footprint scaling demands all device dimensions to shrink together while delivering maximum performance at reduced voltage. This is a strict demand as device dimensions shrink into the sub-micron regime, where short-channel effects (SCEs) in MOSFETs become severe and catastrophic. SCEs, include mobility degradation, drain-induced barrier lowering (DIBL), degradation of subthreshold swing, leakage, and device variability. To resolve the serious challenges of planar CMOS, which cannot maintain acceptable SCEs at deeply-scaled technology nodes anymore, innovations in device structure need to be made. The evolution of device structures can be depicted in Fig. 1-3. To mitigate SCEs, extremely-thin-body-silicon-on-insulator (ET-SOI) MOSFET and multi-gate transistor structures have been proposed to enhance the electrostatic control of the gate to the channel. In ET-SOI, the channel is made sufficiently thin so that it is fully depleted. In multi-gate transistors, the channel is under control of more than one gate. For example, in a FinFET (double-gate MOSFET), the fin channel is fully depleted when it is narrow enough, so that the

entire channel is controlled by the gates on both sides. These non-planar MOSFET designs in effect shorten the characteristic electrostatic length of the device:

$$\lambda_{non-planar} = \sqrt{\frac{\epsilon_s}{\epsilon_{ox}} x_{ox} \frac{x_{Si}}{n}} = \lambda_{planar} \sqrt{\frac{x_{Si}}{n \cdot x_d}} \tag{1-3}$$

where ε_s and ε_{ox} are the electron permittivity of channel and gate oxide, x_{Si} and x_{ox} is the width or thickness of the channel and gate oxide, x_d is depletion width of the ideal planar MOSFET, and n is the gate number. The gate number is 1 for ET-SOI, 2 for FinFETs, 3 for Tri-gate MOSFETs, and 4 for gate-all-around (GAA) MOSFETs, respectively. Therefore, alternative designs can reduce the characteristic electrostatic length, $\lambda_{planar,}$ and in other words, mitigate short-channel effects.

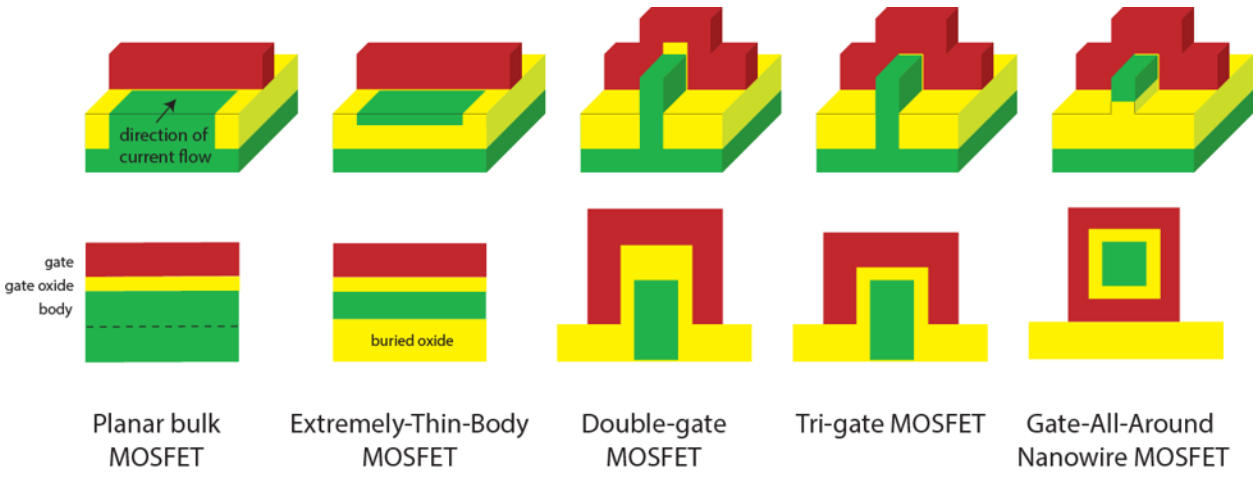


Fig. 1-3: Evolution of MOSFET structural designs with increasing electrostatic gate control of the channel from left to right, reproduced from [14].

The multi-gate designs bring more complexity to device fabrication, as the channel is no longer planar but three-dimensional. However, much progress has been made in research and industry

along this path, and tri-gate MOSFETs have already been the mainstream of today's logic transistors. It is believed that futures innovations in device technology will enable more advanced device structures, such as GAA nanowire MOSFETs and even vertical nanowire (VNW) MOSFETs.

1.3 Advantages of III-V channel MOSFETs

The other possible solution to maintaining or even improving transistor performance while continuing to reduce the supply voltage is to introduce alternative channel materials. In recent years, much research has been made on several novel materials, such as 2-D materials (graphene, MoS_2) [11], carbon nanotubes [12], germanium [13], gallium-nitride, and III-V compound semiconductors [1], [14]. In this regard, III-V compound semiconductors have been regarded as one of the most attractive and feasible candidates to replace Si [1]. III-V compound semiconductors, especially the arsenide-based ($In_xGa_{1-x}As$) and antimonide-based ($In_xGa_{1-x}Sb$) compounds, are known for their superior electron and hole transport properties over silicon. III-V compound semiconductors are advantageous over other candidates, because the fabrication process of III-V transistors is mostly compatible with present Si CMOS process. Moreover, GaAs and InP are the most advanced in terms of large wafer manufacturing, which can be achieved on wafer-scale substrates. Therefore, in the last few years, increasing research efforts have been devoted to the development of III-V nanometer-scale logic transistors.

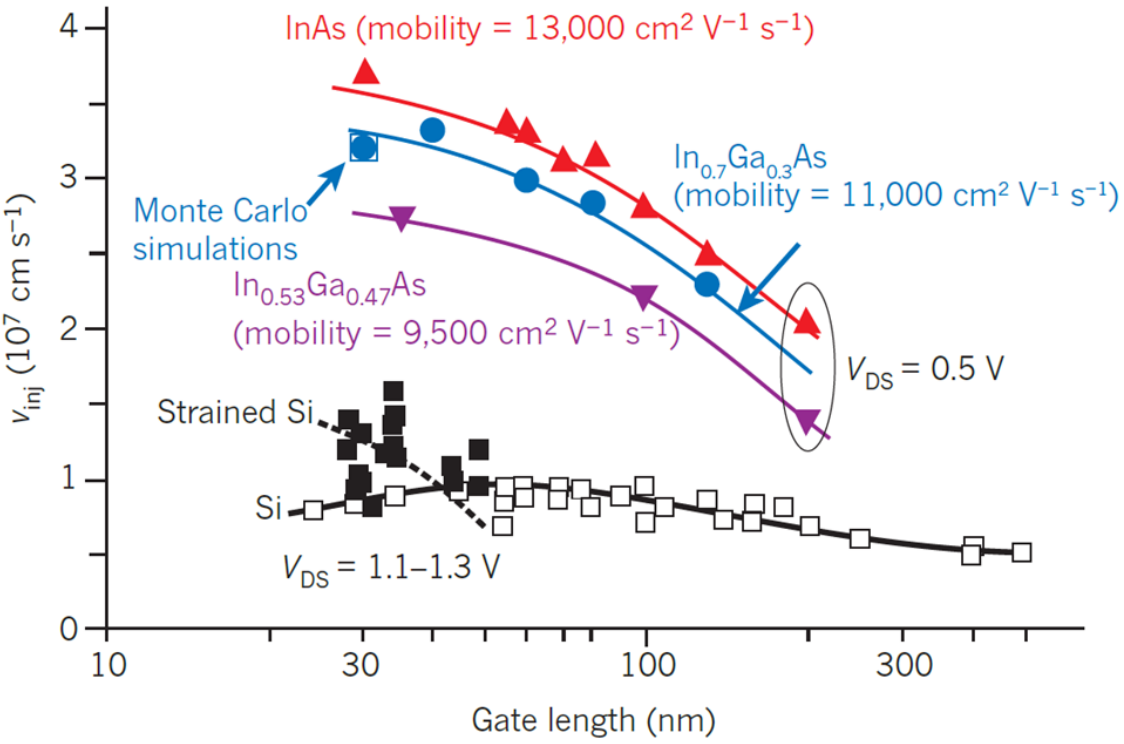


Fig. 1-4: Electron injection velocity in Si and strained-Si MOSFETs, InGaAs and InAs HEMTs with different channel compositions, as a function of gate length. Figure reproduced from [1].

For modern deeply-scaled MOSFETs, according to the virtual source model [4], [15], I_{ON} can be described as:

$$I_{ON} = v_{inj} \cdot Q_{inj} \cong v_{inj} \cdot C_{inj} \cdot (V_{DD} - V_T) \quad (4)$$

, where v_{inj} is the source injection velocity, Q_{inj} is the charge density at the virtual source, and (V_{DD} – V_T) is the gate overdrive. Fig. 1-4 illustrates the injection velocity extracted from Si and strained Si MOSFETs at V_{DS} of 1.1-1.3 V, along with the injection velocity of InGaAs and InAs high electron mobility transistors (HEMTs) at V_{DS} = 0.5 V, as a function of gate length. For Si

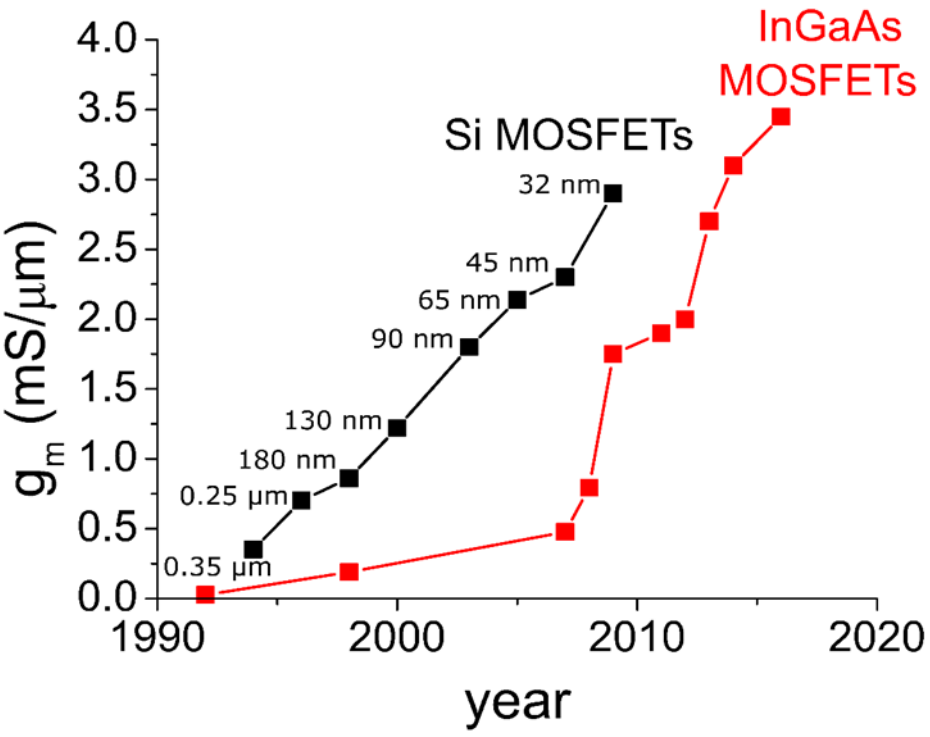


Fig. 1-5: Transconductance of Intel's Si MOSFETs and InGaAs MOSFETs vs. year of demonstration.

MOSFETs, v_{inj} maximizes at $1 \cdot 10^7$ cm/s, and reaches nearly $1.5 \cdot 10^7$ cm/s for strained Si. For InGaAs and InAs at less than half the operating voltage, v_{inj} can be more than twice or three-times that of Si MOSFETs. This is attributed to the lower electron effective masses of $In_xGa_{1-x}As$. In the ballistic limit, the injection velocity is closely related to the carrier mobility as $v_{inj} \sim \mu^{\alpha}$, where $\alpha \leq 0.5$ [4]. Therefore, to the first order, III-V channels with a higher mobility promise to provide higher v_{inj} in short channel devices. Fig. 1-5 shows the evolution of transconductance, g_m, of Intel's Si MOSFETs and InGaAs MOSFETs, as a function of year of demonstration. In recent years, rapid progress has been made in InGaAs MOSFETs, demonstrating a maximum g_m of 3.45 mS/µm, with L_g = 70 nm and V_{DS} = 0.5 V [16], [17], exceeding that Si's 32 nm node, which is the last node

before Intel introduced 3-D Tri-gate transistors. This progress shows that III-V MOSFETs have great potential for future logic and high-speed circuits even at lower operating voltage.

1.4 P-Channel MOSFETs

CMOS technology requires high performance from both NMOS and PMOS transistors. Because of the lower carrier mobility of holes, much research efforts in p-type MOSFETs are needed in order to match the performance of n-type MOSFETs. In unstrained Si, the current density of the PMOS transistor has been about one-third that of the NMOS transistor, and circuit designers learned to work with this imbalance. On the other hand, because of the development of strain engineering, the Si PMOS has been able to catch up with the performance of the NMOS. In state-of-the-art Si CMOS with 7^{th}-generation strained Si, I_{ON} of PMOS transistors has improved to ~80% of that of NMOS transistors [18]. Uniaxial compressive stress is usually obtained by creating stressed regions in the source and drain sides using SiGe, which has a larger lattice constant than Si. The compressive strain to the Si channel changes its lattice constant and therefore the band structure, and it can result in a significant enhancement in the hole mobility. In fact, over several generations of strained Si CMOS technologies, the hole concentration of Ge in the source and drain has kept increasing, and the source/drain/gate spacing has kept decreasing, resulting in increased stress [19]–[22].

For alternative p-type MOSFETs, Ge has the highest bulk hole mobility among conventional semiconductors and it looks very promising for its maturity in Si CMOS technology. SiGe regrowth and strain engineering has already been applied in the standard CMOS fabrication process. Current state-of-the-art high-Ge-content strained SiGe p-channel FinFETs have shown record g_m as high as 2.7 mS/μm, by aggressive L_g and EOT scaling (L_g ~ 25 nm, EOT ~ 0.7 nm) [23]. On the other hand, Ge NMOS transistors have not been able to demonstrate performance as

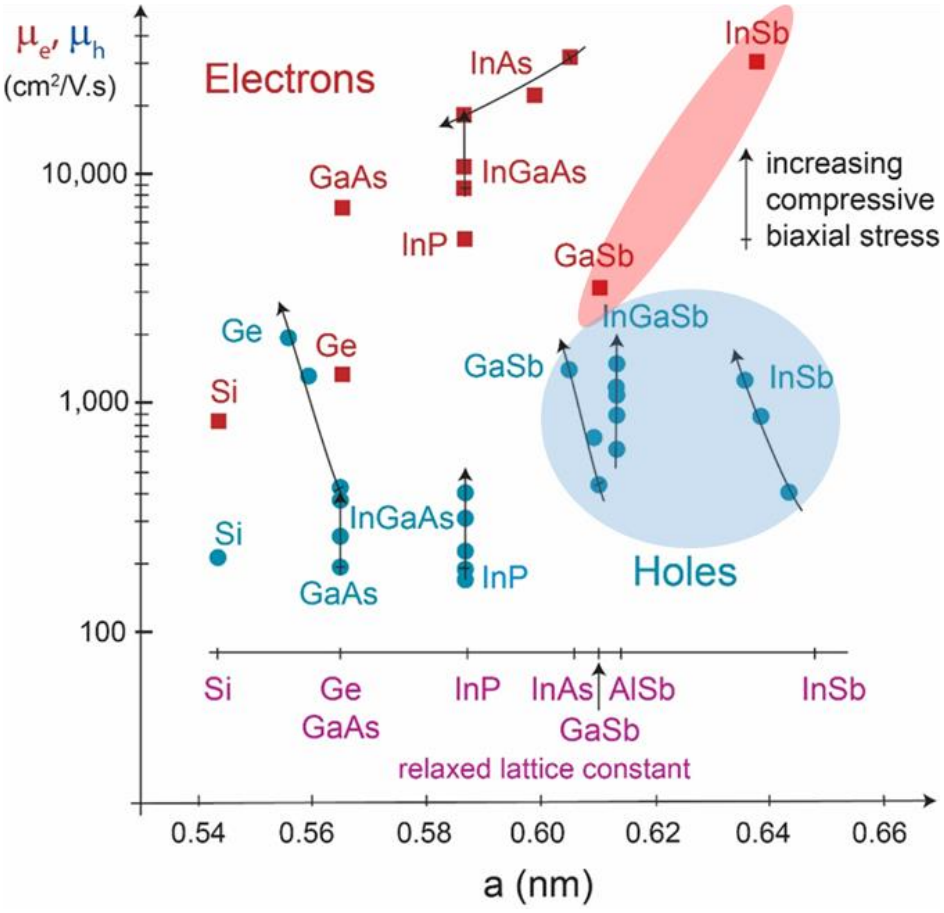


Fig. 1-6: Electron (red) and hole (blue) mobility of III-V compound semiconductors in reported devices, plotted against the lattice constant. The arrows represent the impact of biaxial strain on mobility. The hole mobility of InGaSb can be almost doubled under application of compressive strain.

good as InGaAs/InAs NMOS transistors. Therefore, future technology integrating InGaAs as NMOS channel and Ge as PMOS will pose challenges to the manufacture process [24].

Another option for alternative-channel CMOS is the III-V CMOS technology. Fig. 1-6 plots the highest room-temperature electron and hole mobilities of the III-V family, together with Si and Ge, reported in inversion layers and quantum wells, as a function of the corresponding lattice constant [1]. The impact of increasing compressive biaxial stress is indicated by the arrows. For electron transport, the $In_xGa_{1-x}As$ system and InSb shows the highest electron mobility. Unfortunately, $In_xGa_{1-x}Sb$ n-channel MOSFETs have never been demonstrated, though the electron mobility is predicted to be quite high. For hole transport, the $In_xGa_{1-x}Sb$ family shows the highest hole mobility, and it is the only III-V compound which can match the hole mobility of pure Ge. In addition, the hole mobility of $In_xGa_{1-x}Sb$ can be increased effectively by strain [25], bridging the gap between the electron mobility of other III-Vs. A hole mobility as high as 1,500 $cm^2/V{\cdot}s$ has been reported in strained $In_{0.4}Ga_{0.6}Sb$ [26], 1,300 $cm^2/V{\cdot}s$ in GaSb [27], and 1,200 $cm^2/V{\cdot}s$ in InSb [28].

While the carrier transport properties favor the III-V compound semiconductors, transistor technology must be developed and high-performance devices need to be delivered to advance the III-V technology in real-world applications. Recently, rapid progress has taken place in the development of InGaAs-based n-channel MOSFETs. From HEMTs [29], [30] to planar MOSFETs [31] to FinFETs [32] and to nanowire gate-all-around MOSFETs [33], [34], we have witnessed a surge of device innovations and demonstrations in the InGaAs system.

Besides the InGaAs system, antimonides, particularly $In_xGa_{1-x}Sb$, are attractive materials due to their extraordinary hole mobility, as well as their decently high electron mobility, as highlighted in Fig. 1-6. In addition, the antimonide family has lattice constants around 6.1 Å. Therefore, a fully III-V-on-Si CMOS integration scheme is possible by integrating InGaAs NMOS and InGaSb PMOS, or InGaSb for both NMOS and PMOS, as shown in Fig. 1-7. Several preliminary work has demonstrated such III-V CMOS integration [35]–[37]. Moreover, the antimonide family owns a wide range of band gap energies. Therefore, band gap engineering can be exploited using the antimonides without excessive lattice mismatch. For example, antimonides have been found essential in building InAs/GaSb tunnel FETs (TFETs) [38]–[40], which is a very promising class of novel devices for ultra-low power applications.

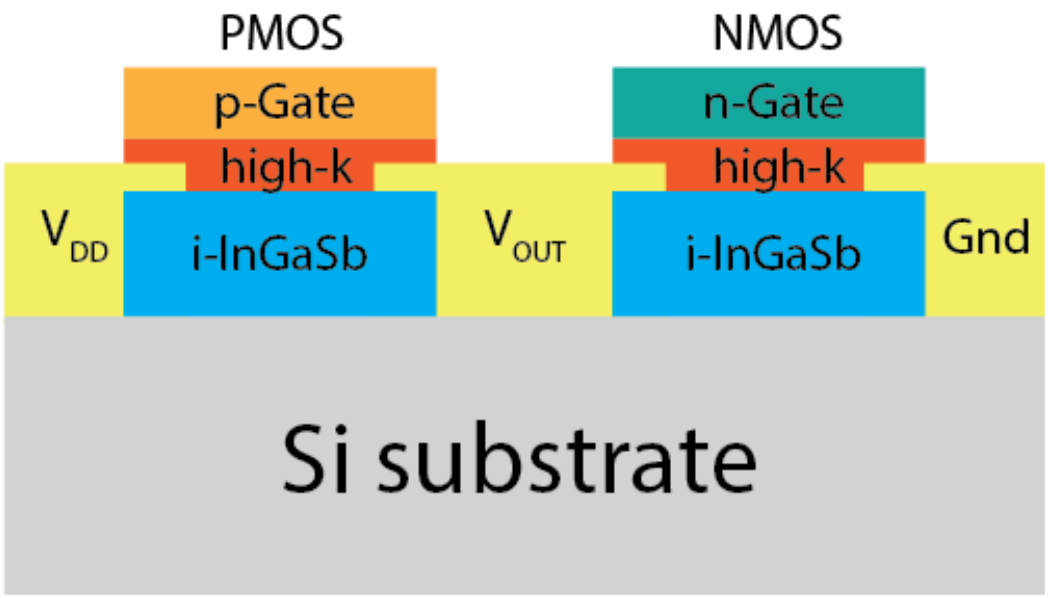


Fig. 1-7: Possible III-V-on-Si integration scheme using intrinsic InGaSb single channel for future CMOS technology with alternative channel materials.

However, the progress in the antimonide-based transistors, until recently, became stagnant. Although a few demonstrations of InGaSb p-channel transistors have been demonstrated [35], [37], [41]–[48], the device performance has been unsatisfactory. Fig. 1-8 shows two works that have had great influence on antimonide device research. Since around 2000, B. R. Bennett et al. have demonstrated InAs/AlSb p-channel HEMTs and have developed the key technology of molecular beam epitaxy growth of antimonides [45], [49]. In 2010, A. Nainani et al. have demonstrated the first InGaSb buried-channel pMOSFET that outperforms Ge with two times higher hole mobility [44], [50]. Nevertheless, the gap between antimonide device research with Ge or InGaAs devices remains large. Fig. 1-9 shows the number of publications on InGaAs transistors in IEEE journals alone vs. the number of Sb-based devices in all literature to date. This comparison not only shows the large discrepancy in the research effort, but also implies the tremendous challenges in antimonide device research. In fact, antimonide-based multi-gate MOSFETs have never been demonstrated before this thesis work.

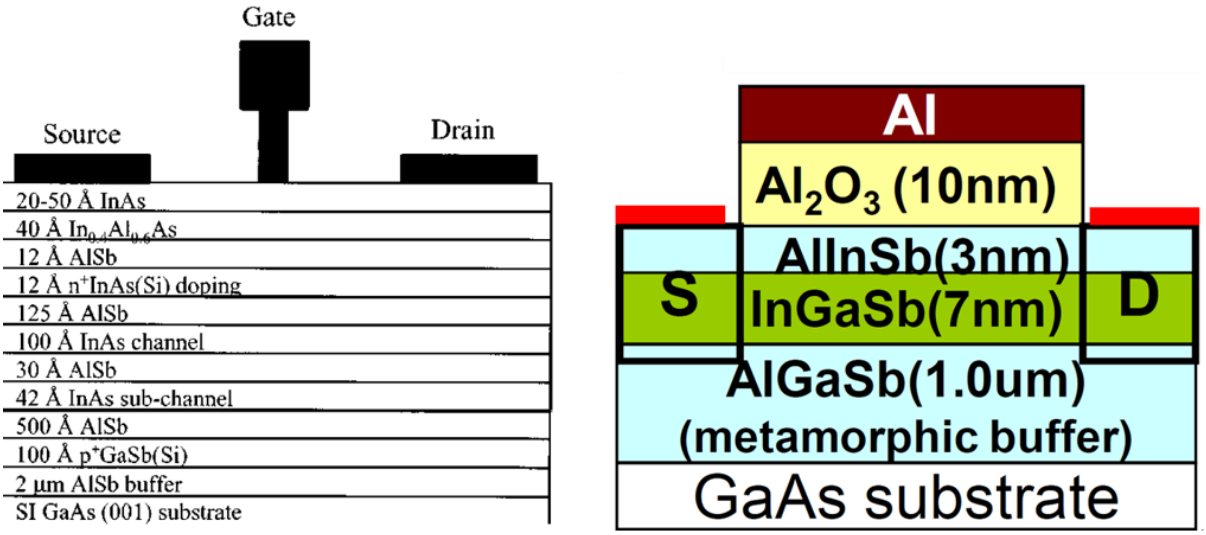


Fig. 1-8: Cross-section of (left) InAs/AlSb p-channel HEMT [45] and (right) InGaSb buried-channel pMOSFET [44].

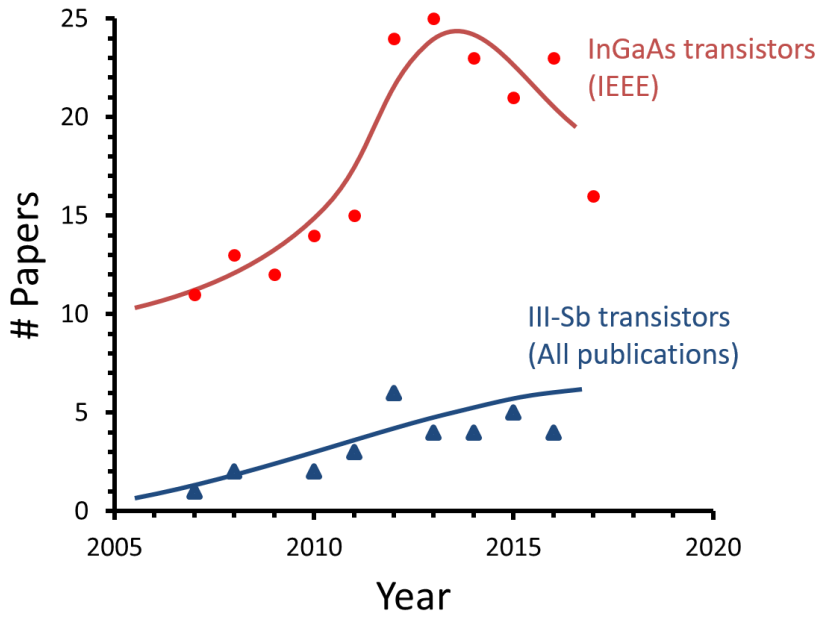


Fig. 1-9: Count of publications of InGaAs transistors in IEEE journals alone vs. the count of Sb-based devices in all literature to date.

The lack of InGaSb multi-gate MOSFETs can be attributed to several challenges. First, antimonide-based fin or nanowire etching has never been successfully demonstrated. All existing approaches to form 3-D antimonide structures rely on epitaxial growth, which is not CMOS-compatible. Second, antimonide compounds are highly reactive and unstable. When exposed to air, a layer of native oxide is instantly formed. What is even worse, the oxidation of InGaSb is not self-limiting, and the formed oxide layer severely degrades the device performance because of Fermi-level pinning. In addition, antimonides are easily attacked by most acids and bases used commonly in CMOS fabrication, such as hydrofluoric acid and sulfuric acid, which are common etchants, and tetramethylammonium hydroxide (TMAH), which is widely present in photolithography developer. Therefore, cleaning and passivation of the Sb surface are difficult, and formation of a high-quality interface with the high-k gate dielectric is challenging. Third, high-quality nano-scale ohmic contacts for antimonide-based devices have not been thoroughly

investigated. As device footprints scales, source-drain contact resistance plays an important role by limiting the drive current. Eventually, and most critically, a 3-D process integration of a FinFET is a significantly more complex fabrication task than that of a planar transistor, even after resolving the challenges enunciated above. Therefore, it is the goal of this thesis work to overcome the above challenges and demonstrate the first InGaSb p-channel FinFET.

1.5 Thesis Overview

This thesis studies some novel solutions to the formerly described challenges with InGaSb p-channel multi-gate transistor technology. Since no such device has ever been demonstrated, this thesis aims to start from the fundamental technological challenges of antimonide-based process, then move on to the demonstration of full transistor fabrication, and eventually investigate the electrical performance of the InGaSb multi-gate transistors.

In the first part of the thesis, the goal of demonstrating the first InGaSb multi-gate transistor is tackled in a divide-and-conquer approach. Prior to a complete device design and fabrication, solutions to the challenges mentioned above need to be delivered. Three main technologies are developed. First, an antimonide reactive ion etching (RIE) technology is developed. The RIE technology can etch antimonide fins and vertical nanowires with high aspect ratio, anisotropy, and sidewall smoothness. The new technology is also compatible with other III-V compounds, such as arsenides. Second, a novel alcohol-based digital etch technology is invented. It can be used to perform fin and nanowire sidewall etching with nanometer precision, at a rate of 2.0 nm/cycle. It also serves as a surface passivation method for the high-k dielectric/InGaSb interface. It has been

found to improve interfacial quality and reduce leakage current. Last, a nanoscale ohmic contact scheme using nickelide is developed and analyzed. It achieves record ultra-low contact resistance. A composite p^+ capping layer is designed to reduce the contact resistance.

In the second part of the thesis, the above technologies are applied and integrated into the design of an InGaSb p-channel FinFET. The design principles of the heterostructure and process of the proposed device are explained. Then, a fabrication process is developed for the actual transistors. The smallest devices achieved feature a minimum fin width of 10 nm and a gate length of 20 nm. It is the not only the first demonstration of InGaSb FinFET, but also one of the most aggressively scaled devices based on III-V compound semiconductors. Lastly, the electrical characteristics of three generation of fabricated InGaSb FinFETs are reported and analyzed.

CHAPTER 2. PROCESS DEVELOPMENT

2.1 Process Overview

Following the introduction in Ch. 1, this chapter describes the key processes required in a self-aligned InGaSb p-channel FinFETs. As described in the previous chapter, various critical fabrication technologies for antimonide-based multi-gate transistors are not yet available. To achieve the goal of demonstrating the first InGaSb FinFET, this part of the thesis focuses on an effort to develop three major components of an antimonide-based FinFET. Those are (1) fin reactive ion etching, (2) ohmic contact, and (3) digital etch. All the three components are supposed to be optimized for the antimonides, and even better, compatible to the arsenides.

2.2 Fin Reactive Ion Etching

2.2.1 Antimonide RIE – First Generation

As its name suggests, fin etching is the essential process module for the FinFET fabrication. In our top-down approach, the reactive ion etching (RIE) is required to deliver high-aspect ratio fins with nanometer-scale fin width. In addition, the RIE process needs to be both anisotropic and gentle so that the profile of the fin is vertical and uniform while the sidewall of the fin remains smooth.

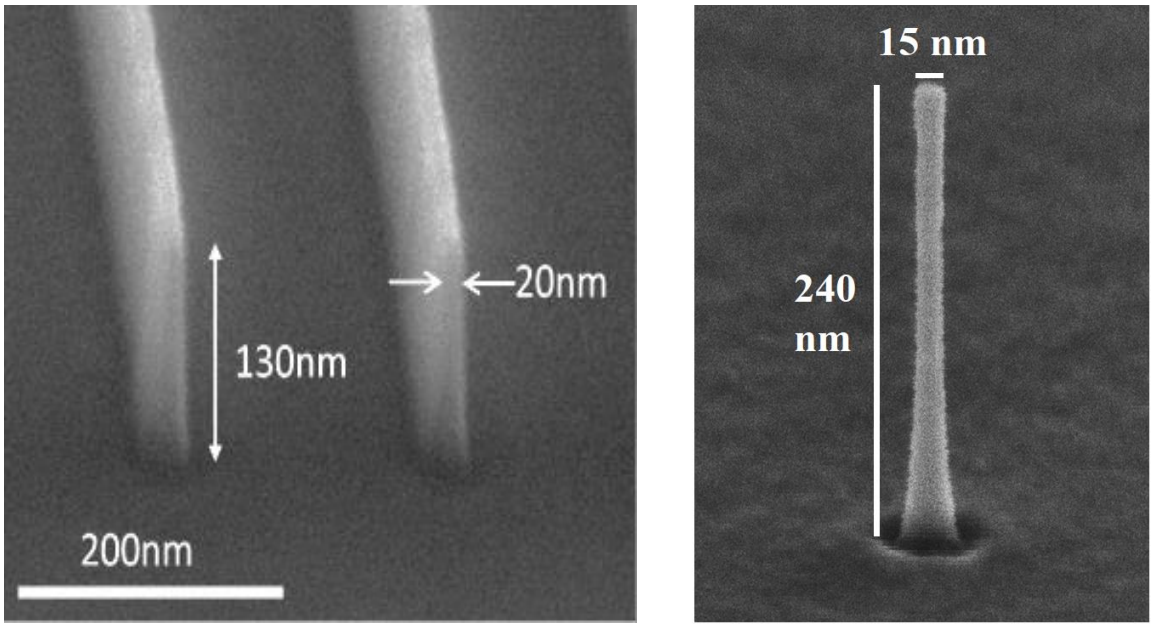


Fig. 2-1: (left) InGaAs fins with 20 nm fin width and (right) InGaAs VNW with 15 nm diameter, patterned by optimized RIE process.

In recent years, there have been tremendous breakthroughs in the InGaAs FinFET and vertical nanowire (VNW) MOSFET technologies [32], [34], [51]–[54]. The InGaAs fin RIE process is the backbone of these high-performance transistors. Fig. 2-1 shows some examples of the state-of-the-art InGaAs fin and VNW etching developed in our group [34], [55].

An example of the starting antimonide-based heterostructure used in this thesis is shown in Fig. 2-2. It is grown by molecular-beam epitaxy. The details of the design and properties of the heterostructure will be described in Ch.3 when the FinFET process integration is discussed. Note that a typical etched fin structure contains multiple compounds, such as InGaSb, AlGaSb, and

perhaps InAs, with different compound compositions. Therefore, the RIE process needs to show ideally negligible material selectivity for various III-V compounds semiconductors.

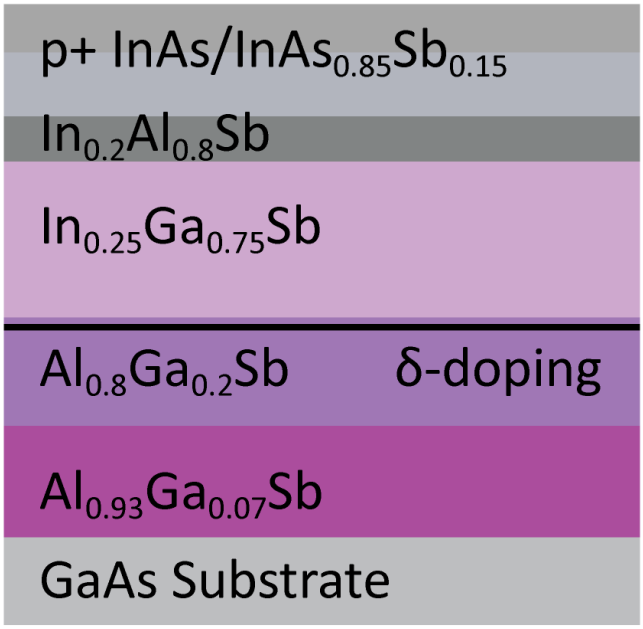


Fig. 2-2: An example of the heterostructure contained in an etched fin or VNW.

The first step to form the fins is the patterning of etching hardmask. In this thesis, HSQ[1] (hydrogen silsesquioxane), a flowable oxide diluted with methyl-isobutyl-ketone (MIBK), is used as the hardmask and patterned by electron beam lithography (EBL). Before HSQ spin coating, the wafer surface is coated with a thin Si_3N_4 film deposited by plasma-enhanced chemical vapor deposition (PECVD) at 300 °C to promote adhesion of HSQ on the semiconductor surface. The thickness of the PECVD Si_3N_4 film is only 2-3 nm to avoid complication of the fin etching process. Then, the

[1] HSQ is distributed by Dow Corning as XR-1541™.

HSQ is spin coated onto the sample and exposed in the Elionix ELS-F125 EBL system, with a resolution of 5 nm at a beam current of 1 nA. After the e-beam exposure, the HSQ is immediately developed in 25% tetramethylammonium hydroxide (TMAH) for 60s, followed by DI water and isopropanol rinsing, and blown dry gently by N_2 gun. It has been suggested that development of HSQ in low temperature in cold-water bath improve the resist contrast for high-resolution patterns [ref]. From our experience, room temperature development suffices for patterns as small as 15 nm. On the other hand, it is worth noting that both the exposure and development of HSQ are done in the same day to avoid HSQ crosslinking in air, in order to ensure best optimal resolution and reproducibility. Fig. 2-3 shows scanning electron microscopy (SEM) images of developed HSQ nanowire pattern with 26 nm diameter and fin patterns with various widths.

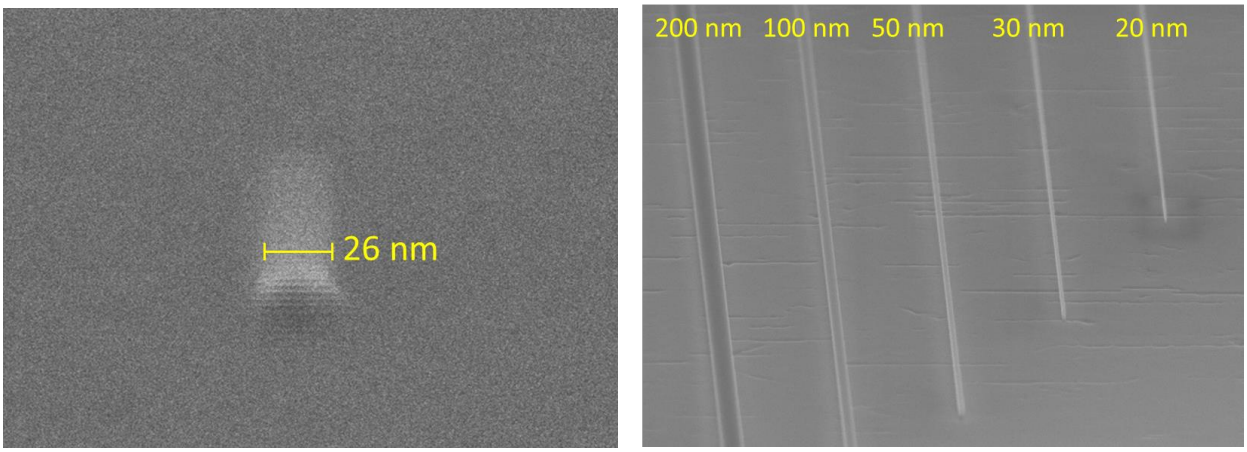


Fig. 2-3: Tilted SEM images of developed HSQ (left) nanowire pattern with dimeter of 26 nm and (right) fin patterns with different width.

Then, the RIE is carried out in an inductive coupled plasma (ICP) etcher SAMCO RIE-200iP, using chlorine-based etching. Samples were cleaved into approximately 1x1 cm^2 small rectangular pieces and loaded on a 6-inch ceramic wafer carrier, one sample at a time. In the case of very small

sample size (e.g. for testing or calibration purpose), a 1x1 cm^2 GaAs wafer is put next to the sample to avoid variability caused by loading effect.

After iterative tuning of the gas chemistry, gas flow, and etching power, the etching condition is optimized to the following condition: 13.5/5.5 sccm BCl_3/N_2, 20 W ICP power, 280 W RF platen power, 0.2 Pa chamber pressure, and substrate temperature of 250°C. The etched fins are shown in Fig. 2-4. This is the first demonstration of RIE of antimonide-based fins. Both single fins and arrays of fins with vertical profile and smooth surface can be obtained. The resulted fins have high aspect ratio > 10 and no noticeable undercutting or trenching is observed.

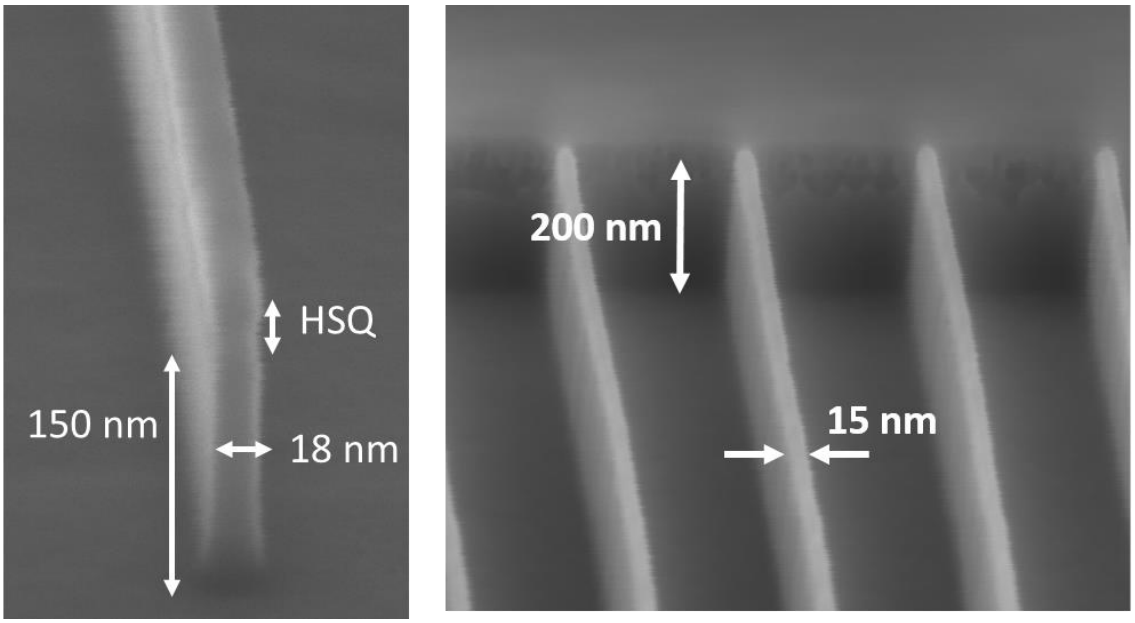


Fig. 2-4: SEM images of (left) single InGaSb fin with 18 nm fin width and 150 nm fin height, and (right) fin array with 15 nm fin width and 200 nm fin height.

The developed RIE process is also able to etch InGaSb vertical nanowires with high aspect ratio and vertical profile (> 85°). Moreover, dense fin arrays can be achieved, with 20 nm fin width and 20 nm fin spacing, which is equivalent to a full pitch size of 40 nm. Examples of an InGaSb VNW and tight-pitch fin array is demonstrated in Fig. 2-5.

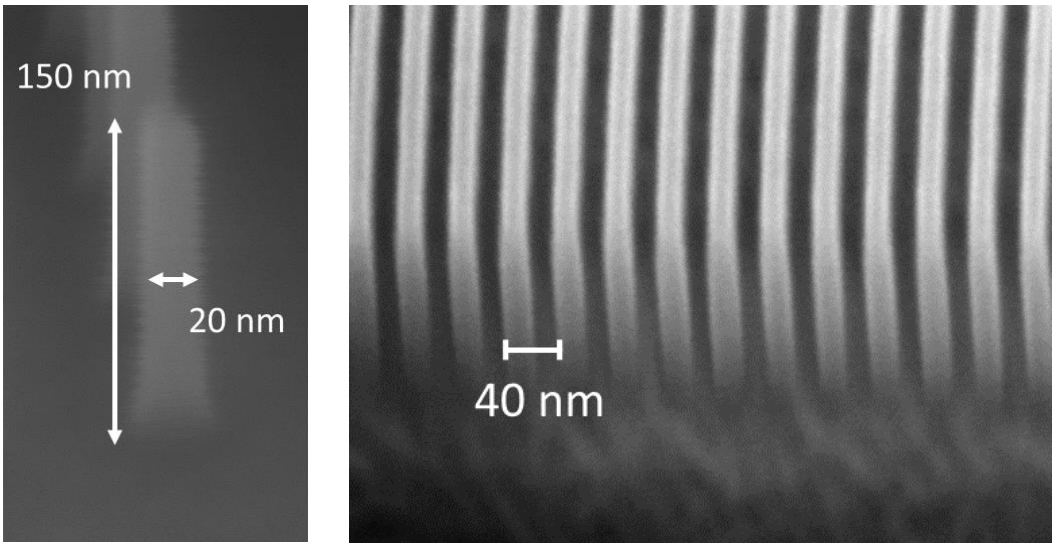


Fig. 2-5: SEM images of 20 nm InGaSb vertical nanowire and dense fin array with 20 nm fin spacing.

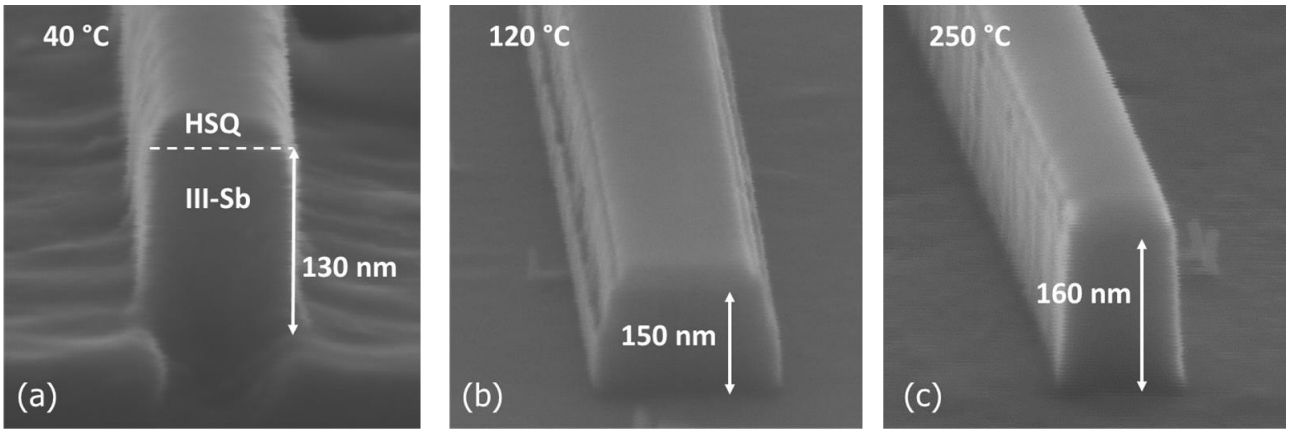


Fig. 2-6: RIE of antimonide-based heterostructures using HSQ hardmask at: (a) 40°C, (b) 120°C, and (c) 250°C substrate temperature during etching.

The dry etching of antimonide fins shows strong dependence on the substrate temperature. Fig. 2-6 shows the same etched fin test structures at different substrate temperatures of 40°C, 120°C and 250°C, with other etching conditions unchanged. It illustrates the clear benefit of higher substrate temperature during the RIE. Raised substrate temperature increases the etch rate from 65 nm/min at 40°C to 80 nm/min at 250°C. At higher substrate temperature, the resulted fin sidewall and surface are smoother and profile is more vertical. Raised temperature etching is necessary because of the lower volatility of $InCl_x$. In this thesis, the etching temperature is kept at 250°C because it is the highest temperature the heater allowed in our etcher system. Nevertheless, the results are already indeed very promising. Fig. 2-7 shows that the etched fin sidewall appears to be very smooth with little material selectivity.

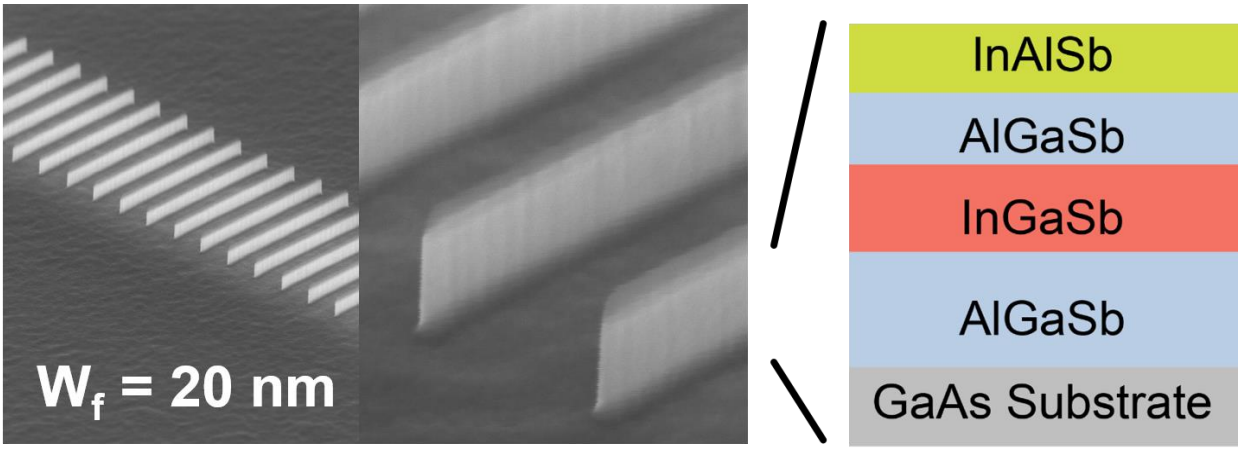


Fig. 2-7: Tilted SEM images of the side-view of the antimonide-based fin sidewall.

2.2.2 Antimonide RIE – Second Generation

The RIE process described above has been integrated into the first and second generation of InGaSb FinFET fabrication. In the third generation of FinFETs, an improvement has been made to the RIE technology so that it is compatible to arsenide-based compound semiconductors. The RIE is performed by ICP plasma with the following condition: 3:11:0.4 sccm BCl_3/Ar/$SiCl_4$, 280 W RF platen power and 20 W ICP power, at 250 °C. Fig. 2-8 shows fins comprising of 50 nm InAs on GaSb with the former mentioned RIE process and the improved one. With the former RIE condition, the InAs layer is roughened during dry etch, while a smooth sidewall is obtained with the improved RIE condition. Therefore, simultaneous etching of antimonides and arsenides is achieved without roughening any of the surface. This is significant when dry etching of both types of III-V compounds is required, for example, for InAs/GaSb tunnel FETs (TFETs). This etching condition is deployed in the third generation of the InGaSb FinFETs.

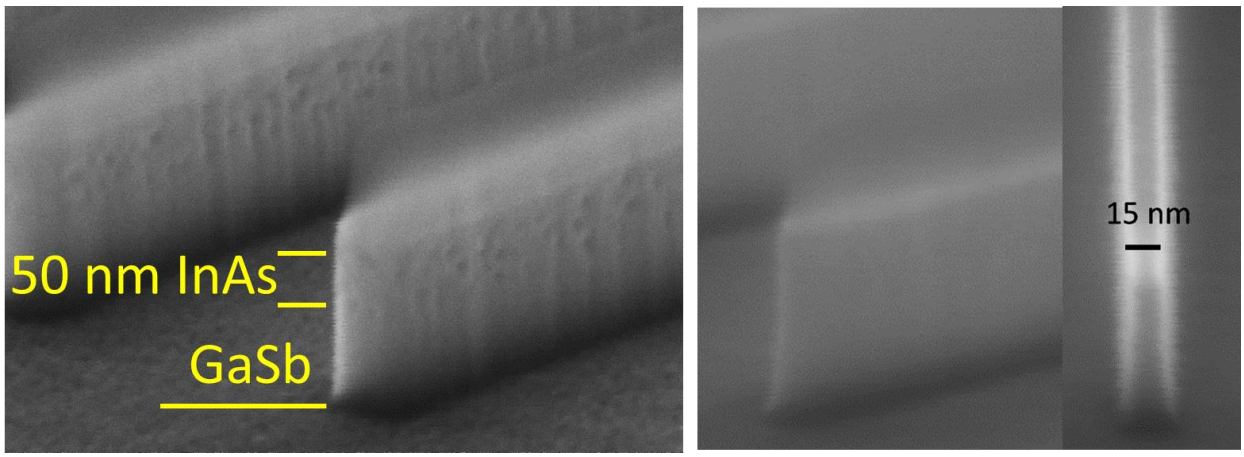


Fig. 2-8: SEM images of RIE-etched InAs/GaSb fins using the RIE condition of (left) 13.5/5.5 sccm BCl_3/N_2, and (right) 3:11:0.4 sccm BCl_3/Ar/$SiCl_4$. In both cases, RF platen power is 280 W, ICP power is 20 W, chamber pressure is 0.2 Pa, and substrate temperature is 250°C.

It is worthy to note here that high-quality fin RIE is necessary but not sufficient for highly scaled high performance III-V FinFETs. First, the fin width is limited by the capability and resolution of the electron-beam lithography, which is limited to around 15 nm and beyond. Second, considerable plasma damage might be taking place, especially when anisotropic etching is desired and therefore the plasma power is large. Surface cleaning technique for etched fin sidewalls is particularly important for antimonides because of their highly reactive and unstable surface. Therefore, a digital etch technology is in dire need to scale the device further down and improve the surface quality. In the following chapter, I will describe the development of the digital etch technology for antimonides.

2.3 Digital Etch

One issue of top-down approach of FinFET fabrication via dry etching is the RIE damage of the fin sidewall, due to the high energy ion bombardment during the etch [56]–[59]. The quality of the fin sidewall is one of the most critical factors for FinFETs operation, because the current flows in close proximity to the surface. In Si FinFETs fabrication, high temperature forming gas annealing at above 850°C has been shown to be an effective approach to improve the surface quality [60]. However, processing at such high temperature is unfeasible for III-V transistors. Although direct wet etching can be used to remove RIE-damaged semiconductor, it is not applicable to deeply scaled devices, which require nanometer-scale etching precision.

In silicon CMOS technology, the concept of digital etch (DE) or atomic layer etching (ALE) has been proposed and developed for more than twenty years [61], [62]. In such process, layer-by-layer etching of Si can be achieved at atomic scale, by a repeated process of surface modification and surface removal reactions that are self-limiting. For example, the surface modification step can be fluorination of the Si surface by CF/O_2 plasma to form a thin layer of volatile silicon fluoride. Then, the surface removal step can be achieved by Ar^+ ion irradiation. As the semiconductor industry has entered the era of atomic-scale dimensions, the digital etch technology is turning to be increasingly critical and is under active research. Inspired by the Si digital etch and atomic layer etching technology, the similar concept has been introduced to III-V compound semiconductor as well, as an effective to modify the surface property at atomic-layer scale.

2.3.1 III-V Digital Etch

In the quest for enhanced performance over silicon devices at advanced nodes, III-V multi-gate transistors such as FinFETs or nanowire gate-all-around MOSFETs are being pursued with fin width or nanowire diameter in the sub-10 nm range. In this dimensional regime, precise etching control is paramount. Inspired by the atomic layer etching technology developed in silicon [61], [62], the concept of digital etch was introduced to III-V compound semiconductors [63]–[68]. In digital etch, the oxidation and oxide removal steps characteristic of chemical etching are performed separately. This makes them both self-limiting affording nanometer-scale control of the etching process. Digital etch has been shown to preserve high aspect ratio features, enabling precise device dimension engineering. It has also been shown to be effective in mitigating surface damage induced by RIE [69]. In the last few years, the use of digital etch has enabled demonstrations of

aggressively scaled III-V FinFETs and NW-MOSFETs with exciting electrical characteristics [38], [39], [70]–[72].

Prior to this thesis work, digital etch has been solely applied to arsenide-based III-Vs, achieving a smallest feature size of 5 nm in fins [55] and 7 nm in lateral nanowires [38]. On the other hand, demonstrations of high aspect ratio vertical nanowires (VNW) by digital etch have been limited to 11-15 nm in diameter [39]. Firstly, conventional digital does not work on antimonide-based III-Vs, because antimonides are typically highly reactive, and therefore a benign and stable self-limiting digital etch process is challenging. Secondly, a digital etch process for fins and VNWs with critical dimensions approaching sub-10 nm regime is desired to push the state-of-the-art of III-V transistor technology, regardless in InGaAs or InGaSb.

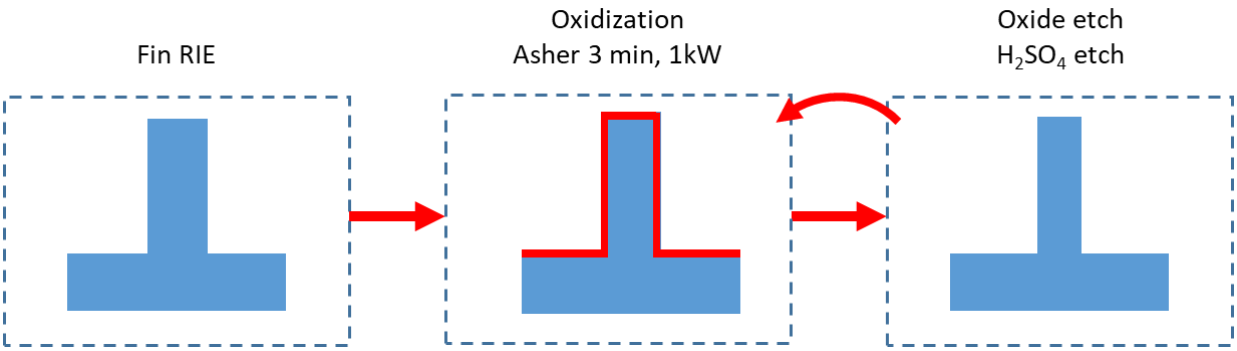


Fig. 2-9: Process flow of a typical digital etch process.

Fig. 2-9 illustrates the typical process flow of a digital etch process. After the fin RIE, the surface is oxidized using oxygen plasma (in an asher), at 1 kW for 3 minutes. Other oxidation agents include hydrogen peroxide, ozone, or ozonated water. Then, the oxidized surface is etched away by some acid solutions, such as hydrochloric acid or sulfuric acid, which are supposed to have

negligible etching rate on the III-V. Afterwards, the same procedure of oxidation and oxide removal is repeated in cycles, removing one layer at a time.

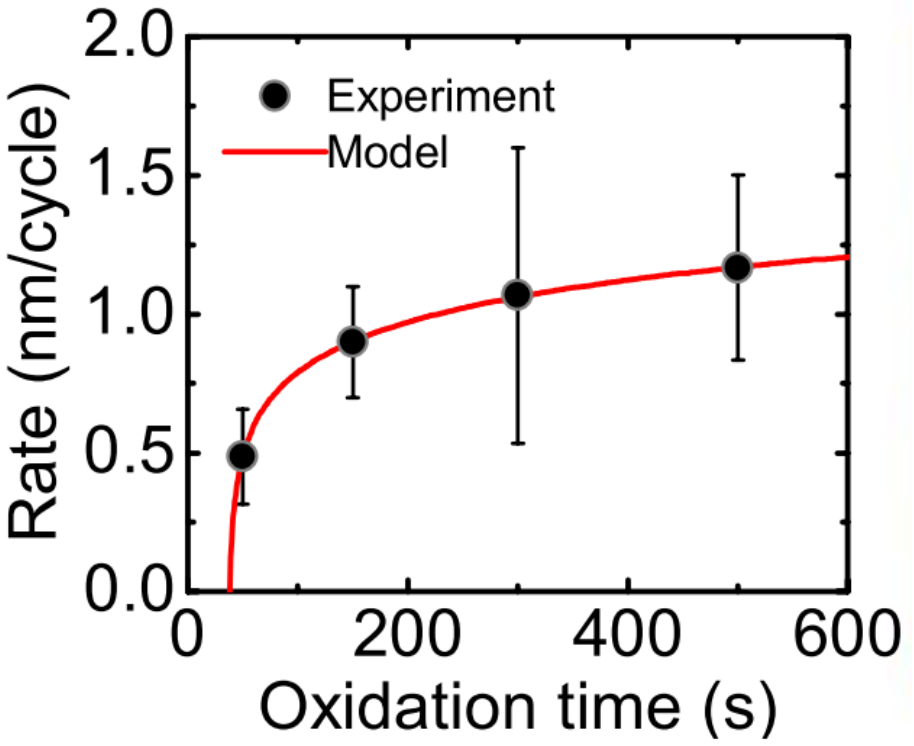


Fig. 2-10: Etch rate per cycle of digital etch on InP planar wafer, as a function of oxidation time under O_2 plasma, adapted from [66], fitted with the Lukeš model [73].

Fig. 2-10, adopted from [67] shows the digital etch rate per cycle on planar InP wafer, as a function of oxygen exposure time, fitted by the Lukeš' rate law [73]. It shows saturation behavior beyond 180 seconds of oxidation time, illustrating the self-limiting property of the oxidation process. Fig. 2-11 illustrates the progress of digital etch process on an InGaAs vertical nanowire in a sequential etch experiment after 2, 5 and 10 cycles of digital etch using oxygen plasma and 10% HCl. The initial diameter of the VNW after RIE is 34 nm, and after 10 cycles, the diameter is reduced to 14 nm. The average etch rate is 1 nm/cycle, on each side of the VNW. It can be observed that the

digital etch preserves well the profile of the VNW. It proves the unpreceded benefit of using digital etch to achieve nanometer-scale scaling of device dimensions.

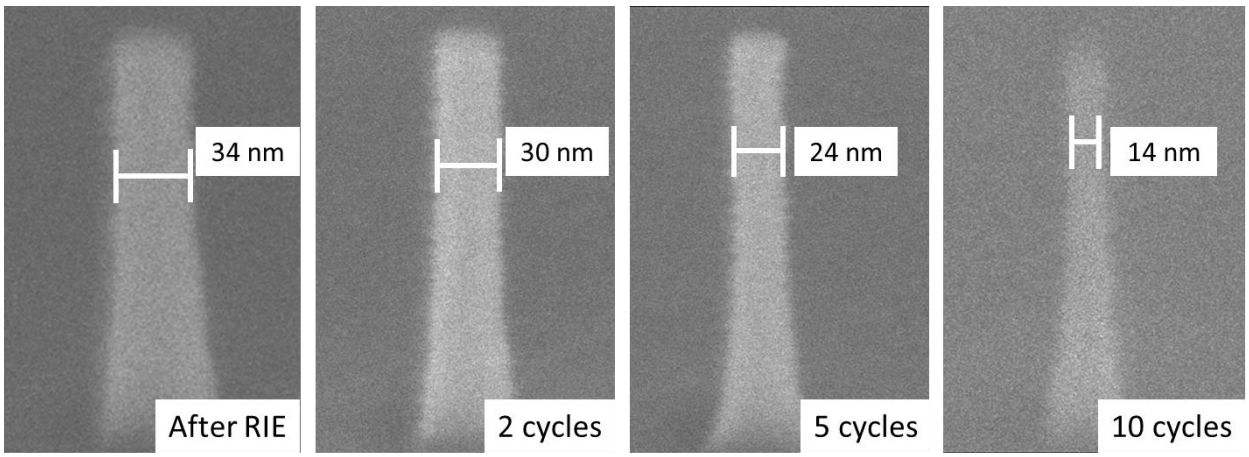


Fig. 2-11: Evolution of InGaAs VNWs in a sequential etch experiment involving different numbers of digital etch cycles in 10% HCl:IPA.

Another important benefit of digital etch is its ability to mitigate plasma damage and improve surface quality. By removing the semiconductor surface layer by layer, the surface roughness is improved and better device performance can be obtained. Fig. 2-12 shows the subthreshold characteristics of InGaAs VNW MOSFETs with and without digital etch (10 cycles), with the transconductance (g_m) characteristics in the inset [69]. While the two transistors are almost identical (D_{final} = 30 nm), the transistor with digital etch shows lower subthreshold swing of 150 mV/dec (at V_{ds} = 50 mV), and higher peak g_m of 280 μS/μm (at V_{ds} = 0.5 V). These results show the significant reduction of sidewall damage by the digital etch process.

Because of the above benefits, digital etch has been widely applied in recent InGaAs MOSFET fabrication. However, there is yet no digital etch reported in the antimonide-based material system. In fact, there is no prior study on surface treatment of antimonide fin/VNW sidewall. Because of the development of our InGaSb fin RIE technology, we are able to study this topic for the first time.

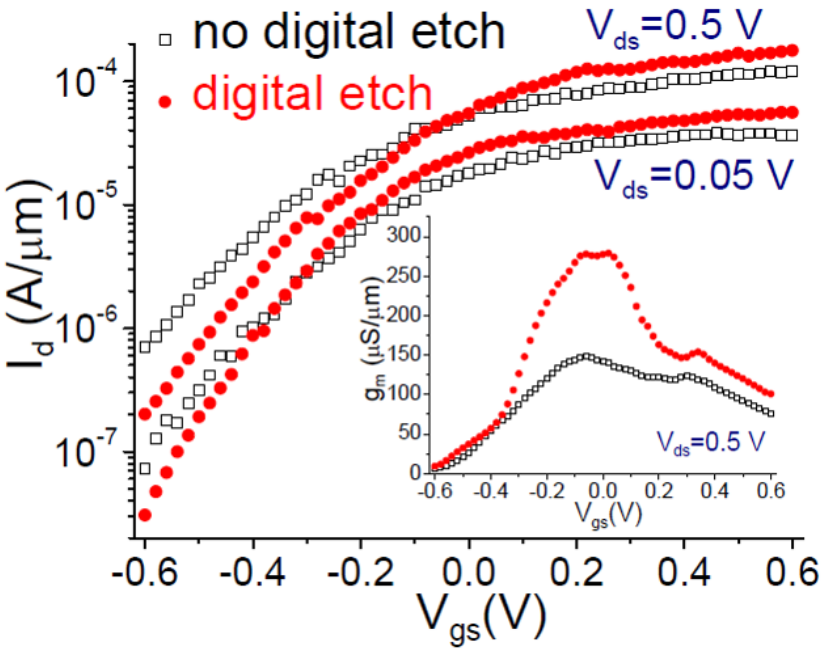


Fig. 2-12: Subthreshold and transconductance (inset) characteristics of InGaAs VNW MOSFETs, with final diameter of 30 nm, with and without digital etch [69].

2.3.2 Antimonide-compatible Digital Etch

In principal, both the oxidation and oxide removal steps of digital etch need to be self-limiting. Unfortunately, neither of the two self-limiting steps can be easily obtained in the antimonide-based devices. First, antimonide-based compounds are notoriously known for its highly reactive surface.

The oxidation of GaSb is very fast even in air, and in theory non-limiting [74]. Second, the etching or removal of native oxide of antimonides is difficult. Chemical cleaning of GaSb surface has been studied for many years [74]–[80], and there is no demonstration of effective approaches. Most popular methods is using diluted HCl:H_2O solution, which has been shown to remove GaO_x and SbO_x from the surface [75]. In addition, all existing studies on antimonide surface cleaning or passivation were carried out on planar wafers, and the impact is characterized by surface studies such as x-ray photoelectron spectroscopy (XPS), or electrically by capacitance-voltage (C-V) measurement. There is little information about the etch rate and selectivity, particularly in the case of vertical structures or sidewalls.

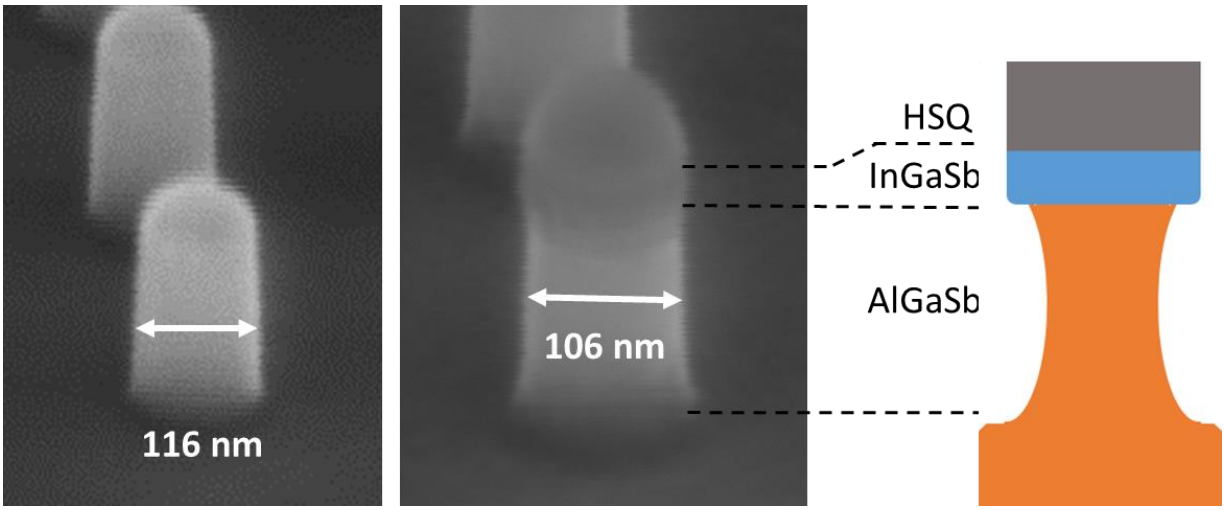


Fig. 2.-13: Etched InGaSb/AlGaSb VNWs with initial diameter of 116 nm (left), after 30s dipping in 1% HCl:H_2O (middle). (Right) schematic showing the effect of HCl cleaning on antimonide-based vertical sidewalls.

Fig. 2-13 shows tilted SEM images demonstrating the effect of a 30 second dipping of InGaSb VNW in 1% diluted HCl:H_2O solution. It is observed that the diameter of the VNW decreased from 116 nm to 106 nm. Also, the shape of the VNW is changed, and the surface of the sidewall

becomes rougher. This shows that HCl:H_2O treatment is not self-limiting and is etching the InGaSb and AlGaSb at different rates. This poses the first major challenge of developing digital etch for the antimonides.

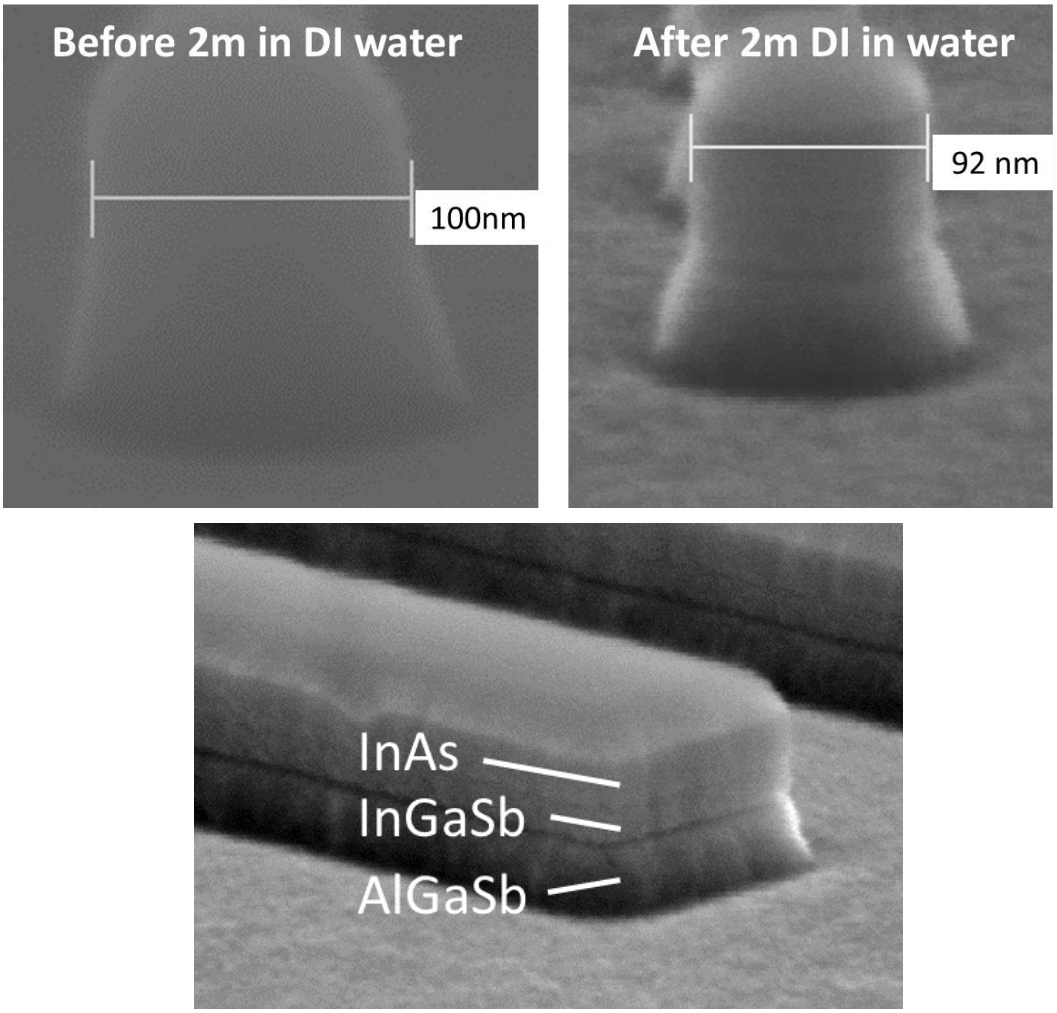


Fig. 2-14: Antimonide vertical nanowire (a) before and (b) after dipping in DI water for 2 minutes. (c) Antimonide fin test structure after dipping in DI water for 2 minutes.

After trying several other etchants, it seems that simply replacing the HCl does not solve the problem. Another aspect of the etching process is the solvent of the etchant, which is usually neglected. Fig. 2-14 shows InGaSb nanowire and fin test structures before and after a two-minute dipping in DI water. The diameter of the antimonide nanowire decreases from 100 nm to 92 nm.

Also, the nanowire sidewall and AlGaSb field surface become rough and notches appear at the heterojunction interfaces. The antimonide fin sidewall in Fig. 2-14(c) also reveals such etching effect. This etching of antimonide-based heterostructures in water has not been reported prior to this work. It implies that the surface of the antimonides is chemically highly reactive and unstable, so that it can react even in water. Also, the notches formed near the interfaces of heterostructure layers imply preferential etching, possibly due to defects from the epitaxial growth and strain, as antimonide-based heterostructures are generally not lattice matched.

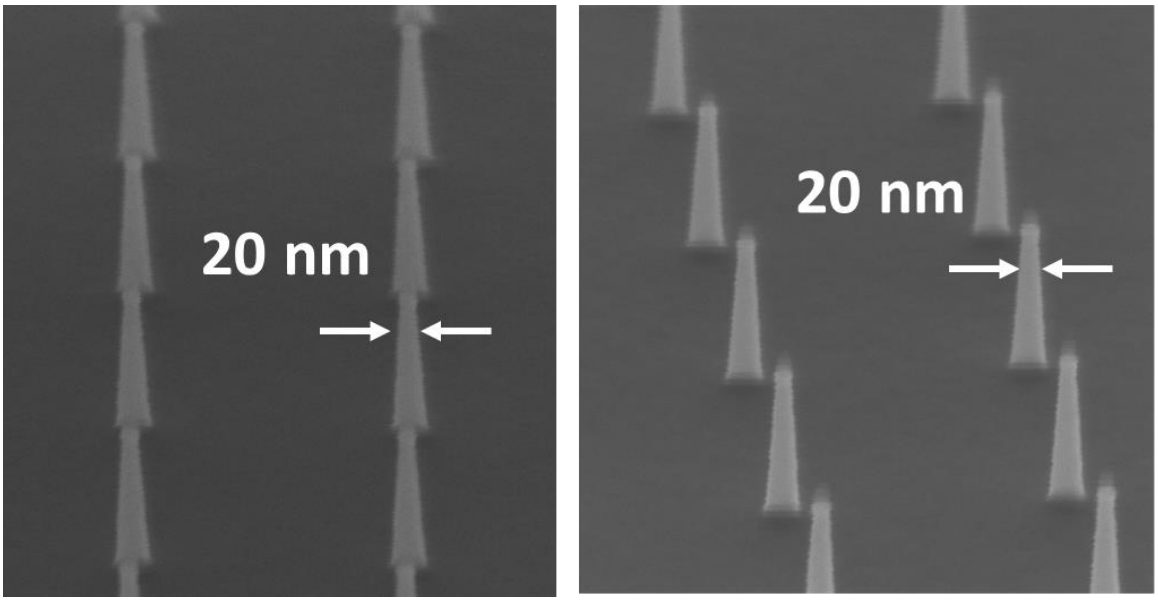


Fig. 2.15: Antimonide vertical nanowires with 20 nm diameter (left) before and (right) after dipping in 10% HCl:IPA for 2 minutes.

Therefore, to develop the digital etch, water needs to be removed first from the process. Fig. 2-15 shows VNWs with the same antimonide heterostructure, before and after 2 minutes dipping in 10% HCl:IPA[2]. By using an alcohol-based solution, no noticeable sidewall etching or surface

[2] The solvent-based acids are commercially available at Sigma-Aldrich Co as HCl-2-proponol solution.

damage is observed, implying that the etching is self-limiting after removing the layer of oxide on the sidewall.

With the discovery of the alcohol-based treatment, digital etch on antimonide-based heterostructure is demonstrated for the first time. Oxidation is done by O_2 plasma as described above. Fig. 2-16(a) shows the radial etch rate of digital etch on InGaSb vertical nanowires as a function of number of digital etch cycles. Fig. 2-16(b) shows an InGaSb fin obtained after 5 cycles of digital etch in HCl:IPA, with a minimum fin width of 10 nm.

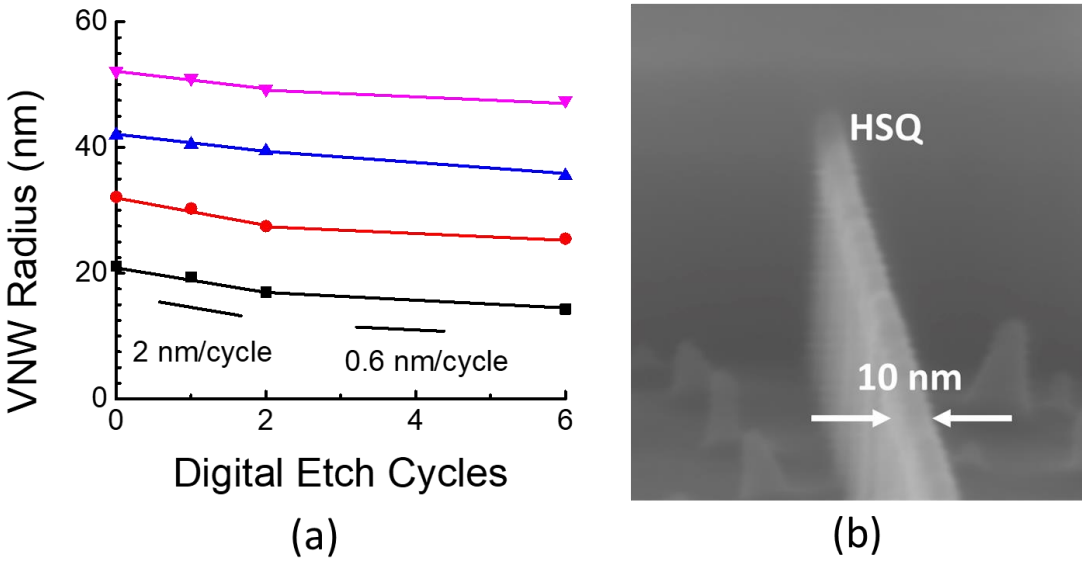


Fig. 2-16: (a) Evolution of antimonide VNW diameter with number of digital etch cycles in 10% HCl:IPA. (b) 10 nm wide InGaSb fin obtained after 5 cycles of digital etch in HCl:IPA.

Fig. 2-16(a) reveals an additional issue that remains with antimonide-based digital etch. For the first two cycles of digital etch, the radial etch rate is 2 nm/cycle, while the etch rate decreases to an average only 0.6 nm/cycle for the next 4 cycles. This is inconsistent with how digital etch works

with InGaAs, in which the etch rate is independent of the numbers of cycles. Therefore, it is possible that the surface property of the antimonide is changed after the treatment of digital etch.

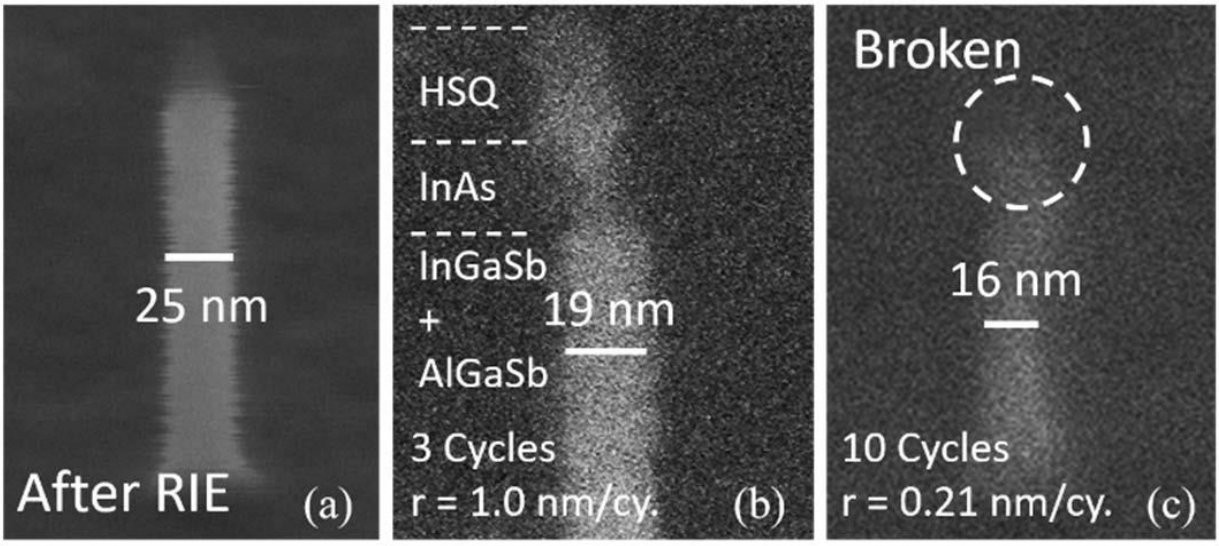


Fig. 2.17: (a) InAs/InGaSb/AlGaSb vertical nanowire after RIE with initial diameter of 25 nm, and after (b) 3 cycles and (c) 10 cycles of digital etch in O_2 plasma and HCl:IPA. Noted are the average radial etch rates of the InGaSb/AlGaSb portion of the nanowires.

Fig. 2-17 shows sequential digital etch of antimonide-based vertical nanowire in O_2 plasma and HCl:IPA. The nanowire contains InAs, as well, so that the digital etch on the two compounds can be monitored together. Alcohol-based digital etch produces an initial radial etch rate on the III-Sb of 1.0 nm/cycle, while the etch rate on the InAs portion of the nanowire is about 2.0 nm/cycle. It is an important deficiency that the digital etch has material selectivity. In addition, after 10 cycles, the average etch rate on the III-Sb portion of the structure is drastically reduced to about 0.21 nm/cycle, while all the InAs has already been etched away. Other researchers have reported a similar much reduced etch rate of GaSb in water-based digital etch [40].

Oxide etch \ Oxidation	UV ozone	H_2O_2	Organic peroxides	O_2 plasma	O_2
H_2SO_4:methanol	Damage	Damage	Damage	Damage	Damage
Citric acid:IPA	No etching	No etching	No etching	No etching	No etching
Acetic acid:IPA	No etching	No etching	No etching	No etching	No etching
HCl:IPA	No etching	Rate $\rightarrow$ 0	Rate $\rightarrow$ 0	Rate $\rightarrow$ 0	2 nm/cycle

Table 2.1: Summary of digital etch process using various methods for oxidation and oxide etching.

The issue lies in the oxidation properties of antimonides. For GaSb, strong oxidizing agents, like hydrogen peroxide, are known to produce Sb_2O_5 in addition to Ga_2O_3, Sb_2O_3 and elemental Sb on the surface [76]. However, Sb_2O_5 is insoluble in most acids or alkali. Therefore, when enough Sb_2O_5 has been formed on the surface during the oxidation process, digital etch effectively stops. To address this issue, we have investigated digital etch by multiple combinations of oxidation and oxide removal methods, as shown in table 2.1. In the end, our study reveals that 3 minutes exposure in pure O_2 chamber at room temperature combined with 10% HCl:IPA etch for 30 seconds gives the most satisfactory results. The O_2 atmosphere prevents strong oxidation of antimonide surface to avoid formation of Sb_2O_5. Fig. 2-18(d) shows the same vertical nanowire in Fig. 2-17, treated with 2 and 4 cycles of digital etch in O_2 atmosphere and HCl:IPA. A consistent radial etching rate of 2.0 nm/cycle for all of InAs, InGaSb, and AlGaSb is demonstrated. The smallest fabricated InGaSb vertical nanowire has a diameter of 9 nm. Fig. 2-18 also shows SEM images of examples

of some other oxidation and etching methods and their effects, including O_2 plasma and H_2SO_4 in methanol, hydrogen peroxide and citric acid, O_2 plasma and HCl in IPA, and O_2 atmosphere and HCl in IPA.

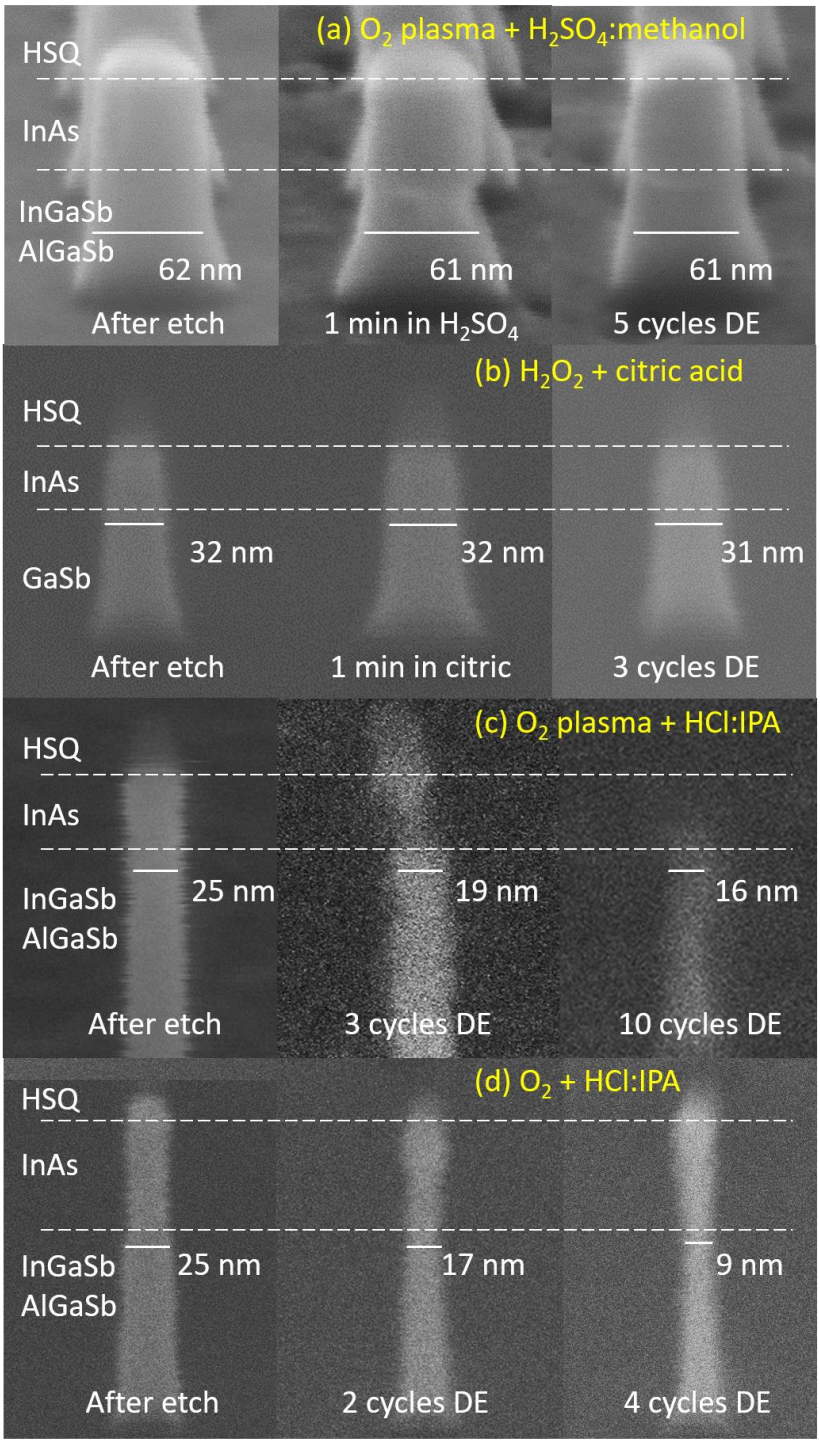


Fig. 2-18: InAs/InGaSb/AlGaSb vertical nanowire after RIE and various cycles of digital etch using: (a) O_2 plasma and H_2SO_4 in methanol, (b) H_2O_2 and citric acid, (c) O_2 plasma and HCl in IPA, and (d) O_2 atmosphere and HCl in IPA.

2.3.3 Alcohol-based Digital Etch on InGaAs

This novel non-aqueous digital etch has shown tremendous benefits on InGaAs-based devices as well. As previously mentioned, high aspect ratio InGaAs vertical nanowires with sub-10 nm diameter have never been demonstrated, due to the strong mechanical force that the wires experienced during the wet etch part of the digital etch. Using an alcohol-based solution can effectively reduce the surface tension in the acids so enhance the survivability of very narrow nanowires. For example, water has a surface tension of 72 mN/m, but methanol and isopropanol have surface tension of 22 and 23 mN/m, respectively, which is less than a third of that of water.

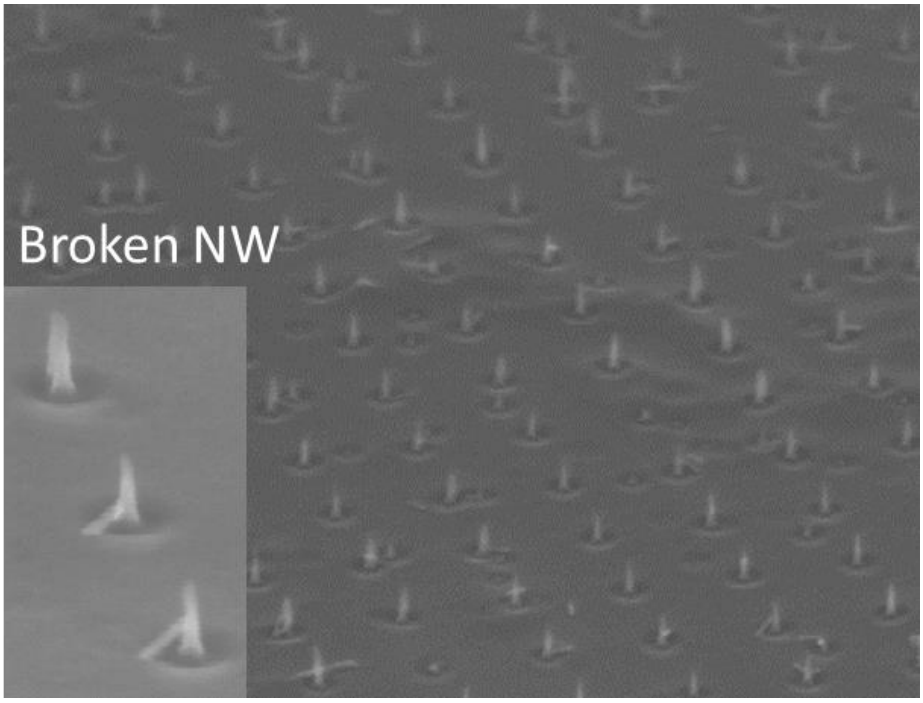


Fig. 2-19: InGaAs vertical nanowire array after 7 cycles of digital etch in 10% HCl in DI water.

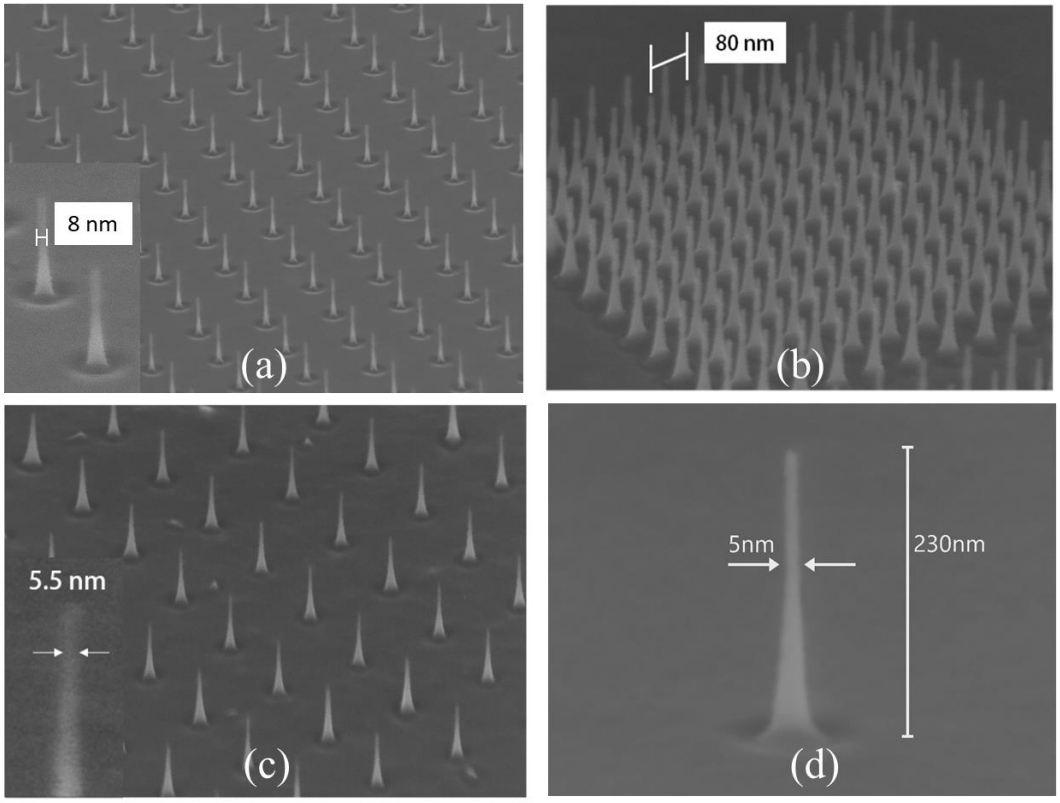


Fig. 2-20: (a) InGaAs VWN array after 7 cycles of digital etch in 10% HCl in IPA, with final diameter of 8 nm. The mechanical yield is over 97%. The inset shows a close-up of the etched structures. (b) Tightly-packed VNW array. The NW diameter is 10 nm and spacing is 80 nm. (c) InGaAs VNW array after 7 cycles of digital etch in 10% H_2SO_4 in methanol. The inset shows a close-up. The nanowire mechanical yield is 90%. (d) Narrowest InGaAs VNW with 5 nm diameter, 230 nm height obtained after 10 cycles of digital etch in H_2SO_4:methanol.

Three oxide removal chemistries are studied: 4.1 M H_2SO_4 in DI water, 2.0 M H_2SO_4 in methanol, and 0.1 M HCl in isopropanol (IPA). In all cases, oxidation is performed in oxygen plasma for 3 minutes. The starting arsenide heterostructure consists of multiple layers of $In_{0.53}Ga_{0.47}As$ and $In_{0.52}Al_{0.48}As$ grown by MBE on an InP substrate. The etched InGaAs vertical nanowires is about 230 nm in height. After RIE, the samples are dipped in BOE for 30 seconds to remove the HSQ hardmask and native oxide. At this point, various cycles of digital etch are performed. Between

cycles, the nanowire diameter is measured by SEM to monitor the etch rate. Fig. 2-19 shows the nanowire array after 7 cycles of conventional water-based digital etch. The targeted final diameter is 8 nm. In this case, all the wires are broken (mechanical yield = 0%).

Fig. 2-20(a) shows an identical sample processed side-by-side in HCl:IPA. 8 nm diameter vertical nanowires are demonstrated with over 97% mechanical yield. Moreover, no clustering is observed for tightly spaced vertical nanowires [81]. This is revealed in Fig. 2-20(b) where a 10 nm diameter vertical nanowire array spaced by 80 nm is shown. Fig. 2-20(c) shows a vertical nanowire array of diameter of 5.5 nm, obtained by 7 cycles of digital etch in H_2SO_4 in methanol. Fig. 2-20(d) shows the narrowest InGaAs nanowire with minimal diameter of only 5 nm and height of 230 nm (aspect ratio = 46), obtained after 10 cycles of digital etch in H_2SO_4:methanol. This is the narrowest diameter and highest aspect ratio in any vertical nanowire obtained via a top-down approach. The narrowest part of nanowire (D = 5 nm) is uniform for half of the total height.

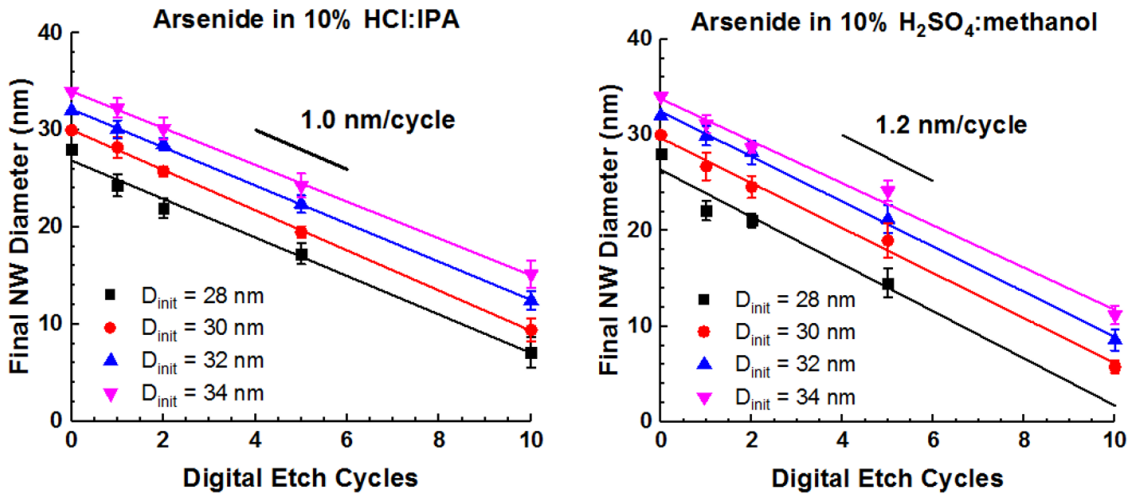


Fig. 2-21: Evolution of InGaAs VNW diameter with number of digital etch cycles in (left) 10% HCl:IPA, and (right) 10% H_2SO_4:methanol. The legends indicate the initial VNW diameter after RIE.

Etch-rates were obtained by measuring VNW diameter by SEM on sets of 12 NWs of various initial diameters. Figs. 2-21(a) and (b) show the evolution of InGaAs VNW diameter in alcohol-based HCl and H_2SO_4, respectively. The average radial etch rate in HCl:IPA is 1.0 ± 0.04 nm/cycle, and in H_2SO_4:methanol is 1.2 ± 0.05 nm/cycle. These values closely match the etch rate using H_2SO_4:DI water which is 1.0 ± 0.1 nm/cycle.

Fig. 2-22 summarizes the mechanical yield of arsenide VNWs after 7 cycles of digital etch with various nominal final diameters, for water as well as alcohol-based acids. The superiority of the new technique is clearly demonstrated. It should be noted that 7 cycles of digital etch cycles is a harsh process that is typically unnecessary in actual device fabrication. In a shortened process, higher mechanical yield at smallest diameter may be possible.

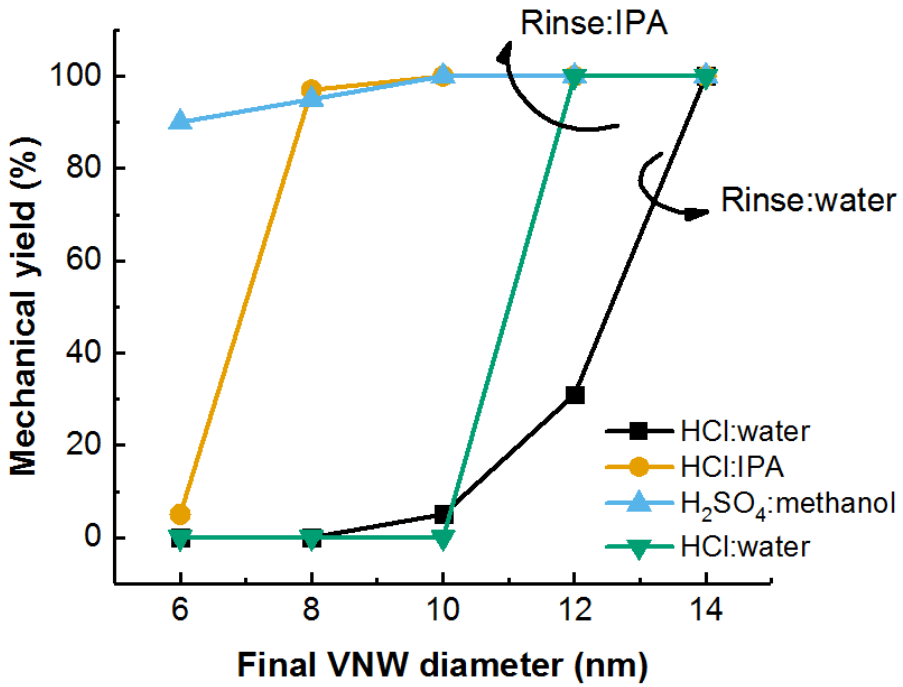


Fig. 2.22: Mechanical yield of InGaAs VNWs after 7 cycles of digital etch in various solutions and rinsing methods.

It is worth noting that, although there are other drying methods which can minimize surface tension, such as critical point drying (CPD), a solvent-based treatment is most compatible with CMOS manufacturing. Drying methods such as CPD are often time consuming, making them infeasible when multiple cycles of digital etch are required. In addition, breakage of VNWs can happen in both the etching and rinsing steps. To clarify the relative contributions, Fig. 2-22 includes the mechanical yield of methanolic H_2SO_4 digital etch, and aqueous HCl digital etch with IPA rinsing. Although IPA rinsing improves the mechanical yield of HCl:H_2O digital drastically at D = 12 nm, it shows no difference at and below 10 nm. Interestingly, the H_2SO_4:methanol digital etch shows much better yield at the smallest final diameter of 6 nm than the HCl:IPA digital etch, even though both methods are rinsed with IPA. This is shown in Fig. 2-20(c) and (d). This is likely due to the smaller viscosity of methanol (0.54 cP) than that of IPA (2.0 cP), which causes less drag force on the nanowire during etching. These results indicate that NW breakage occurs during the oxide etch process as well as while rinsing, and the oxide removal is the most aggressive step below D = 10 nm.

In summary, by developing a digital etch scheme utilizing alcohol-based etchants, digital etch on the antimonide material system is demonstrated for the first time. Moreover, this alcohol-based etching scheme exhibits excellent mechanical yield at sub-10 nm critical dimensions. This section, together with the previous one, focused on the fabrication of the intrinsic part of the transistor. The next section will focus on an extrinsic part of the device, the ohmic contact.

2.4 Ohmic Contact

2.4.1 III-V Ohmic Contacts

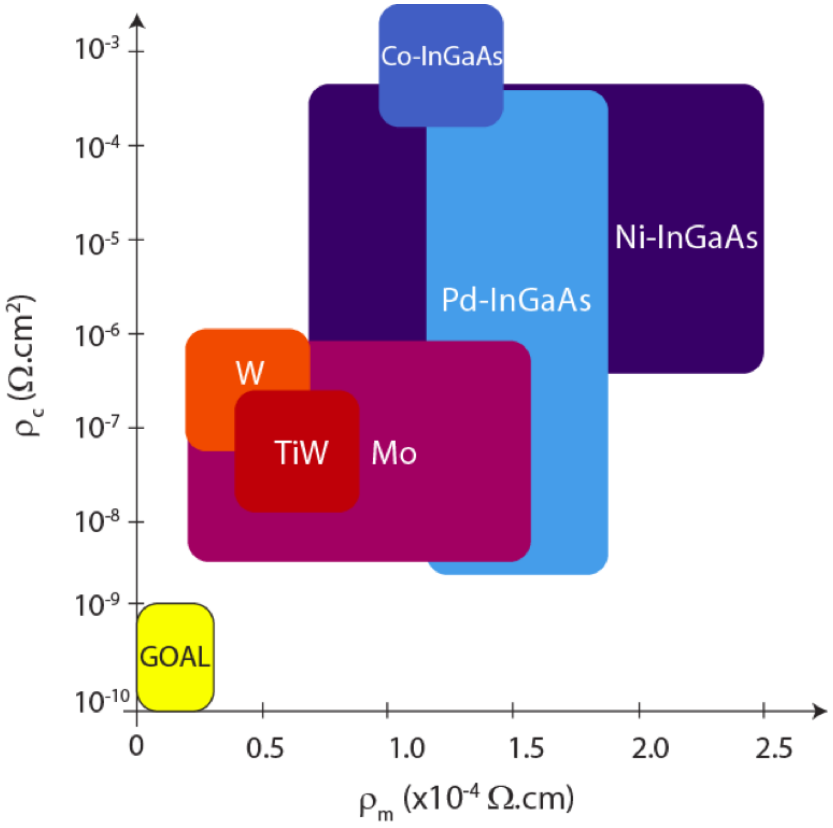


Fig. 2-23: Contact resistivity vs. metal film resistivity of Si-compatible ohmic contacts to n^+-InGaAs. The target regime is the bottom left corner.

As the scaling of the CMOS transistors advances by following Moore's law, the device pitches of Si and III-V logic MOSFETs have reached below 50 nm [1], [82]. At this device dimension, transistor performance is largely limited by extrinsic parasitics, such as the ohmic contact resistance. In current technologies, the transistor saturation current is estimated to be degraded by more than 40% by series resistance. At the insertion point, III-V logic transistors must have

contacts in the 10-15 nm length range while delivering a total source resistance 50 Ω·μm [83]. To meet this goal, the metal to semiconductor contact resistivity ρ_c needs to be lower than $5 \cdot 10^{-9}$ $\Omega \cdot cm^2$. Many experimental efforts have been made to optimize Si ohmic contacts to achieve ultra-low contact resistivities [18], [84], [85]. Low values of ρ_c for contacts on n^+-InGaAs between 0.4 and $3.2 \cdot 10^{-8}$ $\Omega \cdot cm^2$ have been reported [86]–[90]. Fig. 2-23 shows a benchmark of various ohmic contacts in the n^+-InGaAs system, with the goal regime noted in the lower left corner. Current III-V contact technology is still insufficient, but it is approaching the goal. On the other hand, a low-resistance ohmic contact technology is particularly lacking for the antimonide-based p-channel MOSFETs. In this section, the goal is to investigate and develop an ohmic contact technology for the InGaSb p-channel MOSFETs.

2.4.2 P^+ Capping Layer Design

For InGaSb-based p-channel MOSFETs, InAs is commonly used as the cap material due to the high etching selectivity that it presents with respect to other antimonide compounds. To our knowledge, the best reported contact resistivity on p^+-InAs is $1.6 \cdot 10^{-6}$ $\Omega \cdot cm^2$ using Pd alloyed contacts [91]. While there are low-resistance contacts for n-type FETs with n^+-InAs caps [92], the contact resistance on p^+-InAs is inadequate to meet the needs of a future nanoscale CMOS technology.

In this thesis, ohmic contacts for InGaSb MOSFETs using a novel p^+-InAs/InAsSb composite capping layer are fabricated and characterized. The motivation for this new cap structure stems from the lattice mismatch that typically exists between the InAs cap and the InGaSb, which has an

InSb composition that ranges from 20% to 40%. The resulting 1.2% to 3.2% lattice mismatch limits the InAs cap thickness. Thicker caps develop a large density of defects that yield a poor contact resistance. To address this problem, we introduce a thick p^+-InAsSb subcap that is lattice-matched to the channel under a thin p^+-InAs surface layer. This maintains a high crystallographic quality through the entire cap, resulting in better electrical characteristics.

Two study the effect of the capping layer designs, two InGaSb quantum-well FET heterostructures were grown by molecular beam epitaxy on GaAs substrate [26]. Fig. 2-24 shows a cross-section of one of them which features a highly-doped p-type InAs/$InAs_{0.85}Sb_{0.15}$ 5/30 nm bilayer cap (N_A = $1 \cdot 10^{19}$ cm^{-3} for both layers). The second heterostructure is identical, except for the cap, which is comprised of a 20 nm InAs single-layer with the same doping.

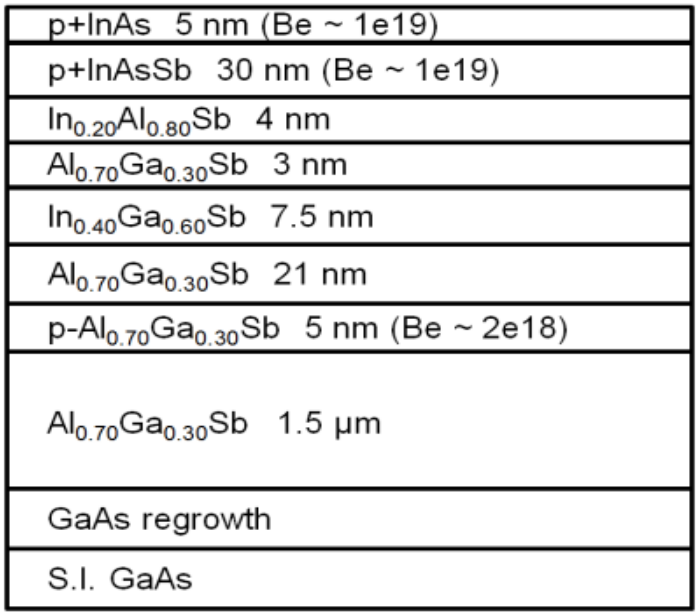


Fig. 2-24: Schematic of heterostructure with p^+-InAs/$In_{0.85}As_{0.15}Sb$ composite capping layer.

Conventional and circular transmission line model (TLM and CTLM, respectively) test structures were fabricated to characterize various ohmic contact schemes. The samples are first degreased with organic solvents and patterned by photolithography. They are then ashed for 5 minutes in oxygen plasma to remove residual resist and dipped in 10% HCl to remove the native oxide. The samples are immediately transferred to an electron-beam deposition system and brought to vacuum ($< 1 \cdot 10^{-6}$ Torr). Metal contacts are subsequently deposited and lifted-off in acetone. Mesa isolation is then performed by means of BCl_3 plasma etching (mesa isolation can be skipped for CTLM). Finally, sequential rapid thermal annealing (RTA) at increasing temperature is carried out on each sample in N_2 in 3 min steps. During RTA, samples are placed upside down on a GaAs proximity cap to prevent surface decomposition [93].

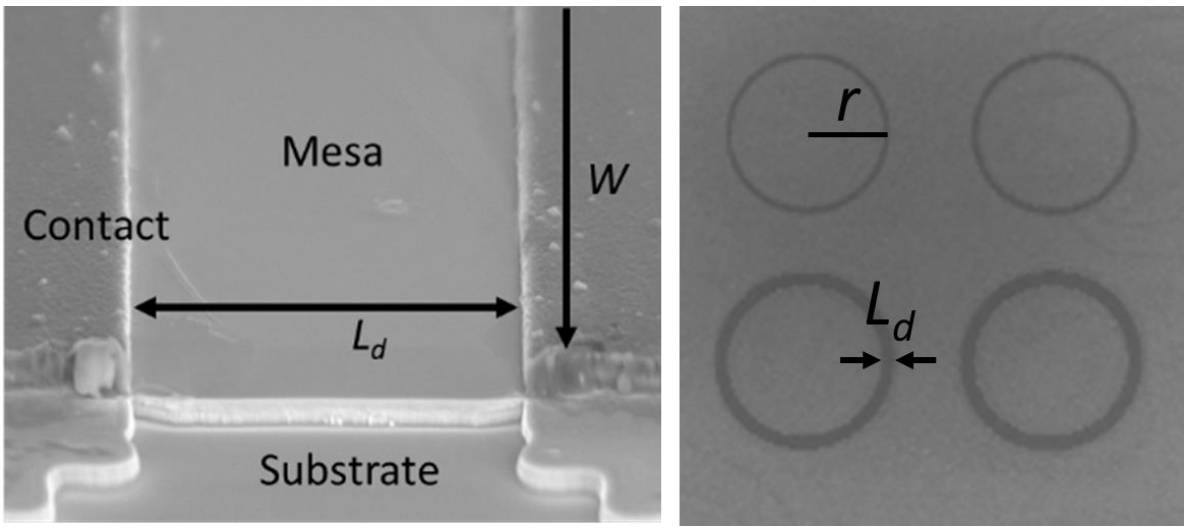


Fig 2-25: SEM picture of fabricated (left) TLM structure with Ni/Pt/Au (15/10/100 nm) contacts that are 6 µm apart, and (right) CTLM structures with the same Ni/Pt/Au contacts, with various contact spacings.

Fig. 2-25 shows SEM images of finished TLM and CTLM test structure. TLM characterization is performed using four-point probe measurements. Results obtained from TLMs and CTLMs were

found to be consistent. All sample dimensions were measured by SEM to ensure accuracy in extracting low values of the contact resistance, R_c.

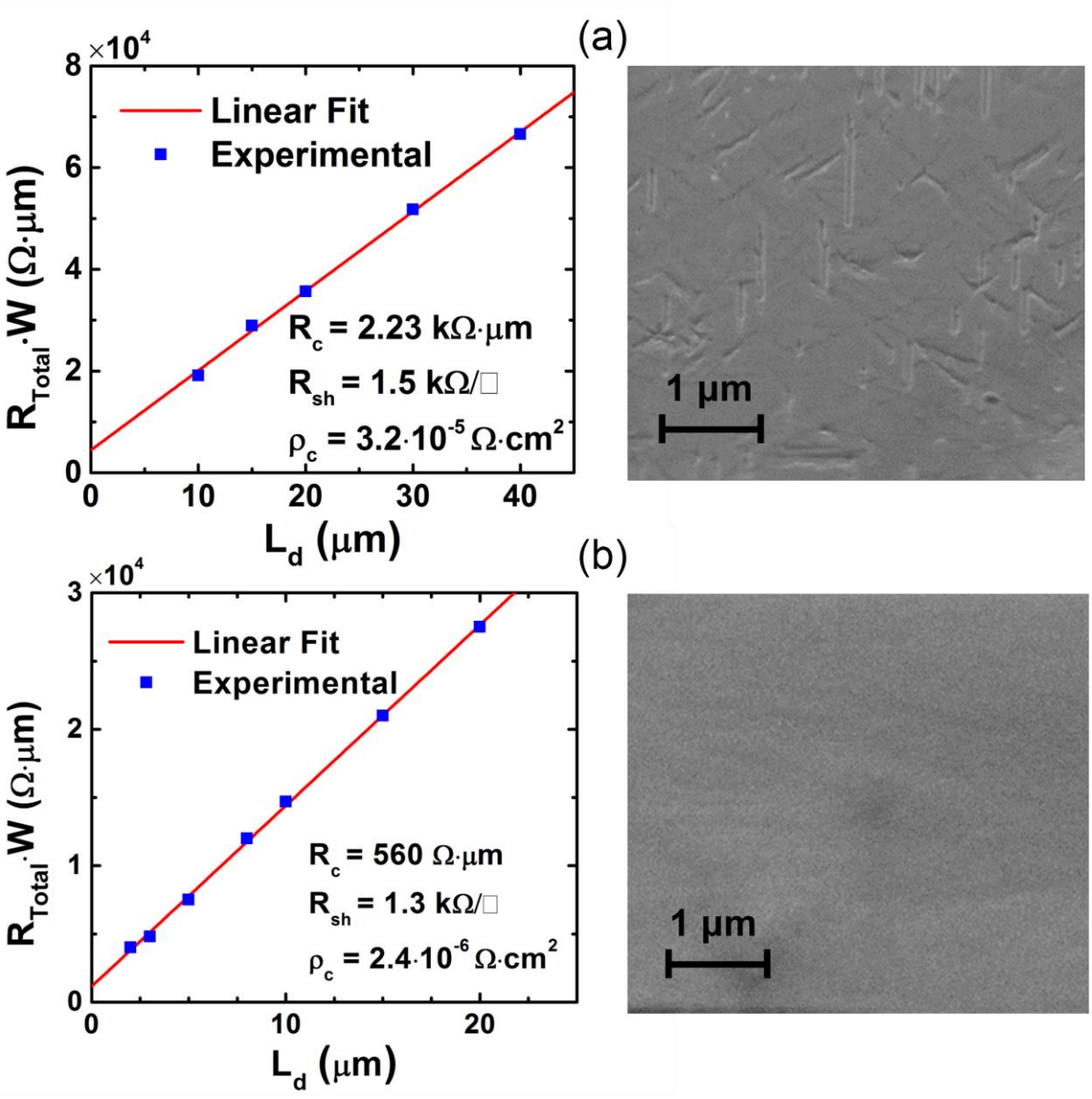


Fig. 2-26 Electrical measurements (left) and SEM images of virgin semiconductor surface (right) on Pd/Pt/Au TLMs on: (a) p^+-InAs cap structure, and (b) p^+-InAs/InAsSb cap structure.

Fig. 2-26 shows TLM data for Pd/Pt/Au (10/10/150 nm) contacts on both heterostructures, annealed at 200°C. The sample with the InAs single-layer cap exhibits R_c of 2.23 kΩ·µm and R_{sh} of 1.56 kΩ/□, while the InAs/InAsSb bilayer cap yields R_c of 560 kΩ·µm and R_{sh} of 1.32 kΩ/□, representing a 4X improvement in the contact resistance. 200°C was selected as optimum annealing temperature based on the literature and our own sequential annealing experiments confirmed the same temperature behavior, as shown in Fig. 2-29(b).

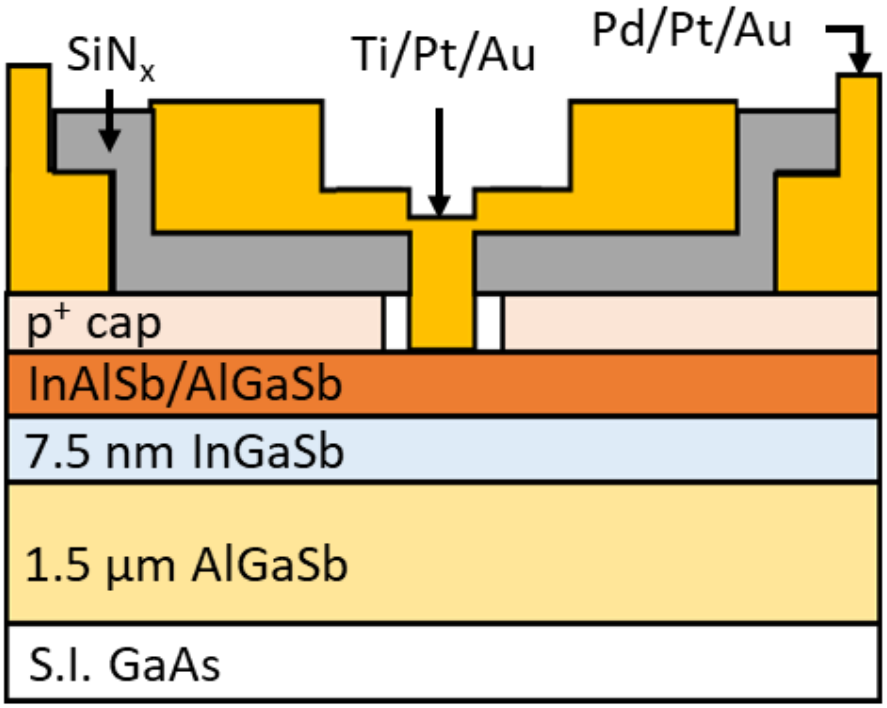


Fig. 2-27: Schematic cross-section of InGaSb p-channel QW-FETs fabricated. Device heterostructure is the same as in Fig. 2-24.

The improvement in the contact resistance in the composite cap is likely due to its better crystallographic quality. SEM examination of the sample with the pure InAs cap reveals a surface with line defects, as shown in Fig. 2-26(a), presumably due to excessive tensile stress in the InAs layer. Use of a lattice-matched InAsSb sub-cap mitigate the stress and results in a smooth surface,

as shown in Fig. 2-26(b). To verify that the observed defects in the InAs cap sample were limited to the cap, the InAs layer was removed in a 3:1 citric acid/H_2O_2 solution, with the 4 nm InAlSb layer serving as an etch stop. The exposed surface underneath was smooth and free of the line defects.

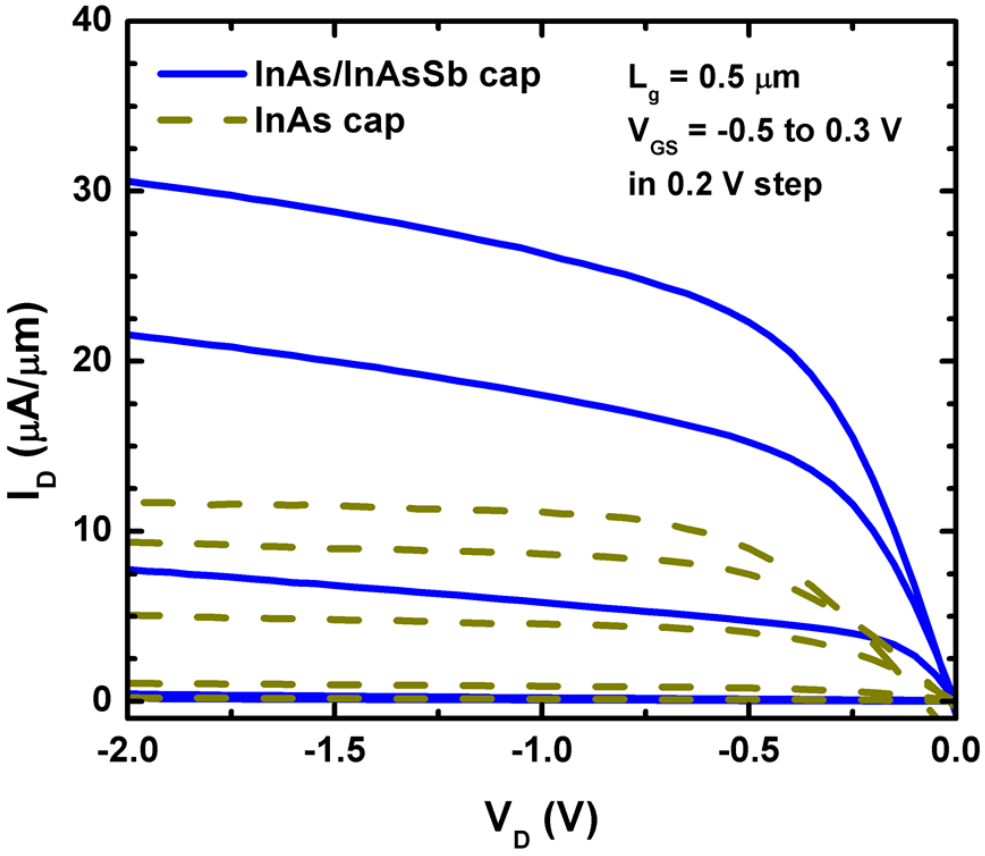


Fig. 2-28: Output characteristics of InGaSb QW-FETs with L_g = 0.5 μm and with both cap designs. V_{GS} is swept from -0.5 V to 0.3 V, in -0.2 V steps.

To confirm the device worthiness of the new cap design, InGaSb p-channel quantum-well-FETs (QW-FETs) with L_g = 0.5 μm using both cap structures were fabricated. Fig. 2-27 shows a cross-section schematic of the finished device. After standard cleaning, Pd/Pt/Au contacts are deposited, lifted-off, and annealed at 200°C as in the TLM process. Then, 50 nm of SiN_x is deposited via

PECVD. The 0.5 µm gate opening is defined by electron-beam lithography and the dielectric is patterned by SF_6/O_2 plasma etching. The cap is recessed in 3:1 citric acid:H_2O_2 solution. The Ti/Pt/Au gate metal is patterned and deposited by electron-beam evaporation and lifted-off. Finally, mesa etch is done using lactic acid/H_2O_2 wet etch.

Fig. 2-28 shows the output characteristics of a finished InGaSb p-channel QW-FETs. The device with the composite cap exhibits a drain current that is more than double that of the device with the InAs-only cap. Both devices have ungated region of 2 µm on both sides of the channel. The InAs cap sample has R_{on} of 34.4 kΩ·µm and a peak g_m of 26 µS/µm at V_{DS} = -2 V. By contrast, the bilayer cap sample has R_{on} of 12.8 kΩ·µm and a peak g_m of 72 µS/µm at V_{DS} = -2 V. These results demonstrate that the greatly improved contact resistance obtained in the new cap structure does translate into major enhancements in device electrical characteristics.

2.4.3 Nickelide Contacts and Nano-TLM Characterizations

It has recently been shown that Ni alloyed contacts are promising candidates for FETs using both n+-InAs and p+-GaSb [18] cap structures. After annealing, Ni reacts with InAs to form an alloyed NiInAs layer like in silicide contacts, and therefore it is sometimes called "nickelide". This motivates us to explore Ni/Pt/Au (15/10/100 nm) contacts on the new p^+ InAs/InAsSb cap structure. Fig. 2-29 shows CTLM resistance measurements and the extracted value of ρ_c in sequential annealing experiments. At 350°C, as shown in Fig. 2-29(a), we obtained the lowest contact resistance of 45.3 Ω·µm. This corresponds to $\rho_c = 1.3 \cdot 10^{-8}$ Ω·cm^2. Above 350°C, the metal starts to delaminate, and the contact resistance increases rapidly. Similar contact resistance

degradation at around this temperature has also been observed in the Ni-InGaAs contact system [94], [95]. In 6 sets of test structures consisting of a total of 72 CTLMs, the average ρ_c obtained after 350°C anneal is $4.5 \cdot 10^{-8}$ Ω·cm^2 with a standard deviation of $3.1 \cdot 10^{-8}$ Ω·cm^2. The average R_c is 70 Ω·μm.

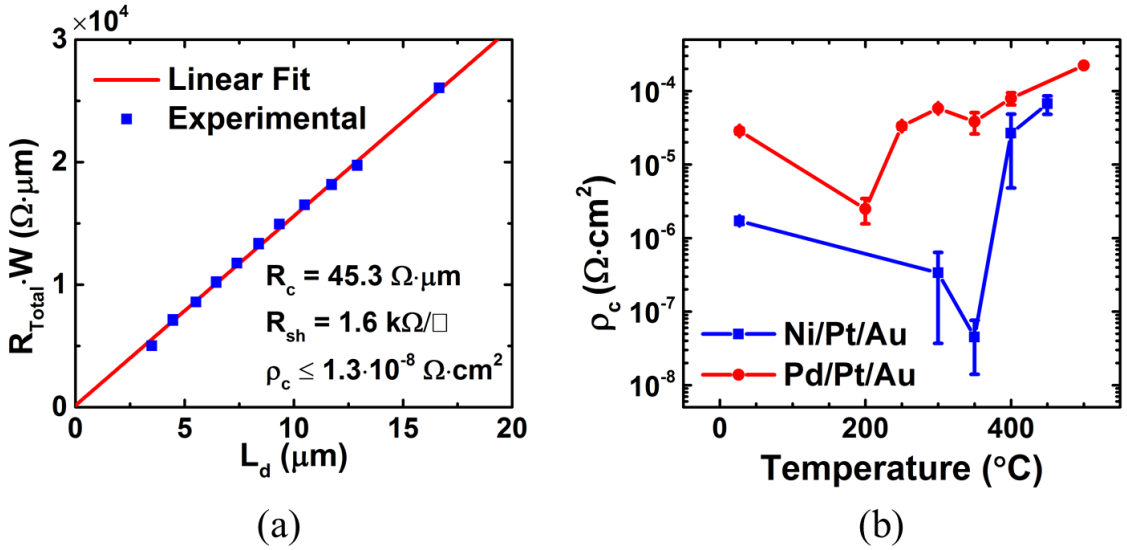


Fig. 2-29: (a) Electrical measurements on a CTLM with Ni/Pt/Au contacts on the p^+-InAs/InAsSb cap structure after 3 min 350°C annealing. (b) Evolution of contact resistivity of Pd/Pt/Au and Ni/Pt/Au contacts in a sequential annealing (3 min) experiment.

To characterize low contact resistance ohmic contacts, it is critical to note that the extraction of R_c is extremely sensitive to the experimental data. As shown in Fig. 2-29, R_c is extracted at the intercept of zero contact spacing, L_d. Here, R_c is very close to zero, and the contact spacing of the conventional TLM structures is larger than 3 μm. Therefore, accurate extraction of small R_c is challenging. To address such challenge, we developed a novel test structure called nano-TLM. Fig. 2-30 (a) shows a schematic of the nano-TLM test structure. The shaded area indicates the active semiconductor region, on which there are two parallel thin metal contacts of length L_c, width W,

separated by a distance L_d. A unique aspect of this structure is that the two nano-scale metal lines can be contacted at both ends separately. This enables the two Kelvin measurement schemes illustrated in Fig. 2-30(b). $R_{||}$, the parallel terminal resistance, is typical of the traditional TLM structure. This structure additionally allows the measurement of $R_{\times}$, the cross terminal resistance. The combination of $R_{||}$ and $R_{\times}$ enables accurate extraction of the contact resistance, R_c, semiconductor sheet resistance, R_{sh}, and ohmic metal sheet resistance, R_{shm}. In actual measurements, the forcing and sensing ports are interchanged so that four sets measurements per elemental test structure can be performed. Another advantage of nano-TLM is that the two contacts can be placed closer at sub-micron L_d, which allows more accurate extraction of R_c. The terminal resistances, $R_{||}$ and $R_{\times}$ can be expressed as:

$$R_{||} = \frac{R_{TLM}}{L_{Tx}} \operatorname{csch}\left(\frac{W}{L_{Tx}}\right) \tag{5}$$

$$R_{\times} = \frac{R_{TLM}}{2L_{Tx}}\left[\operatorname{csch}\left(\frac{W}{L_{Tx}}\right) + \coth\left(\frac{W}{L_{Tx}}\right)\right] - \frac{R_{shm}W}{2L_c} \tag{6}$$

Where L_{Tx} is the transfer length in the longitudinal direction along the metal contacts,

$$L_{Tx} = \sqrt{L_c \cdot \frac{R_{TLM}}{2R_{shm}}} \tag{7}$$

, and R_{TLM} is the conventional TLM resistance.

$$R_{TLM} = 2R_c + R_{sh} \cdot L_d \tag{8}$$

Note that as R_{shm} goes to zero, both (5) and (6) converge to R_{TLM}/W, which coincides with the classic normalized TLM result.

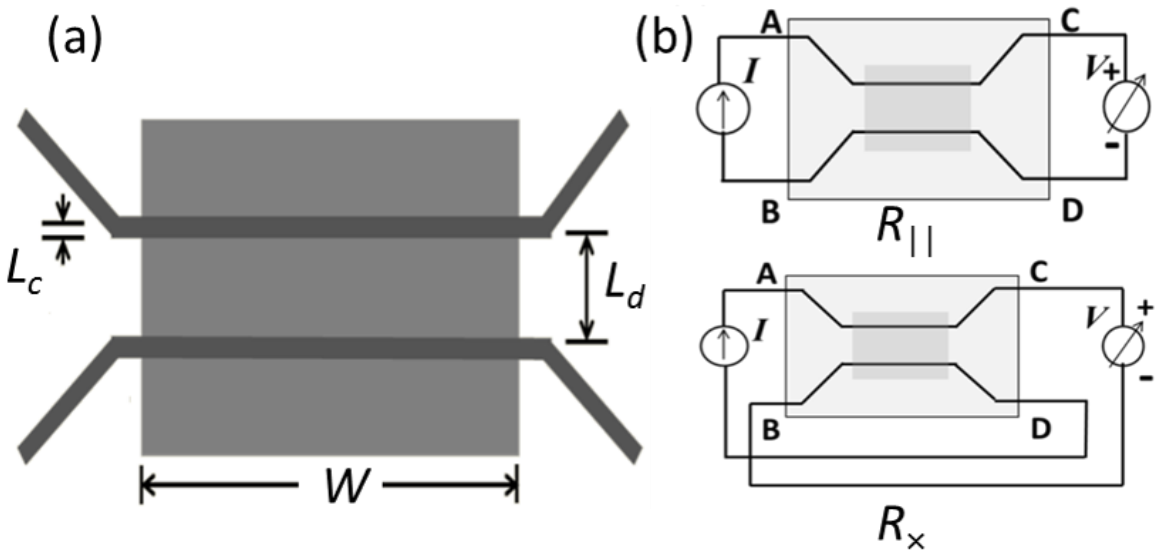


Fig. 2-30: (a) Schematic top down view of a single nano-TLM test structure. (b) Two Kelvin measurement schemes used to characterize the nano-TLM test structure.

Hence, we also fabricated nano-TLMs with contacts length as small as L_c = 80 nm and contact separation down to L_d = 130 nm, patterned by electron-beam lithography. Fig. 2-31(a) shows a fabricated nano-TLM test structure. Fig. 2-31(b) shows measurements of nano-TLMs with nano-contacts of average measured L_c of 100 nm. Because $R_{||}$ and $R_\times$ differ by less than 1%, only data for $R_{||}$ are displayed. From them, we can extract an average R_c of 80.9 Ω·μm (for $L_c > L_T$, the transfer length) and an average ρ_c of $5.3 \cdot 10^{-8}$ Ω·cm^2, with standard deviations of 10 Ω·μm and $2.4 \cdot 10^{-8}$ Ω·cm^2, respectively. This, to the author's knowledge, is the lowest reported contact resistivity to p-type InAs. The nano-TLM results verifies the low R_c extracted by conventional CTLM measurements. The transfer length of the Ni ohmic contacts is estimated about 60 nm, based on the ρ_c and R_{sh} extracted by the nano-TLMs.

Table 3-2 summaries the Au-based ohmic contacts schemes that have been attempted in this thesis work and in the literature on other types of capping semiconductors, such as GaSb, GaAsSb, and InGaSb.

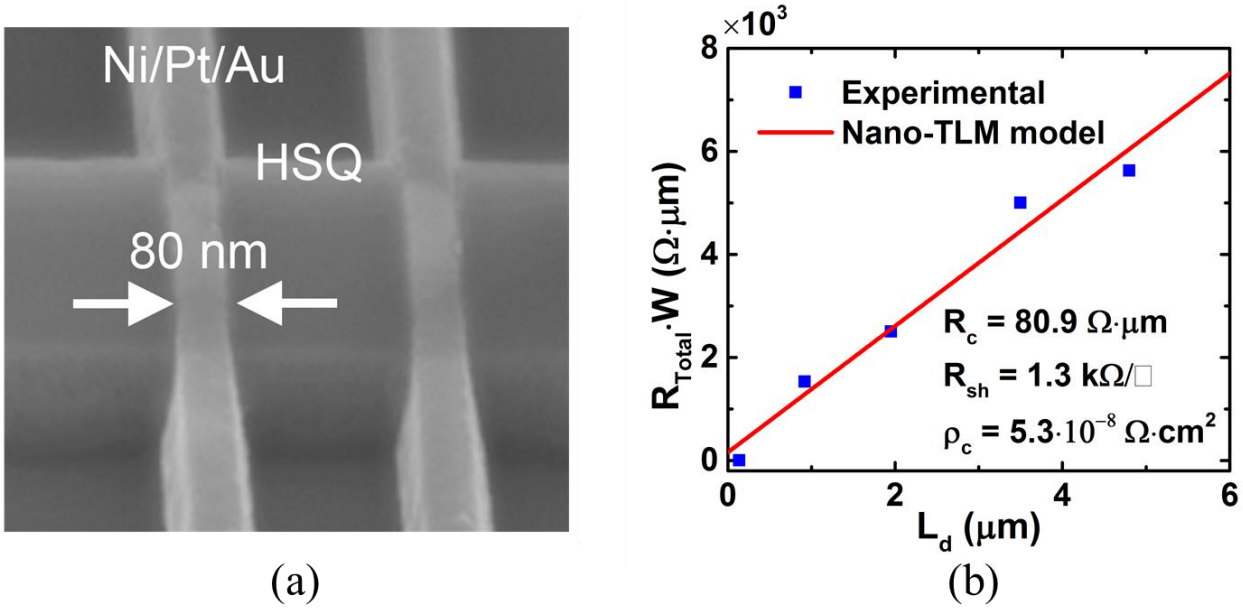


Fig. 2-31: (a) SEM image of a Ni/Pt/Au nano-TLM test structure, with 80 nm contact length, 130 nm contact spacing, and 1 μm width. (b) Electrical measurements in nano-TLMs with nano-contacts with average L_c of 100 nm.

The work of Oxland et al. indicates that under similar conditions to ours, Ni on InAs forms a shallow reaction region of 8.5 nm depth [96], as shown in the right TEM image in Fig. 2-32. We also observed a very shallow nickelide region of approximately 10 nm in other III-V compounds such as GaAsSb after 1 min RTA at 350°C, as shown in the left TEM image in Fig. 2-32. This suggests that the top InAs layer in our cap structure is mostly consumed in the reaction and that a direct contact exists between the reacted region and the underlying InAsSb layer. This might be the key behind these excellent results.

Semiconductor	Metal	N_A (cm^{-3})	R_c (Ω·μm)	ρ_c (Ω·cm^2)	Reference
InAs	Pd/Pt/Au	2.0E+19	N/A	1.6E-06	Lysczek, 2006
		1.0E+19	2230	3.2E-05	Our work
InAs/InAsSb	Pd/Pt/Au	1.0E+19	560	2.4E-06	Our work
	Ni/Pt/Au		45.3	< 1.3E-08	
	Mo/Ti/Au		1700	1.6E-05	
	Ni/Ti/Pt/Al		93	4.1E-08	
GaSb	Ti/Pt/Au	6.6E+16	N/A	5.8E-06	Vogt, 1996
	Ni/Au	6.2E+18	N/A	1e-5~3e-5	Vogt, 1999
	Pd/Au		N/A	3e-7~3e-6	
	Pd/Pt/Au	2.0E+19	12.7	7.4E-06	Our work
	Mo/Ti/Au		8.7	2.9E-06	
	Ni/Pt/Au		10.2	3.0E-06	
GaAsSb	Ni/Pt/Au	2.0E+19	36.1	1.7E-07	Our work
InGaSb	Pd/W/Au	1.8E+18	80	3.0E-07	Wang, 2003

Table 2-2: Summary of ohmic contact schemes for InGaSb p-channel MOSFETs in this thesis work and in relevant literature.

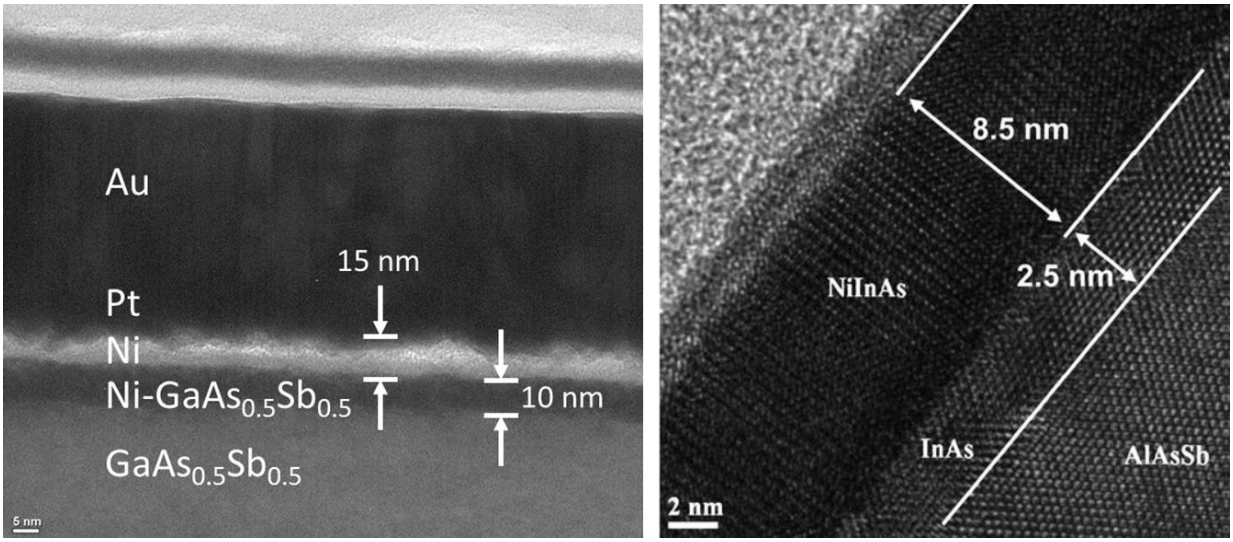


Fig. 2-32: Cross-section TEM images of (left) Ni-GaAsSb contact region after 1 min RTA at 350°C, and (right) Ni-InAs contact after 1 min RTA at 350°C [40].

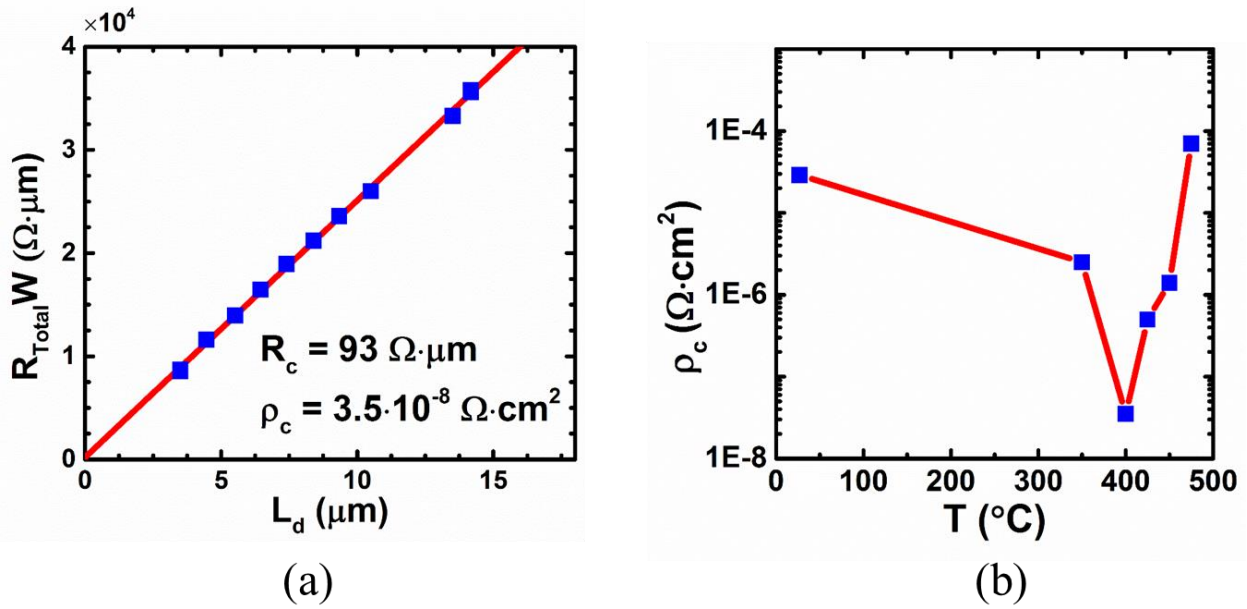


Fig. 2-33: (a) Si compatible Ni/Ti/Pt/Al ohmic contacts to p^+-InAs measured by circular TLM. (b) ρ_c evolution vs. annealing temperature in sequential annealing experiments.

2.4.4 Si-compatible Nickelide Contacts

As we have found that Ni/Pt/Au as a promising contact scheme to p^+-InAs, we have also investigated Si-compatible ohmic contact schemes using either Ni alone or a Ni/Ti/Pt/Al stack. Similar circular TLM test structures were fabricated with a Ni/Ti/Pt/Al (15/10/15/100 nm) contact stack on the p^+-InAs/$InAs_{0.85}Sb_{0.15}$ composite cap ($N_A = 1 \cdot 10^{19}$ cm^{-3}). Fig. 2-33 shows contact resistance measurements, and the evolution of ρ_c versus annealing temperature in sequential annealing experiments. A minimum contact resistivity of $3.5 \cdot 10^{-8}$ $\Omega \cdot cm^2$ is obtained after one-minute annealing at 400°C. This is the first demonstration of Si-compatible contacts to p^+-InAs with ultra-low contact resistivity. This is the contact process that is incorporated in the first and second generation of InGaSb p-channel FinFETs described elsewhere in this thesis.

In the third generation of the InGaSb p-channel FinFETs, the doping of the composite capping layer is increased from $1 \cdot 10^{19}$ cm^{-3} to $3 \cdot 10^{19}$ cm^{-3} to further improve the contact resistance. Fig. 2-34 (a) shows the ohmic contact resistances, measured by circular TLM test structures. Fig. 2-34 (b) benchmarks R_c obtained by Mo, W, and Ni contacts as a function of annealing temperature. Ni ohmic contacts annealed at 350°C for 3 min yield a record contact resistance of 22.3 Ω·μm, which corresponds to an ultra-low contact resistivity of $4.6 \cdot 10^{-9}$ Ω·cm^2. This R_c is a factor of 4X better than the previous demonstration [97], and it arises from the use of a composite p$^+$-cap and a much higher doping level.

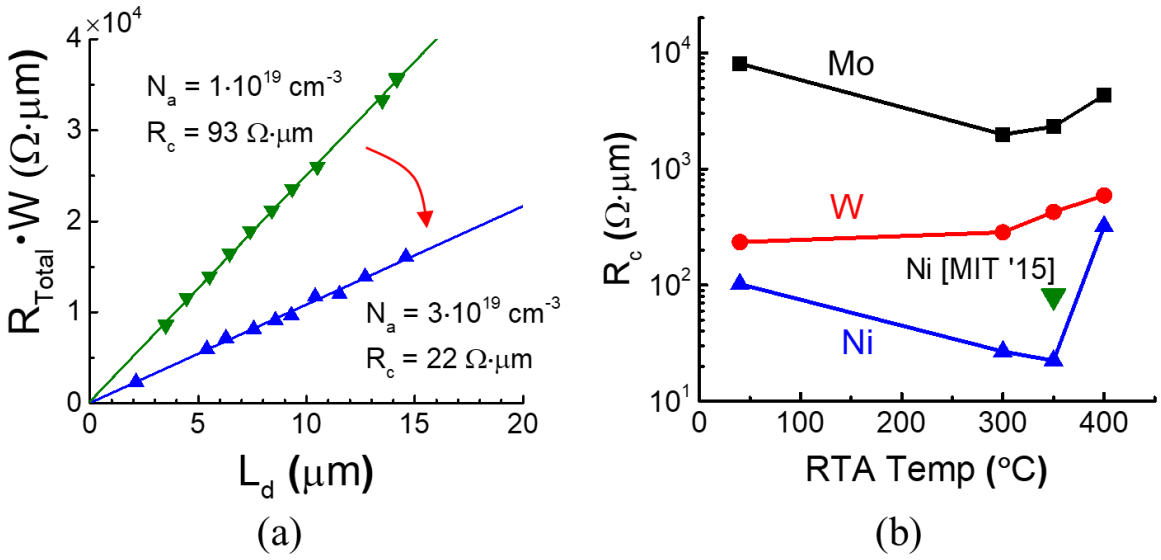


Fig. 2-34: (a) Ni TLM resistance on p$^+$-InAs with $N_A = 1 \cdot 10^{19}$ cm^{-3} and $3 \cdot 10^{19}$ cm^{-3}. (b) R_c benchmark of Mo, W, and Ni ohmic contacts and earlier results.

2.5 Chapter Summary

In this chapter, we have discussed three key process components required in InGaSb p-channel FinFET fabrication. First, a fin reactive ion etching technology for InGaSb devices has been developed for the first time. This novel RIE technology is able to etch antimonide-based and arsenide-based fins and vertical nanowires simultaneously, with high aspect ratio (W_f < 20 nm, H_f > 200 nm), vertical profile (> 85°), and smooth sidewall.

Second, to mitigate dry-etch damage and further scale-down the device size, an antimonide-compatible digital etch process has been developed for the first time, as well. The antimonide-compatible digital etch features a novel fully alcohol-based process and shows a consistent digital etch rate of 2 nm/cycle on the antimonides. Antimonide fins and vertical nanowires with fin width of 10 nm and diameter of 9 nm have been demonstrated using this technique. Also, alcohol-based digital etch addresses the limitations of the conventional digital etch approach in enabling structures with sub-10 nm 3D features. It shows remarkable improvement in the mechanical yield of the nanowires at sub-10 nm regime. Narrowest InGaAs nanowires with diameter of 5 nm and height of 230 nm (aspect ratio = 46) have been demonstrated.

Last, ohmic contacts on antimonide-based p-channel MOSFETs have been systematically studied and surveyed. To improve the contact quality, a novel composite InAs/InAsSb capping layer has been introduced to relieve the stress-induced defects. Si-compatible nickelide contacts have been fabricated with contact resistance of 93 Ω·μm. With increased p^+ doping level at $3 \cdot 10^{19}$ cm^{-3}, the

contact resistance is further reduced to 22 Ω·μm. This corresponds to a record-low contact resistivity of $4.6 \cdot 10^{-9}$ Ω·cm^2.

The critical technologies described in this chapter lay the foundation for the device integration and demonstration in Chapter 3 and Chapter 4. Using these process components, the demonstration of the first InGaSb p-channel FinFETs becomes possible. The detailed process integration of InGaSb FinFETs with be discussed in the next chapter.

CHAPTER 3. InGaSb FinFET Fabrication

3.1 Process Overview

In the previous chapter, the key missing technological components of InGaSb p-channel FinFETs are developed. The aim of this chapter is the development and integration of a self-aligned FinFET device architecture that is suited for logic applications. Previous works in our group on InGaAs n-channel planar MOSFETs and InGaSb planar p-channel QW-FETs [31], [98], [99] have developed a successful III-V fabrication template that is well suited as a starting point for the InGaSb FinFET fabrication.

In this thesis, a contact-first, gate-last process scheme is chosen. There are several advantages with this approach. First, a contact-first process helps produce lower contact resistance. Depositing contact metal on the fresh sample surface minimizes surface contamination and damage during the process. Second, a contact-first scheme makes a self-aligned process relatively easy. For a gate-first or fin-first process, processes such as dummy-gate, ion implantation, or III-V regrowth are needed. Second, a gate-last process leaves the fin RIE, digital etch and gate stack deposition near the end of the process, so that the sidewall is exposed to fewer processing steps, especially for processes involving high temperature or plasma.

There are also several disadvantages. First, a contact-first process makes it difficult to remove the HSQ fin hardmask. This is because a spacer layer is needed to prevent a short between gate and source/drain contacts, and the spacer (typically SiO_x or SiN_x) may have poor etching selectivity to HSQ. Second, since the contact metal is deposited first, the highest temperature in the rest of the process cannot be excessively high. In this case, for nickelide contacts, the processing temperature should not exceed 400°C in the subsequent steps, indicated by the annealing experiment in the previous chapter (Fig. 2-33).

In this chapter, three generations of InGaSb p-channel FinFET fabrication processes will be discussed. The first generation will be described in detail as a generic process flow. Processes that are different from this generic process flow in later generations will be described separately. The chapter starts with the designs of the starting InGaSb heterostructures.

3.2 Heterostructure Designs and Growth

A description of the starting heterostructure used for the first generation InGaSb p-channel FinFET fabrication in this work is shown in Fig. 3-1. The wafer was grown by molecular beam epitaxy (MBE) by Sandia National Laboratories. It consists of a 10 nm $In_{0.27}Ga_{0.73}Sb$ quantum-well channel on an $AlAs_{0.16}Sb_{0.84}$ buffer grown by MBE on (100) semi-insulating GaAs substrate. The $In_{0.27}Ga_{0.73}Sb$ channel is compressively stressed at 2.26%. There is a 5 nm Be delta-doping layer ($N_A = 1 \cdot 10^{12}$ cm^{-3}) 5 nm below the $In_{0.27}Ga_{0.73}Sb$ channel, within the $AlAs_{0.16}Sb_{0.84}$ buffer. For the first generation FinFETs, a single p^+-InAs cap layer is used for simplicity of the growth.

	Layer	Thickness (nm)	Dopant	Level (cm^{-3})
cap	InAs	30	Be	$1 \cdot 10^{19}$
etch stop	$In_{0.2}Al_{0.8}Sb$	4		UID
channel	$In_{0.27}Ga_{0.73}Sb$	10		UID
buffer	$AlAs_{0.16}Sb_{0.84}$	5		UID
	δ-doping	5	Be	$2 \cdot 10^{18}$
	$AlAs_{0.16}Sb_{0.84}$	1000		UID
	GaAs Substrate			S.I.

Fig. 3-1: Description of the starting heterostructure used in the first generation InGaSb p-channel FinFET. The wafer was grown by the Sandia National Laboratories.

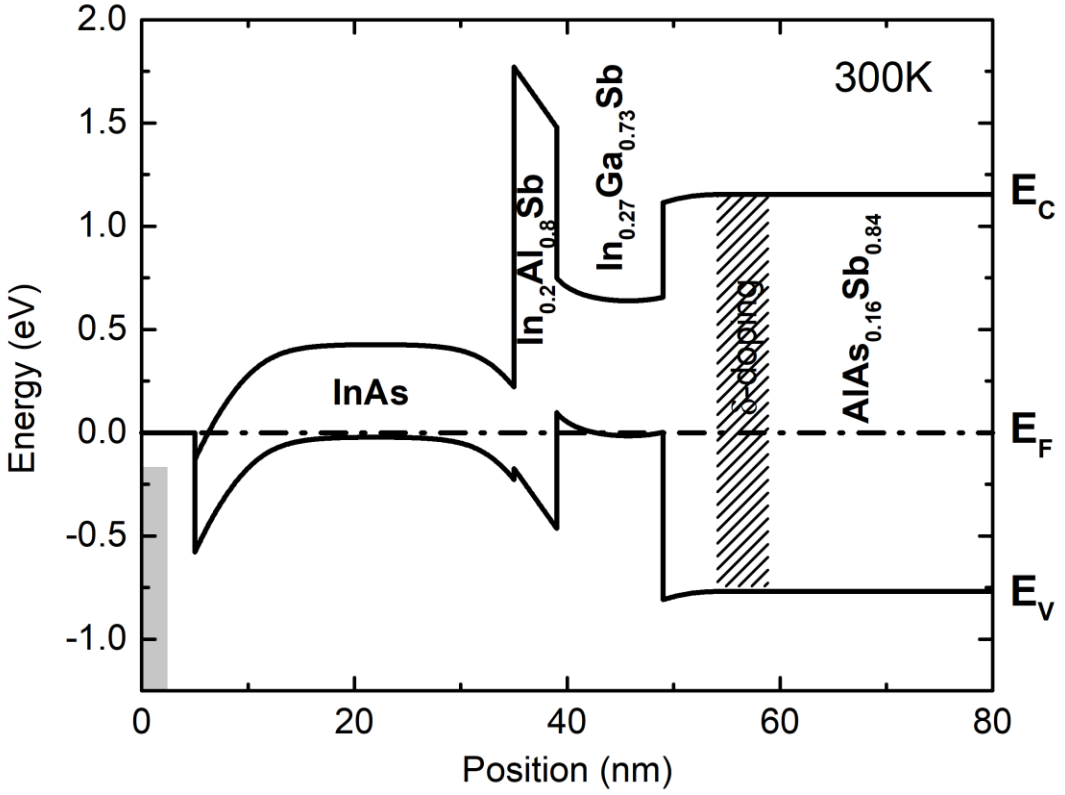


Fig. 3-2: Calculated band structure of the heterostructure in Fig. 3-1.

The 4 nm $In_{0.2}Al_{0.8}Sb$ layer below the p^+-InAs serves as the etch stopper for the recess etching of the cap and also serves as a passivation layer for the $In_{0.27}Ga_{0.73}Sb$ channel. Fig. 3-2 shows the band diagram of the heterostructure at zero bias at room temperature, calculated by 1-D Poisson-Schrödinger software Nextnano. The Fermi-level pinning position for the InAs cap is adopted from [100], which is assumed to be pinned within the conduction band.

	Layer	Thickness (nm)	Dopant	Level (cm^{-3})
cap	InAs	5	Be	$1 \cdot 10^{19}$
cap	$InAs_{0.85}Sb_{0.15}$	30	Be	$1 \cdot 10^{19}$
etch stop	$In_{0.2}Al_{0.8}Sb$	4		UID
	$Al_{0.7}Ga_{0.3}Sb$	3		UID
channel	$In_{0.4}Ga_{0.6}Sb$	7.5		UID
buffer	$Al_{0.7}Ga_{0.3}Sb$	21		UID
buffer	δ-doping	5	Be	$2 \cdot 10^{18}$
buffer	$Al_{0.7}Ga_{0.3}Sb$	1500		UID
	GaAs Substrate			S.I.

Fig. 3-3: Description of the starting heterostructure used in the second generation InGaSb p-channel FinFET. The wafer was grown by the Naval Research Laboratory.

Fig. 3-3 and Fig. 3-4 shows the starting heterostructures used for the second and third generations of InGaSb p-channel FinFET fabrications. The two structures were grown by molecular beam epitaxy by Naval Research Laboratories and the Korea Institute of Science and Technology, respectively. Both of the heterostructures incorporate the composite p^+-InAs/$InAs_{0.85}Sb_{0.15}$ cap, as described in Chapter 2. In addition, both designs were grown on AlGaSb buffer instead of AlAsSb. In Fig. 3-4, the third generation heterostructure, a composite $Al_{0.8}Ga_{0.2}Sb$/$Al_{0.93}Ga_{0.07}Sb$ is grown to increase the resistivity of the buffer.

	Layer	Thickness (nm)	Dopant	Level (cm^{-3})
cap	InAs	5	Be	$3 \cdot 10^{19}$
cap	$InAs_{0.85}Sb_{0.15}$	30	Be	$3 \cdot 10^{19}$
etch stop	$In_{0.2}Al_{0.8}Sb$	6		UID
channel	$In_{0.25}Ga_{0.75}Sb$	23		UID
buffer	$Al_{0.8}Ga_{0.2}Sb$	5		UID
buffer	δ-doping	5	Be	$1 \cdot 10^{18}$
buffer	$Al_{0.8}Ga_{0.2}Sb$	30		UID
buffer	$Al_{0.93}Ga_{0.07}Sb$	970		UID
	GaAs Substrate			S.I.

Fig. 3-4: Description of the starting heterostructure used in the third generation InGaSb p-channel FinFET. The wafer was grown by the Korean Institute of Science and Technology.

The design of channel layer is different for these two heterostructures, as well. For the second generation FinFETs, the indium composition of the $In_xGa_{1-x}Sb$ is increased from 27% to 40%. Higher indium composition increases the hole mobility in the channel. The channel is therefore compressively strained at 2.01%. The trade-off is the lower critical thickness allowed for the channel layer due to the larger lattice mismatch. In this case, the channel height is reduced to 7.5 nm. In the third generation FinFETs, the indium composition is reduced to 25%. This reduction allowed an $In_{0.25}Ga_{0.75}Sb$ layer as thick as 23 nm, to take fuller advantage of a FinFET device architecture. In this case, the channel is less compressively strained, at 0.96%.

Fig. 3-5 shows the calculated band structure of the third generation heterostructure at room temperature. Both of the thickness of the InGaSb channel and the level of compressive strain have

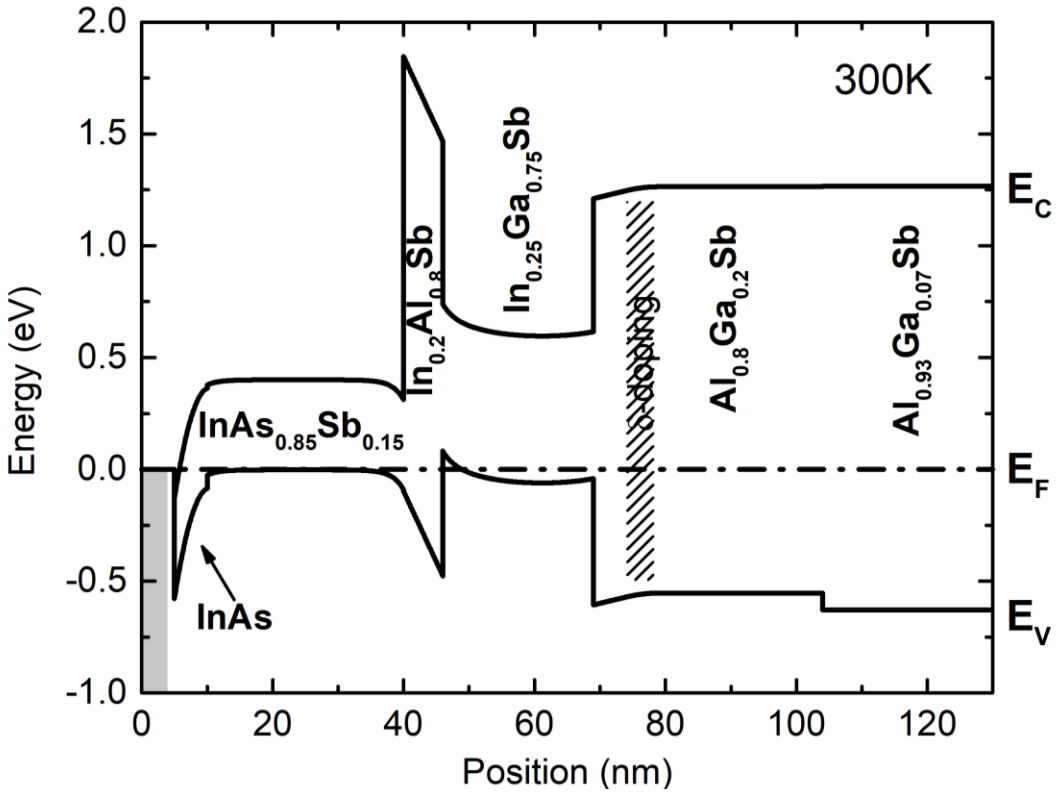


Fig. 3-5: Calculated band structure of the heterostructure in Fig. 3-4.

an impact on the hole mobility in the channel. As shown in Fig. 3-6 [101], for an InGaSb channel with higher strain, the hole mobility is more sensitive to the thickness of the channel. The thicker the channel layer, the lower the mobility. The degraded mobility in thick InGaSb channel is due to lattice relaxation. Therefore, the design of channel thickness and strain needs to be carefully chosen for the FinFET architecture.

Antimonide-based heterostructure epitaxial growth is particularly challenging because of the high lattice mismatch between the epi-layers and the most common substrates, typically GaAs or InP. In this thesis work, it is very fortunate that we have worked with groups with state-of-the-art antimonide growth technologies. As an example of the complexities involved, the growth procedure of the third generation heterostructure is explained in the next paragraph.

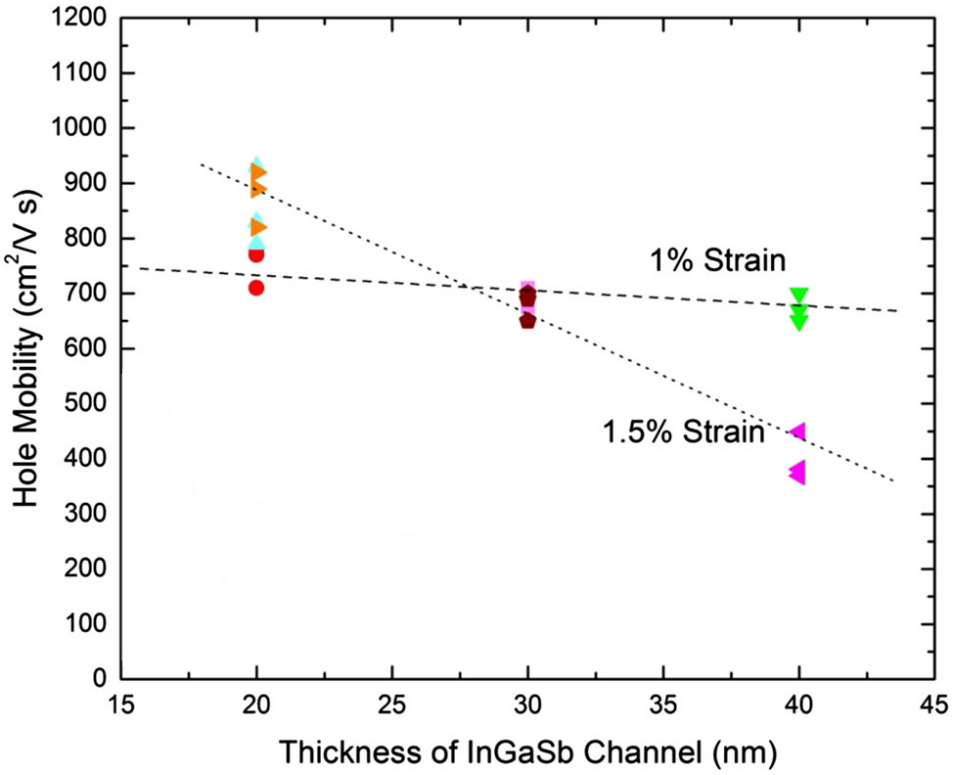


Fig. 3-6: Room temperature hole mobility in InGaSb quantum well with 1% and 1.5% compressive strain and various channel thickness (adopted from [101]).

The wafer was grown using Riber Compact 21E solid source MBE system at KIST. First, a semi-insulating undoped (100) GaAs substrate was introduced to the MBE chamber. Then, the surface oxide was removed thermally at the substrate temperature (T_s) of 630ºC under As dimer ambient, and a 200 nm thick GaAs regrowth layer was grown at T_s = 580 ºC. Subsequently, a 950 nm thick $Al_{0.93}Ga_{0.07}Sb$ metamorphic buffer layer was grown. The $Al_{0.9}Ga_{0.1}Sb$ buffer has high crystal quality, as checked by X-ray rocking curve. The buffer surface has root mean squared roughness (RMS) of 0.45nm, as measured by atomic force microscopy (AFM). The buffer also has a high resistivity of $4.3 \cdot 10^6$ Ω·cm. The $Al_{0.9}Ga_{0.1}Sb$ buffer was grown in optimized condition of T_s =

530 ºC and III/Sb_2 flux ratio of 10/12. From x-ray diffraction (XRD) measurement, the buffer layer is almost fully relaxed, setting the lattice constant of 6.13 Å.

Afterwards, a 30 nm buffer layer of $Al_{0.8}Ga_{0.2}Sb$ was grown to reduce the oxidation rate. Then, the growth temperature is reduced to 430 ºC to grow the channel layer, which consists of 23 nm of $In_{0.25}Ga_{0.75}Sb$. Finally, the top barrier layer of 2 nm $In_{0.20}Al_{0.8}Sb$ and the 30/5 nm p^+-$InAs_{0.85}Sb_{0.15}$/InAs alloy structure is grown through the Sb-for-As anion exchange process between the $In_{0.20}Al_{0.8}Sb$ and p-$InAs_{0.85}Sb_{0.15}$ interface [102], [103]. High p-type doping ($3{\cdot}10^{19}$ cm^{-3}) of p^+-$InAs_{0.85}Sb_{0.15}$ and p^+-InAs were formed using delta-doping method with Be dopant source. P-type delta doping layer ($5{\cdot}10^{11}$ cm^{-2}) in the $Al_{0.8}Ga_{0.2}Sb$ layer was formed at 5 nm below InGaSb channel. From XRD measurement, the $In_{0.25}Ga_{0.75}Sb$ channel is found to be -1.09% compressively stressed.

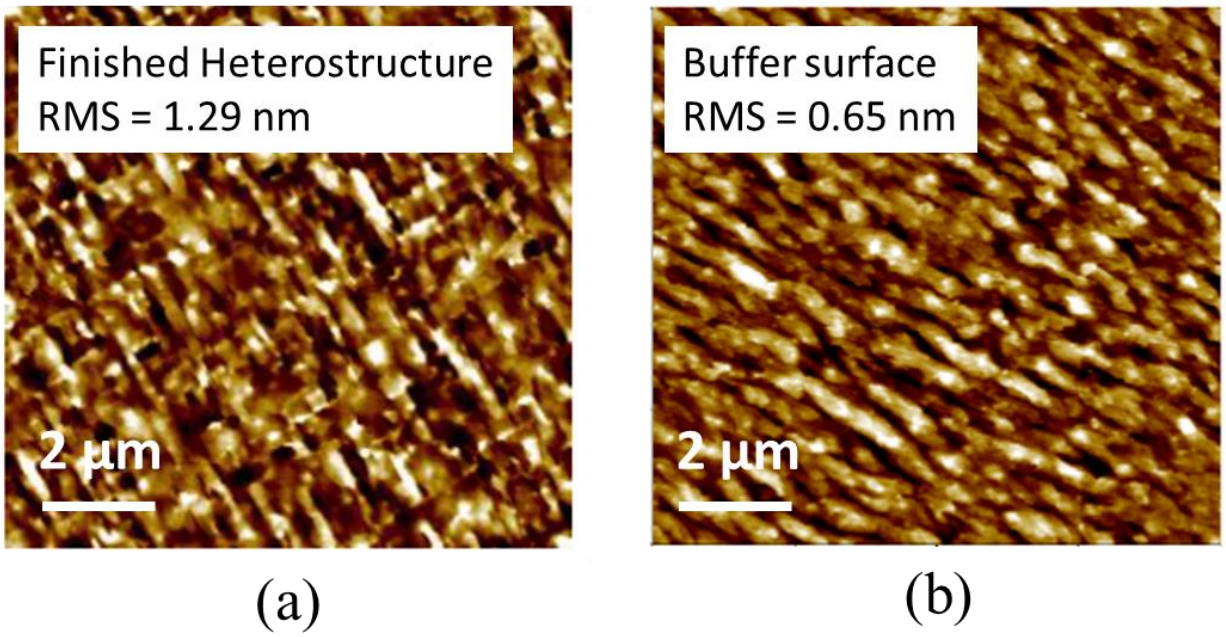


Fig. 3-7: AFM image of (a) as-grown MBE heterostructure and (b) of the surface of the graded buffer structure. Figure courtesy of KIST.

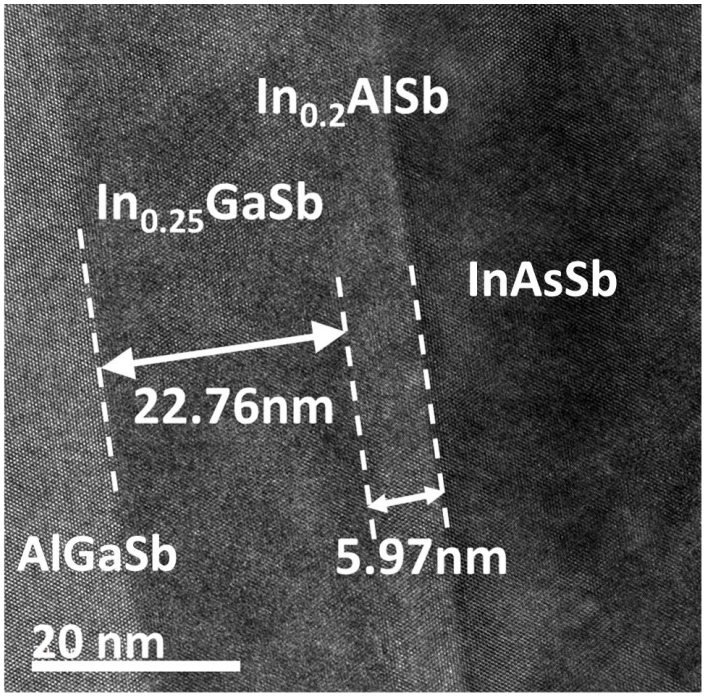


Fig. 3-8: TEM lattice image of the channel and spacer region of the grown heterostructure. Figure courtesy of KIST.

Carrier Concentration N_s	Hole mobility μ_h @ 300K	Hole mobility μ_h @ 77K	Channel resistivity ρ_{ch}	Buffer resistivity ρ_{buff}
$3.7 \cdot 10^{12}$ cm^{-2}	1175 cm^2/V·s	2521 cm^2/V·s	$3.6 \cdot 10^{-3}$ Ω·cm	$4.3 \cdot 10^{6}$ Ω·cm

Fig. 3-9 Summary of electrical properties of the InGaSb heterostructure in the third generation FinFETs.

Fig. 3-7 shows the AFM measurement of the finished heterostructure surface, and the $Al_{0.8}Ga_{0.2}Sb/Al_{0.97}Ga_{0.03}Sb$ graded buffer surface measured before the growth of the layers on top. The RMS roughness is 1.29 nm and 0.65 nm, respectively. Fig. 3-8 shows a cross-section TEM image of the heterostructure. The interfaces are atomically smooth. Fig. 3-9 summarizes the key electrical parameters of the grown wafer. The hole mobility in the channel is as high as

1175 $cm^2/V{\cdot}s$, measured by Hall measurement at room temperature. These characterizations confirm the high-quality growth of the antimonide heterostructure in this work.

3.3 Process Flow

This section will describe the individual modules of the process flow used to fabricate the self-aligned devices. The overall structure of the process flow of the three generations of InGaSb FinFETs is the same. Therefore, a generic description is given.

The device process starts with contact formation. A two-step e-beam lithography process is used to define the gate and mesa region. The first step defines the gate foot and spacers between the source/drain contacts which results in a transistor that is self-aligned. Then, fin RIE is carried out, which also serves the purpose of device isolation. This is followed by digital etch which is introduced in the third generation FinFETs. Afterwards, the gate dielectric and gate metal is deposited and patterned. Lastly, the device is finished with a via process and pad metallization. An overview of the process flow is shown in Fig. 3-10. The detailed recipes are summarized in the Appendix.

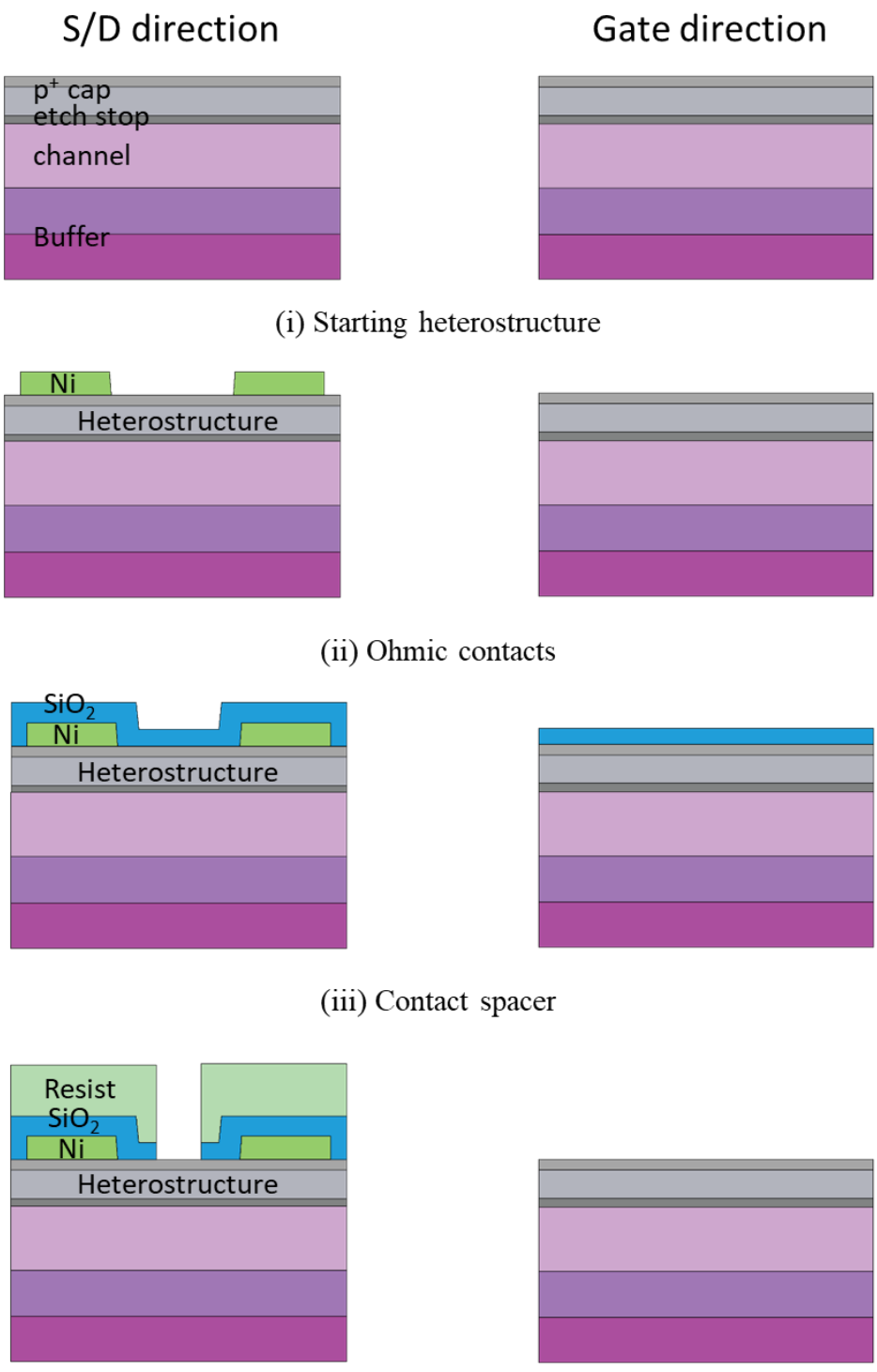


Fig. 3-10: InGaSb p-channel FinFET process flow. On the left column is the cross-section along the source-drain direction and on the right column is the cross-section normal to the channel direction.

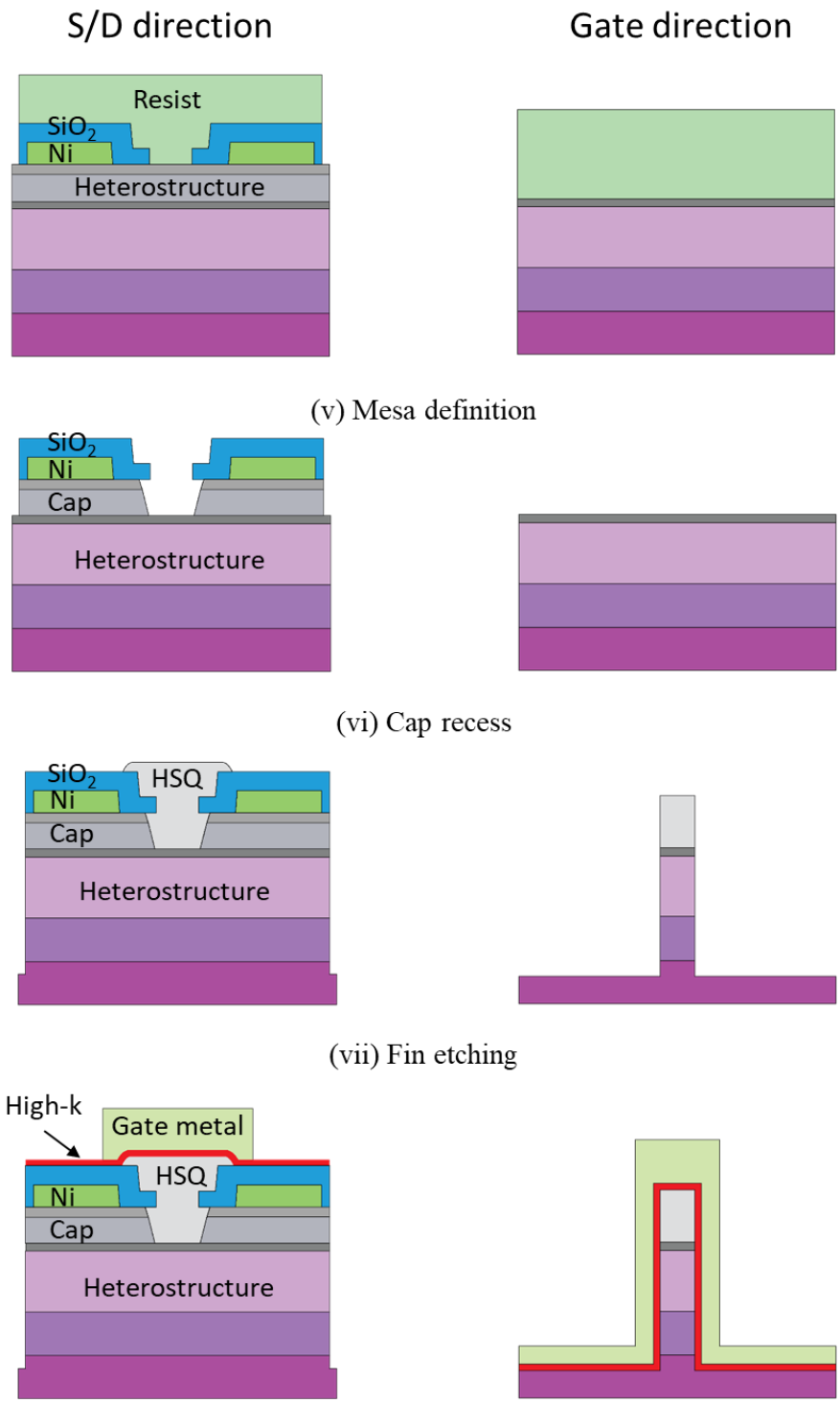
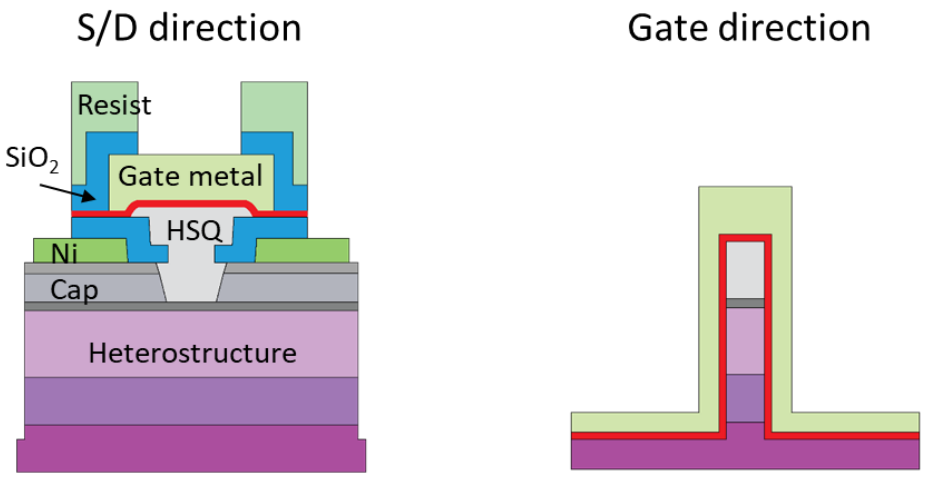
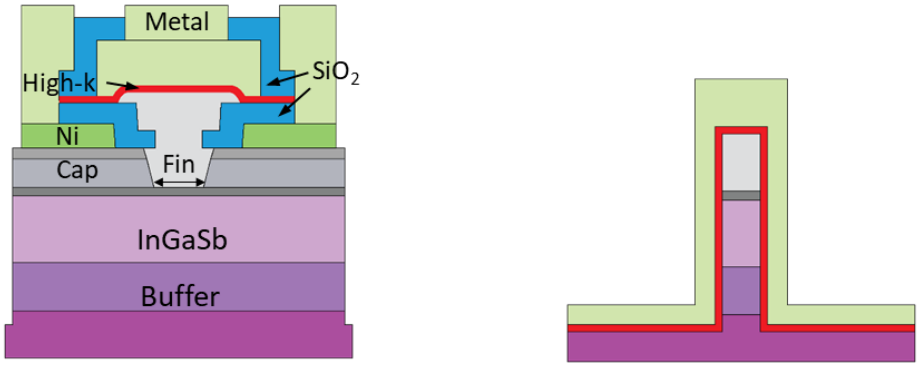


Fig. 3-10: (continued) InGaSb p-channel FinFET process flow. On the left column is the cross-section along the source-drain direction and on the right column is the cross-section normal to the channel direction.



(ix) Via patterning



(x) Metallization

Fig. 3-10: (continued) InGaSb p-channel FinFET process flow. On the left column is the cross-section along the source-drain direction and on the right column is the cross-section normal to the channel direction.

3.3.1 Ohmic contact

The process begins with cleaving the starting heterostructure pieces of 1x1 cm^2 in size, and then the cleaved pieces are cleaned in DI water, acetone, methanol, and IPA in this order, each for about

2 minutes, and are dried by N_2. Then, PMMA (polymethyl methacrylate) is spin coated at spin speed of 3500 rpm for 60s, resulting in a thickness of 500 nm, followed by a baking step of 3 minutes at 180°C. Then, PMMA is exposed using an Elionix ELS-F125 electron-beam lithography system. The electron gun has an accelerating bias of 125 keV, and the beam current is 10 nA, and the electron dose is between 1200 to 1400 $\mu C/cm^2$, depending on the pattern size. After the exposure, PMMA is developed in 1:3 MIBK (methyl isobutyl ketone) to IPA for 90 seconds. After development, the sample is cleaned with 1:3 HCl:H_2O for one minute to remove the native oxide on the p^+-InAs surface. Immediately after the HCl clean, the sample is loaded into vacuum chamber of a Temescal FC-2000 electron beam evaporator. Ni, Pt, and Au of thickness 15, 10, and 20 nm are deposited sequentially at a base pressure lower than $8 \cdot 10^{-7}$ Torr. The deposition rate is 1 Å/s for each metal. After the e-beam evaporation, the sample is soaked in acetone overnight to lift-off of the ohmic contact metal. The alignment marks for later processing step are also patterned and deposited during this step.

It is worthwhile to mention here that it is not an ideal contact-first process, because the Ni ohmic contact is lifted-off. For an ideal contact-first process, the contact metal is supposed to be deposited at the very beginning of the process on fresh semiconductor surface. Such is the case in some InGaAs MOSFETs in which Mo or W is used as the contact metal [31], since Mo and W contacts can be patterned by fluorine-based plasma. On the other hand, for Ni contacts, patterning via dry etching is not readily done because it usually involves chlorine-based plasma, which may also etch the underlying semiconductor. Some other work has developed possible lift-off free Ni contact methods, by alloying Ni and III-V at high temperature (> 300 °C) and then selectively etching away the unreacted Ni by HCl [46], [96]. However, according to our experience, completely clean

wet etching of Ni after RTA is not very successful. For simplicity of the FinFET fabrication process, this technique is not adopted in this thesis.

3.3.2 Gate-foot and mesa definition

After the ohmic contact, the gate-foot and mesa of the transistor need to be patterned to define the intrinsic region of the device. First, a 30 nm SiO_2 layer is deposited by chemical vapor deposition (CVD) at 300°C to serve as the spacer between the source/drain contacts and the gate. Note for a FinFET process, the total thickness of the ohmic contact and SiO_2 spacer cannot be too thick, because the HSQ hardmask for the fin RIE is at most about 100 nm thick. To obtain a CVD SiO_2

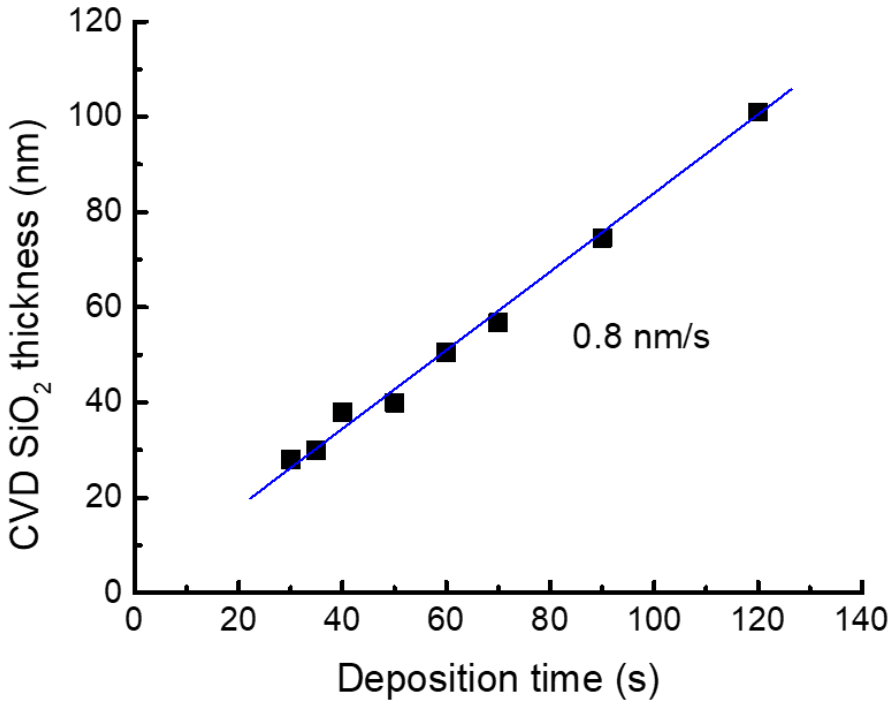


Fig. 3-11: CVD SiO_2 deposition thickness vs. deposition, with pre-deposition stabilization time of 3 minutes.

layer as thin as 30 nm, a long stabilization time is critical. Fig. 3-11 shows the deposition rate of the optimized high-frequency SiO_2 CVD with a pre-deposition stabilization time of 3 minutes.

After spacer deposition, the gate foot is to be patterned. To achieve a short gate length, an e-beam lithography process is needed. Common resist used for this step is PMMA or ZEP-520A diluted in anisol (ZEP is from Nippon Zeon Corp.). Both resists have been used for this step, though ZEP-520A is preferable for its superior dry etch resistance over PMMA. The e-beam resist is spin coated at 3000 rpm and baked at 180°C for 2 minutes. Then, the e-beam resist is exposed by Elionix with a beam current of 1 nA. To obtain narrow gate length, the dose of this exposure is critical, so a dose test is typically done before the actual exposure. After the e-beam lithography, the e-beam resist is developed in 1:3 MIBK:IPA for PMMA, or in xylene for ZEP, respectively.

The SiO_2 is etched with a CF_4 based chemistry, with addition of H_2 to the etching to improve selectivity. The etch rate of SiO_2 is calibrated by a Filmetrics reflectometer, and the etching profile of the SiO_2 is inspected by cleaved test samples using SEM. The ratio of CF_4 to H_2 is optimized to 32:3 to obtain good selectivity and anisotropy. A cross-section SEM image of the etched gate foot region with L_g = 30 nm is shown in Fig. 3-12. After the etching, the resist is removed in hot NMP for one hour, followed by ashing of 10 minutes at 1 kW to clean the resist residuals.

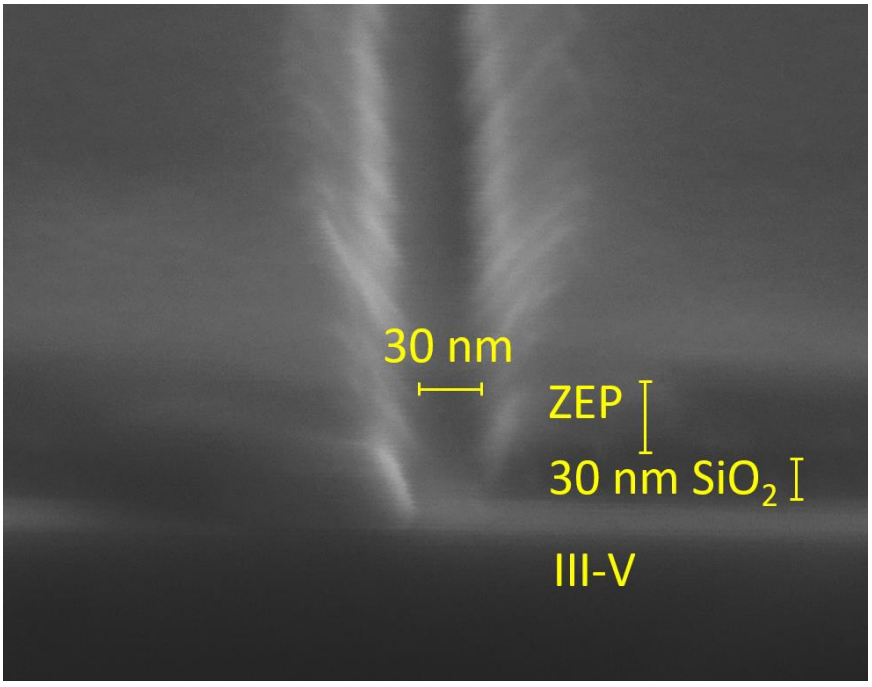


Fig. 3-12: Cross sectional SEM images for the ZEP/SiO_2 patterns for the gate-foot definition.

With the gate-foot patterned, the mesa area is defined next. Negative e-beam resist Ma-N 2403 from Micro Resist Technology is used as the mask. To improve the adhesion of Ma-N, the sample is pre-baked in oven at 120°C for 10 minutes. Then, Ma-N is spin coated at 3000 rpm for 1 minute and baked at 80°C for 1 minute. E-beam exposure of Ma-N is done with beam current of 10 nA and a dose of 400 $\mu C/cm^2$. Ma-N is developed in MICROPOSIT MF CD-26 developer for 1 minute. Then, the SiO_2 spacer is dry etched with the same CF_4/H_2 plasma as in the previous step. After the etching, Ma-N resist is removed in hot NMP. Extra ashing is needed if there is resist residuals remained.

Next step is the recess etching of p^+-InAs/InAsSb cap. With the $In_{0.2}Al_{0.8}Sb$ etch stopper, conventional wet recess using a mixture of citric acid and hydrogen peroxide is used with a volume

ratio of 10:1. The etch rate is approximately 0.7 nm/s. However, for devices with narrow gate foot, the etch rate can be lower and less uniform. Therefore, constant agitation is required during the etching, and sometimes ultrasonic can be used to assist the etching. The cap is slightly overetched to ensure a clean surface, which results in an undercut of 40 nm and above.

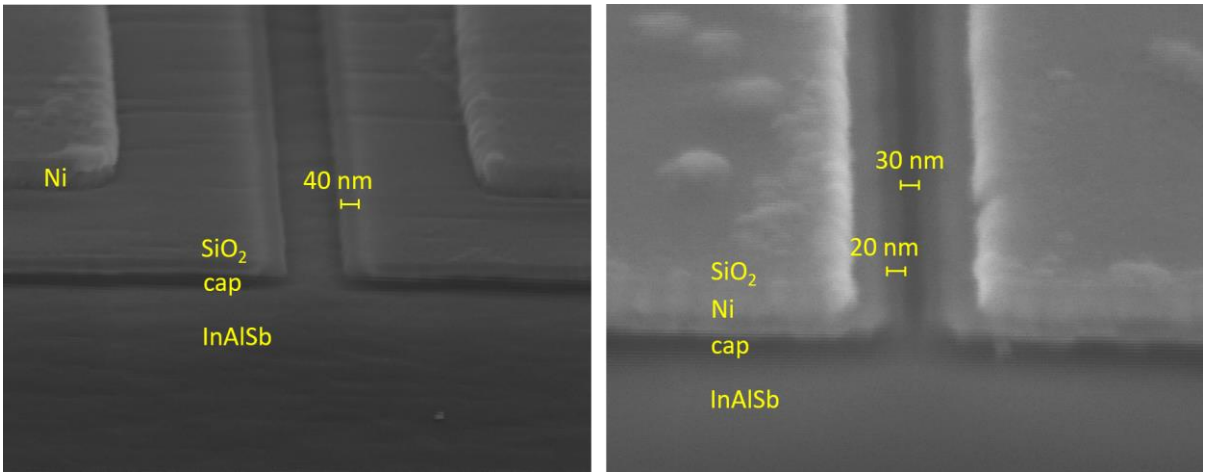


Fig. 3-13: Tilted SEM image of the gate foot region with the p^+-cap recessed with (a) 10:1 citric acid:H_2O_2, and (right) a two-step dry-wet recess process.

In order to reduce the undercut of the wet recess and hence the series resistance, a two-step recess process is employed in the third-generation FinFETs. The first 30 nm of the 35 nm p^+ cap are removed by timed Cl_2/N_2 dry etch. Then the rest is selectively etched in 10:1 citric acid:H_2O_2. This dry-wet recess yields an undercut of the p^+ cap that is less than 20 nm, while maintaining a smooth top surface. In addition, the process has been optimized so that the distance from the edge of Ni ohmic contacts to the edge of the p^+-cap has been reduced from 350 nm in the first generation FinFETs to less than 100 nm in the second and third generation devices, as shown in Fig. 3-13.

3.3.3 Fin and Gate Stack

After the gate recess, the fins need to be patterned. A detailed description of the fin patterning and RIE technology has been given in Section 2.2. In the first generation FinFETs, the fin sidewall is immediately covered by gate dielectric after the fin etch. In the second generation, the fin sidewall is treated with 10% HCl for 30 seconds in IPA right after the fin RIE. In the third generation, the fin sidewall is first treated with 10% HCl, followed by one cycle of digital etch in O_2 and HCl:IPA. The antimonide-compatible digital etch process was discussed in Section 2.3.

After fin RIE and surface treatments, the samples are immediately brought into Cambridge Nanotech atomic layer deposition (ALD) chamber for high-k dielectric deposition. For the three generations of FinFETs, the gate dielectric is 4nm Al_2O_3 at 250°C, 4 nm HfO_2 at 200°C, and 3.5 nm of Al_2O_3 at 175°C, respectively. The film deposition rate is confirmed to be 1.0 Å/cycle by TEM.

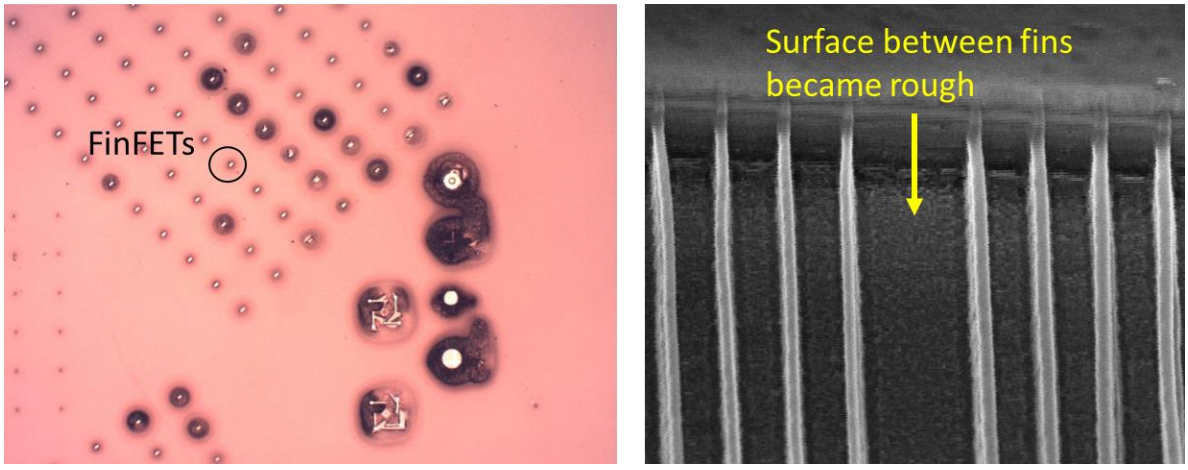


Fig. 3-14: (Left) optical microscopy and (right) SEM images of fin-etched InGaSb FinFET sample surface after 10 minutes of exposure in air.

It is worth noting that the samples should not be exposed to air for too long after the fin etching, because of the highly reactive AlGaSb buffer is unstable in air particularly for highly strained heterostructures when using buffers with high Al composition. Fig. 3-14 shows optical microscopy and SEM images of etched samples exposed in air for 10 minutes. The sample surface becomes brownish and less shiny, and the SEM images reveal that the surface becomes rough after long exposure in air. Therefore, the fin sidewalls need to be passivated quickly. This poses challenges in the antimonide-based FinFET technology and requires more studies in the future.

After gate oxide ALD, the samples are covered with sputtered gate metals, such as Mo or Al, in vacuum using an AJA International ATC-180 sputter deposition system. The power of sputtering is calibrated to maintain a deposition rate of 0.8-1.0 Å/s.

After the formation of the fin and gate stack, the gate is to be patterned. PMMA is spin coated and exposed by e-beam lithography and developed in 1:3 MIBK:IPA. Then, 10 nm and 250 nm Ti and Au is deposited by e-beam evaporation. The Ti/Au stack needs to be thick enough to cover the fin RIE height. The Ti/Au bilayer is lifted off in acetone and serves as the hardmask for gate etch. The Mo gate metal is etched by SF_6/O_2 plasma. A Plasmaquest electron cyclotron resonance (ECR) etcher is used. In the case of Al gate metal, Al is etched by BCl_3/Ar plasma, using a SAMCO 200iP inductively coupled plasma etcher.

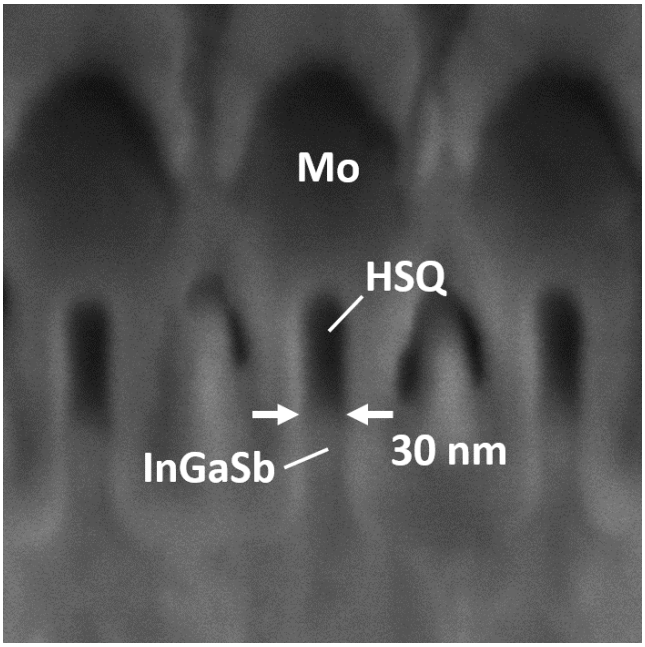


Fig. 3-15: FIB cross-section of first-generation InGaSb FinFET with W_f = 30nm.

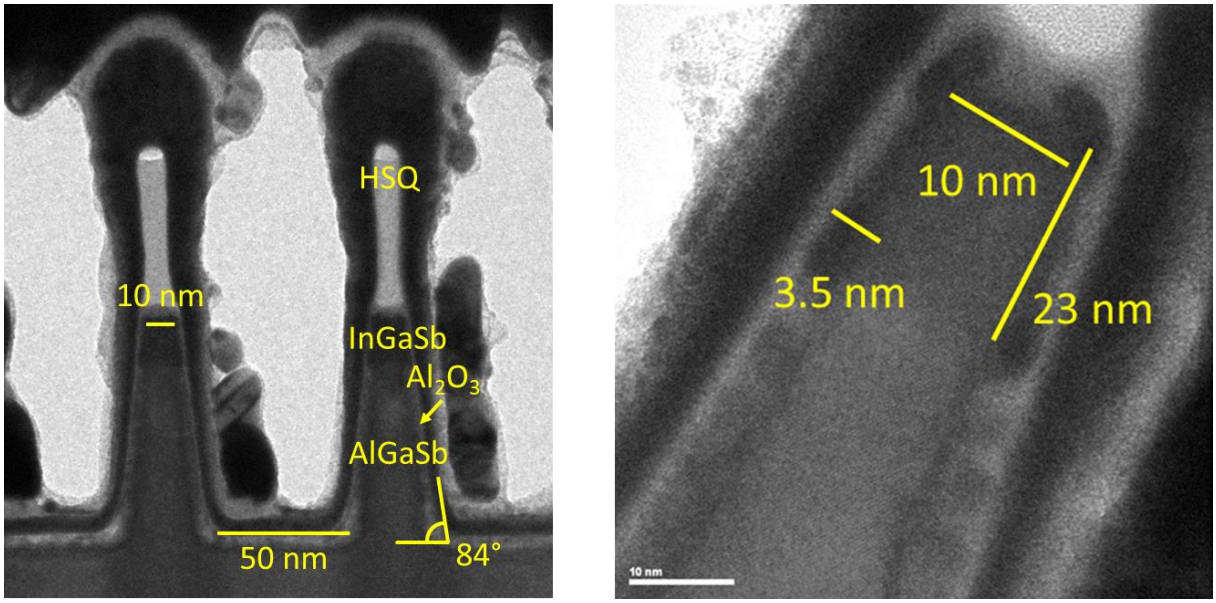


Fig. 3-16: HR-TEM images of third-generation InGaSb FinFET with fin width of 10 nm, fin aspect ratio of 2.3, and 3.5 nm Al_2O_3 gate dielectric.

At this point, the intrinsic portion of the FinFET is finished. Note that the HSQ mask for the fin RIE is left to the end of the process, so the resulting device geometry is a double-gate FinFET,

with the gate modulating only the two sidewalls of the fins. Fig. 3-15 shows a focused-ion-beam (FIB) cross-sectional image of a first generation FinFET with minimum W_f = 30 nm. Fig. 3-16 shows a high-resolution transmission electron microscopy (HR-TEM) image of a third generation FinFET, demonstrating a fin width of 10 nm, fin height of 115 nm and a channel aspect ratio (AR = H_c/W_f) of 2.3. The fin sidewall angle is 84° and the fin pitch is 100 nm.

3.3.5 Back-end Process

The final process for device fabrication is the back-end process. It comprises of interconnect via process and pad metallization. After gate-metal etching, the sample is covered by a 30 nm thick tetraethylorthosilicate (TEOS) spacer. TEOS is deposited conformally by plasma-enhanced chemical vapor deposition (PECVD) using an Oxford PlasmaPro-100 system at substrate temperature of 300°C. Then, the via layer is patterned by e-beam lithography. Undiluted ZEP-520A e-beam resist is spin coated, baked at 180°C for 2 minutes, exposed by Elionix with a dose of 560 $\mu C/cm^2$, and developed in xylene for 1 minute.

The via is etched in three steps. First, the TEOS layer is etched by CF_4 plasma in the Oxford PlasmaPro-100 system. Then, the gate high-k dielectric needs to be removed. A dry etching process is used to etch the Al_2O_3 or HfO_2 [104], [105]. Timed BCl_3/Ar ICP plasma etching is used to remove the high-k dielectrics with slight over-etching. Lastly, the SiO_2 spacer on the ohmic contacts that was deposited in the beginning of Section 3.3.2 is etched by the same CF_4 plasma in the first step. Fig. 3-17 shows the patterned contact via and the close-up image of the via edge.

After the via etch, the ZEP e-beam resist is removed by hot NMP for 2 hours, followed by ashing of 15 minutes.

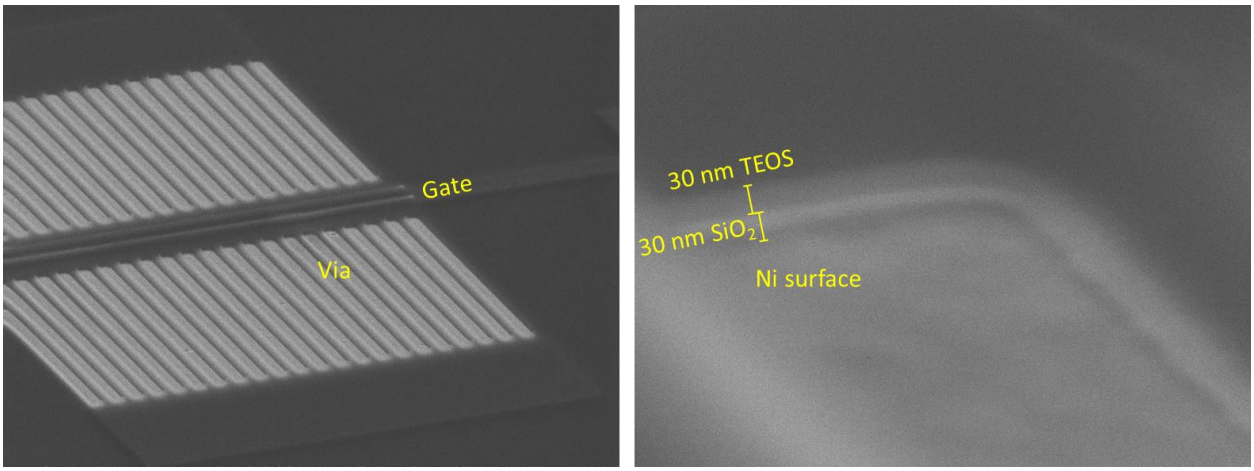


Fig. 3-17: (Left) SEM image of the FinFET after the via etching. (Right) zoomed-in SEM image showing the edge of the contact via.

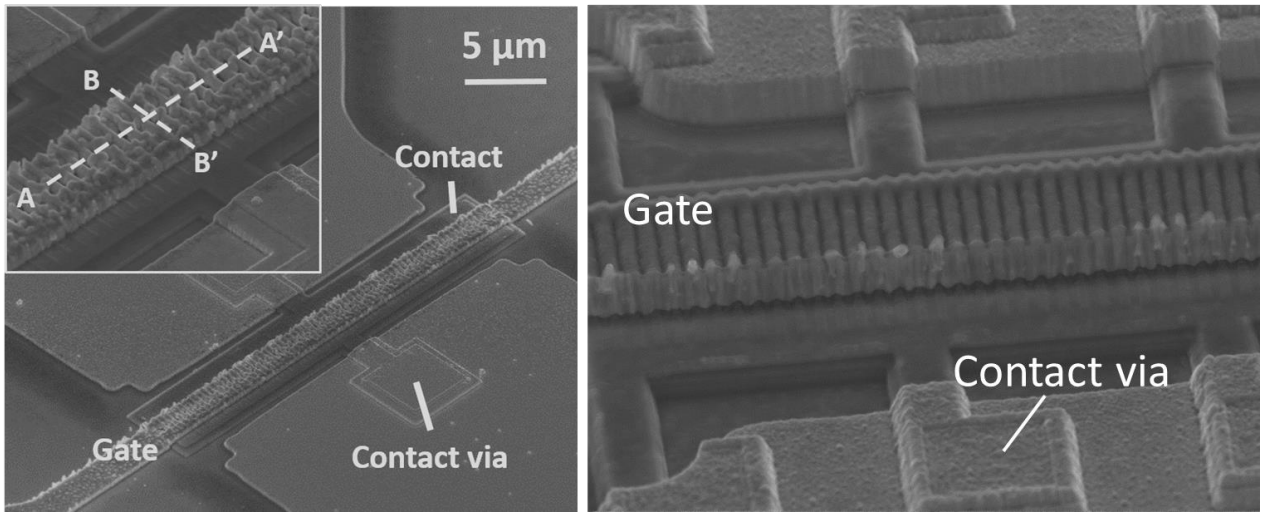


Fig. 3-18: SEM images of the finished devices in the (left) first and (right) third generations of InGaSb p-channel FinFETs.

Lastly, the source, drain, and gate contact pads are patterned by photolithograph. The samples are coated with hexamethyldisilazane (HMDS), followed by spin coating of AZ-5214 image reversal

photoresist and baking of the resist at 80°C for 5 minutes. The photoresist is exposed by a Heidelberg MLA-150 maskless aligner, using a 375 nm laser exposure with a dose of 16 mJ. Then, the image-reversal baking is done at 105°C for 65 seconds, followed by a flood exposure for 35 seconds. The resist is developed in AZ 422 MIF developer for 90 seconds. The Ti/Au pads are then deposited by e-beam evaporation and lifted-off in acetone. The FinFET fabrication is completed. Fig. 3-18 shows SEM images of finished devices in the first and third generations of InGaSb p-channel FinFETs.

3.4 Chapter Summary

In this chapter, the key technological components of InGaSb p-channel FinFETs developed in the previous chapter have been integrated into a process flow that is used to demonstrate the first InGaSb FinFET. Heterostructure design and epitaxy growth process have been discussed. Three generations of InGaSb FinFET process have been developed. A contact-first, gate-last self-aligned process scheme is developed. The FinFET process is CMOS-compatible, Au-free in the front-end, and has very low thermal budget. We have demonstrated the first InGaSb p-channel FinFET with fin widths down to 30 nm and gate lengths down to 100 nm. With the optimizations of the process, we have demonstrated second-generation FinFETs with fin widths down to 18 nm and gate lengths down to 20 nm. With the introduction of antimonide-based digital etch, we have demonstrated third-generation FinFETs with fin widths scaled to 10 nm.

The device fabrication technology developed in this chapter forms the fundamentals of antimonide-based FinFETs process and serves as the basis for the rest of the thesis. In Chapter 4, the electrical characteristics of the fabricated transistors will be shown and analyzed.

CHAPTER 4. InGaSb FinFETs Performance

4.1 Introduction

Chapter 2 and 3 have described the key technologies developed to fabricate InGaSb p-channel FinFETs. In this chapter, the electrical performance of the world's first InGaSb p-channel FinFET is presented, along with the two successive device generations. Table 4-1 summarizes the key device features and differences of the three generations of FinFETs (G1, G2, and G3 for short). The smallest fin width scales from 30 nm in G1, to 18nm in G2, and to 10 nm in G3. The minimum gate length scales from 100 nm in G1 to 20 nm in both G2 and G3. In G1, there is no fin sidewall treatment, and the fin is covered with 4 nm Al_2O_3 immediately after the fin RIE. In G2, the fin sidewall is treated by HCl:IPA before ALD. In G3, both HCl:IPA treatment and antimonide-based digital etch are applied to the devices.

For each generation of InGaSb FinFETs, the DC characteristics are presented, followed by performance benchmarking against other published antimonide-based III-V p-channel MOSFETs. The device characteristics are also analyzed from the perspective of fin width and gate length scaling, strain, ON-state resistance, and OFF-state leakage current. The benefits and limitations

of the post-RIE surface treatment methods are also discussed, revealing the importance of interface quality to the device performance.

	Generation 1	Generation 2	Generation 3
Fin width W_f	30 – 200 nm	18 – 200 nm	10 – 100 nm
Gate length L_g	100 nm – 1 μm	20 – 250 nm	20 nm – 1 μm
Channel	10 nm $In_{0.27}Ga_{0.73}Sb$	7.5 nm $In_{0.4}Ga_{0.6}Sb$	25 nm $In_{0.25}Ga_{0.75}Sb$
δ-doping	$2 \cdot 10^{18}$ cm^{-3}	$2 \cdot 10^{18}$ cm^{-3}	$1 \cdot 10^{18}$ cm^{-3}
Cap design	Single	Composite	Composite
Cap doping	$1 \cdot 10^{19}$ cm^{-3}	$1 \cdot 10^{19}$ cm^{-3}	$3 \cdot 10^{19}$ cm^{-3}
Gate dielectric	4 nm Al_2O_3	4 nm HfO_2	3.5 nm Al_2O_3
Gate metal	Mo	Mo	Al
Post-RIE Treatment	None	HCl:IPA	HCl:IPA + DE

Table 4-1: Summary of device features of three generations of InGaSb p-channel FinFETs.

4.2 Device Electrical Characterization

The following sections present the DC characteristics of each of the three generations of FinFETs in terms of their output and transfer characteristics. Fig. 4-1 shows the device cross-sectional schematics representing the general device structure, and Fig. 4-2 shows the SEM images of a typical finished FinFET. Since the fabricated FinFETs all have a double-gate geometry, unless otherwise specified, drain current (I_D), transconductance (g_m), and ON-resistance are normalized by the total gate periphery, $2N_f H_{ch}$, where N_f is the number of fins per transistor, and H_{ch} is the thickness of the InGaSb channel.

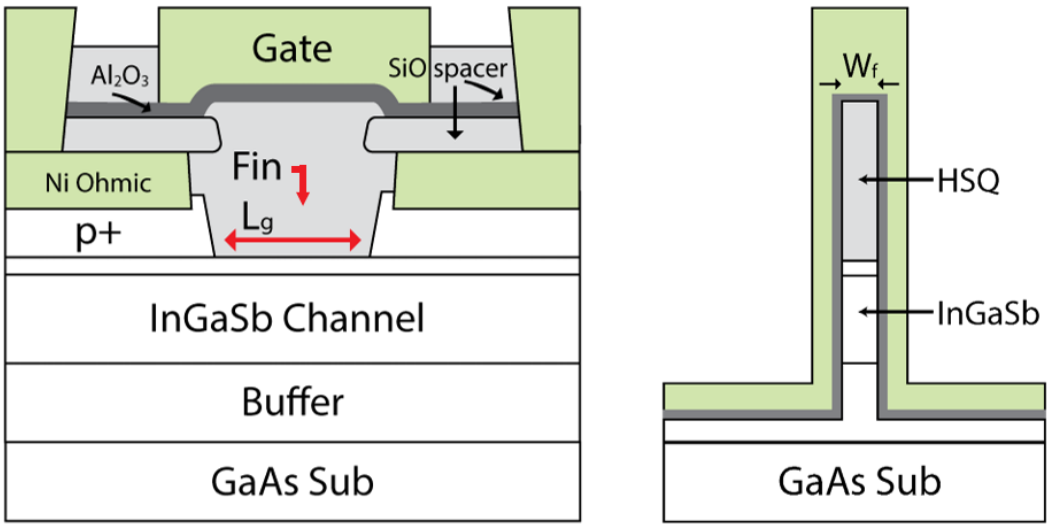


Fig. 4-1: Schematic cross sections of InGaSb p-channel FinFET: (left) along the source-drain direction and (right) across the fin and the channel direction.

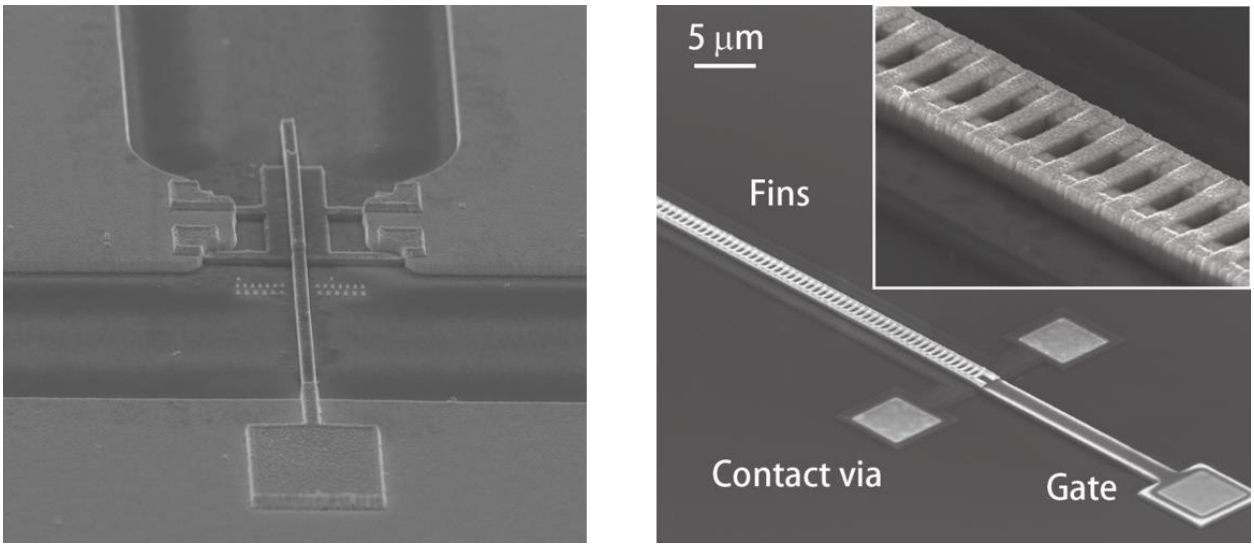


Fig. 4-2: (Left) SEM top-views of a finished InGaSb p-channel FinFET. (Right) Close-up views of the intrinsic area of a FinFET.

4.2.1 First Generation

In the G1 FinFETs, the smallest device that has been fabricated has W_f = 30 nm and L_g = 100 nm. The gate dielectric is 4 nm Al_2O_3 (EOT = 1.8 nm). Fig. 4-3 and 4-4 show typical output and subthreshold characteristics of the smallest FinFETs in G1. Fig. 4-5 and 4-6 show output and transfer characteristics of devices with wider fins and longer channel (W_f = 100 nm, L_g = 1 μm). The gate leakage current is at least two orders of magnitude below the drain current throughout the measurement range, so it is not shown in the figures. The 30 nm fin-width devices exhibit better saturation behavior for all channel lengths. However, as presented in Fig. 4-3 to 4-6, all devices suffer from poor turn-off behavior, and we will study this issue in details in Section 4.3.

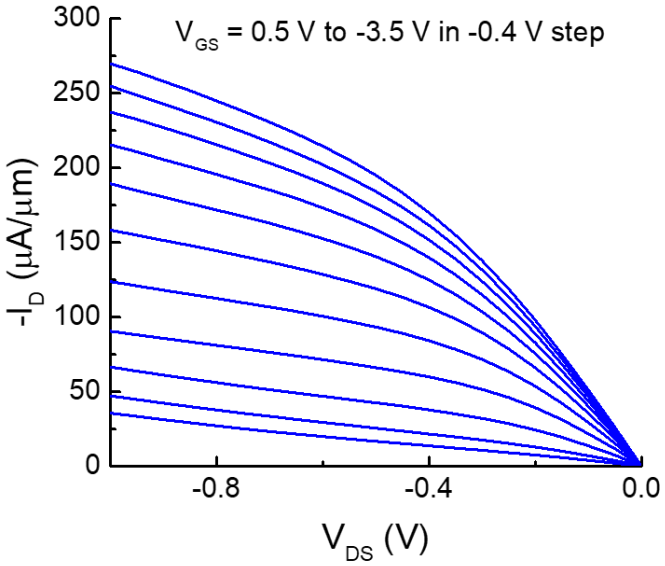


Fig. 4-3: Output characteristics of G1 InGaSb FinFET with W_f = 30 nm and L_g = 100 nm.

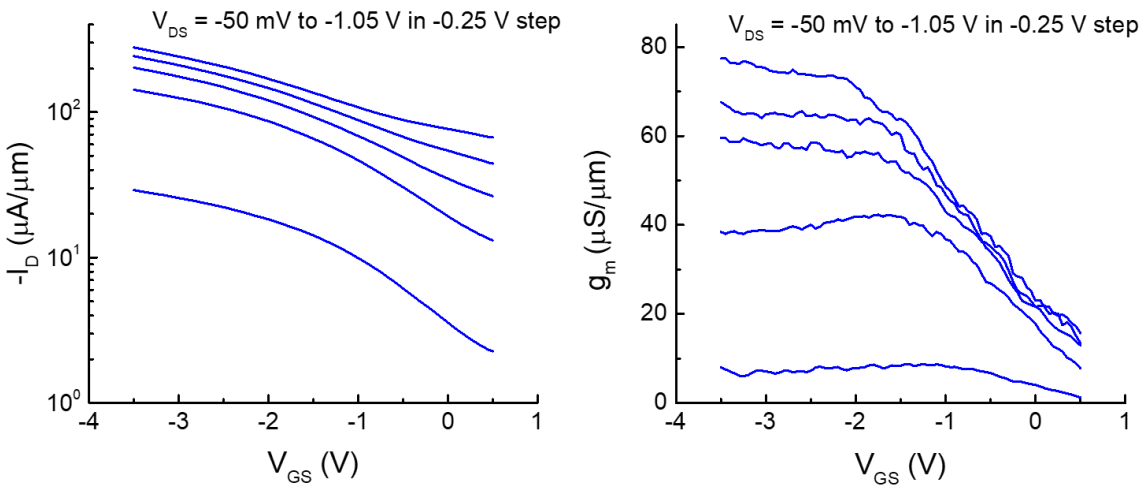


Fig. 4-4: (Left) subthreshold characteristics and (right) g_m characteristics of FinFET with W_f = 30 nm and L_g = 100 nm, at V_{DS} = -50 mV to -1.05 V and V_{GS} = -4 to 1 V.

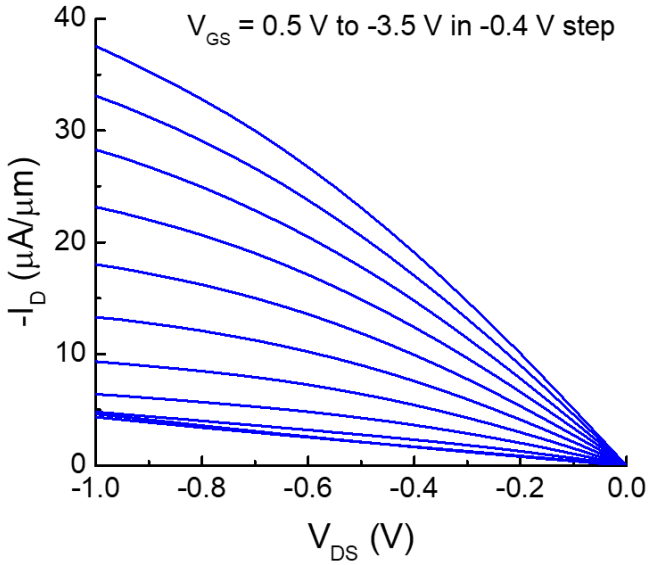


Fig. 4-5: Output characteristics of G1 InGaSb FinFET with W_f = 100 nm and L_g = 1 μm.

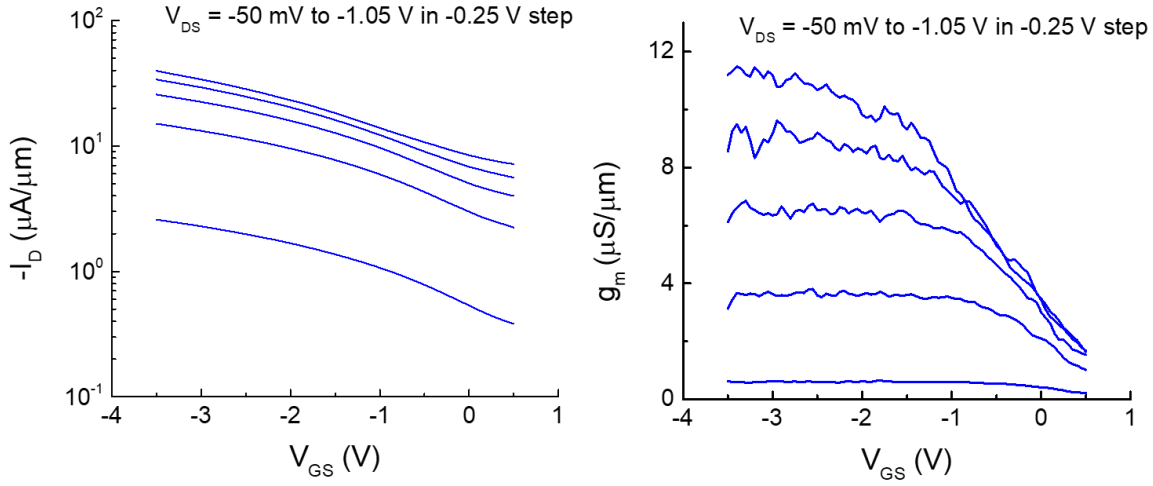


Fig. 4-6: (Left) subthreshold characteristics and (right) g_m characteristics of FinFET with W_f = 100 nm and L_g = 1 μm, at V_{DS} = -50 mV to -1.05 V and V_{GS} = -4 to 1 V.

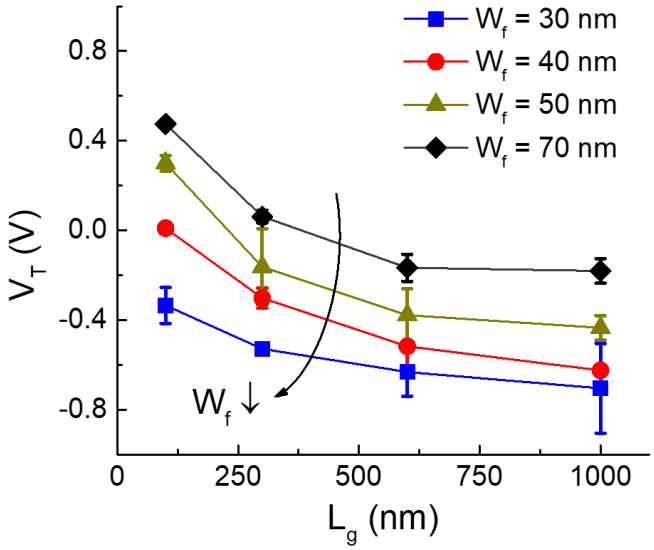


Fig. 4-7: Linear V_T rolloff with gate length of G1 InGaSb FinFETs with various fin widths.

The scaling behavior of key figures of merit of G1 FinFETs is studied as a function of fin width and gate length. Fig. 4-7 shows the threshold voltage (V_T), extracted using the extrapolation method [106] in the linear regime. As W_f is reduced to 30 nm, the V_T rolloff phenomenon is mitigated, showing the beneficial effect of W_f scaling on short-channel effects. The devices also

show a well-behaved dependence of g_m on L_g and W_f, for devices with the same fin orientation ([01$\bar{1}$]), as shown in Fig. 4-8, that the maximum g_m is increased with wider fins or shorter gate length. The highest g_m of 122 μS/μm is obtained in wide-fin and short-channel devices, with W_f = 100 nm and L_g = 100 nm, at V_{DS} = -1.05 V. g_m degrades as W_f decreases.

Fig. 4-9 shows the peak g_m at various L_g of G1 FinFETs with two fin orientations [001] and [01$\bar{1}$], together with published InGaSb p-channel MOSFETs which are all planar devices [43], [44], [46]–[48]. As the first FinFETs demonstrated in the antimonide system, the results that have been obtained are very encouraging, as the maximum g_m approaches the best InGaSb planar MOSFETs.

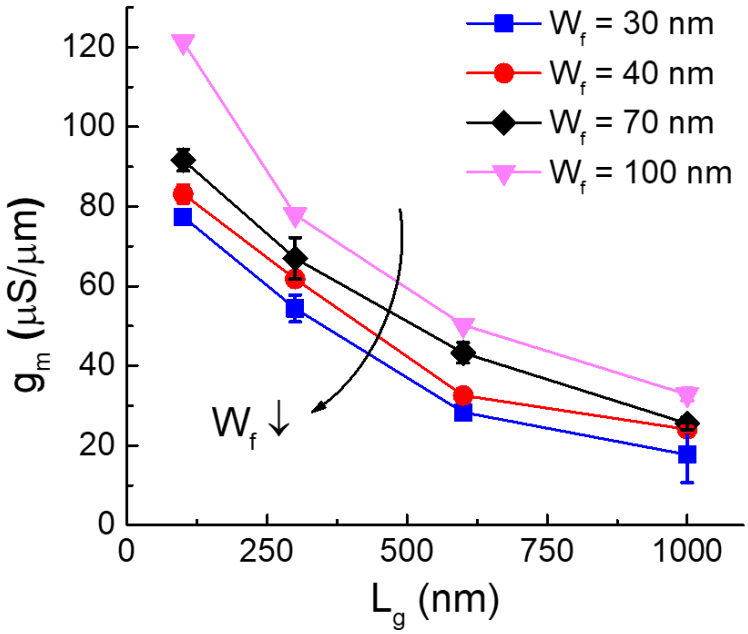


Fig. 4-8: Maximum g_m as a function of L_g for G1 devices with different fin widths (along the same fin orientation [01$\bar{1}$]).

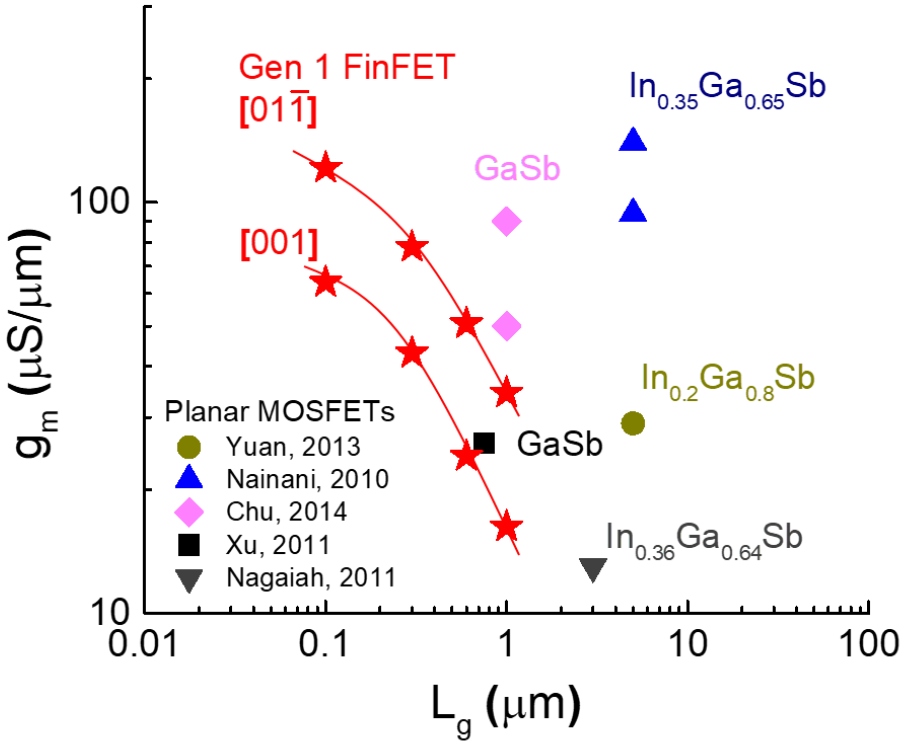


Fig. 4-9: Maximum g_m vs. L_g for G1 InGaSb p-channel FinFETs with fin orientations of [001] and [01$\bar{1}$], benchmarked with antimonide planar MOSFETs.

4.2.2 Second Generation

In the G2 FinFETs, the smallest device has W_f = 18 nm and L_g = 20 nm. Fig. 4-10 and 4-11 show the output and subthreshold characteristics of the smallest devices, respectively, on the [01$\bar{1}$] direction. Compared to the G1 devices, the device exhibits better saturation characteristics and a much-improved minimum linear subthreshold slope S_{linear} = 370 mV/dec at V_{DS} = -50 mV (though it is still insufficient). Fig. 4-12 shows the g_m characteristics of the narrow-fin short-channel (W_f = 18 nm and L_g = 20 nm) device and the wide-fin device with longer channel (W_f = 200 nm, L_g = 250 nm). The narrow-fin devices exhibit a maximum g_m of 186 μS/μm at V_{DS} = 0.5 V, which is

52% higher than the best devices in G1. A maximum g_m of 338 μS/μm is obtained in wide-fin devices. This is a record among GaSb or InGaSb p-channel MOSFETs, and nearly triples the previous record, as shown in the benchmark in Fig. 4-13. These results suggest a substantial improvement of the InGaSb FinFETs in the second generation.

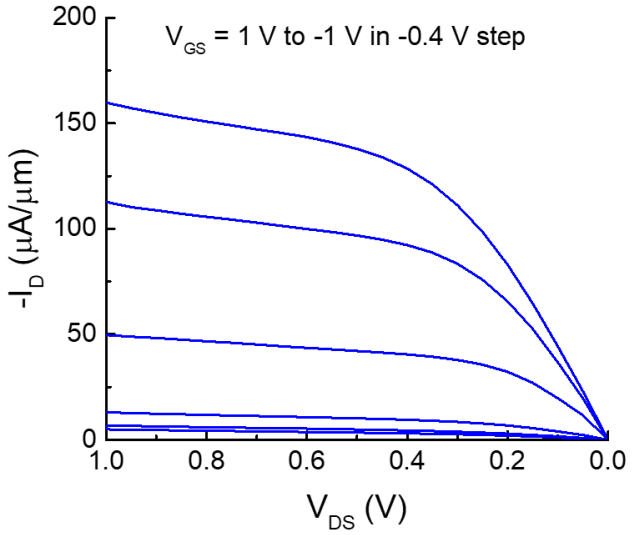


Fig. 4-10: Output characteristics of G2 InGaSb FinFET with W_f = 18 nm and L_g = 20 nm.

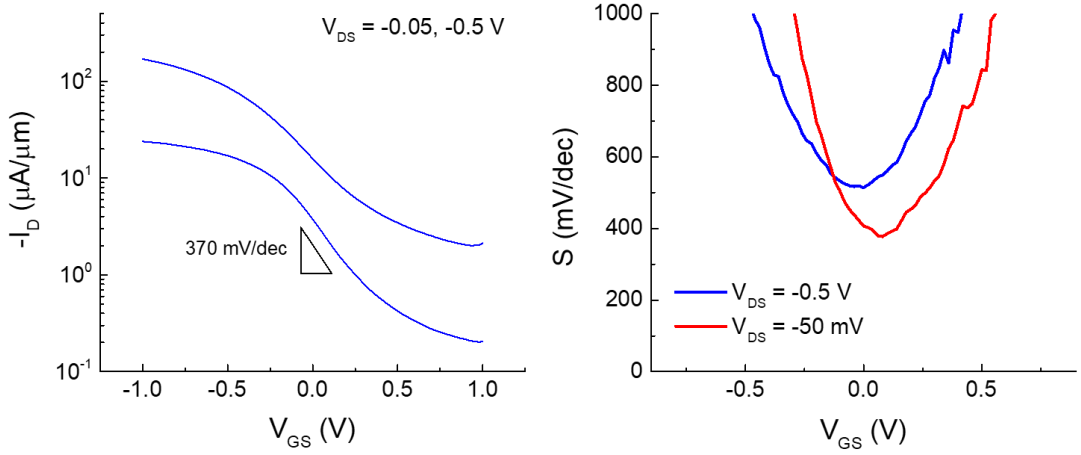


Fig. 4-11: (Left) subthreshold characteristics of G2 InGaSb FinFET with W_f = 18 nm and L_g = 20 nm. (Right) subthreshold slope of the device.

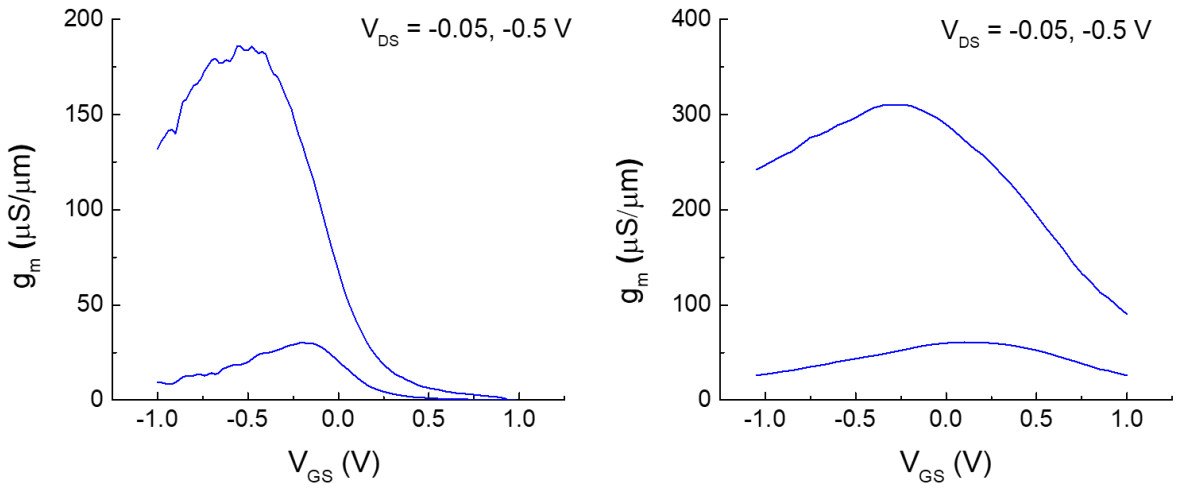


Fig. 4-12: g_m characteristics of G2 FinFETs with (left) W_f = 18 nm and L_g = 20 nm, and (right) W_f = 200 nm and L_g = 250 nm.

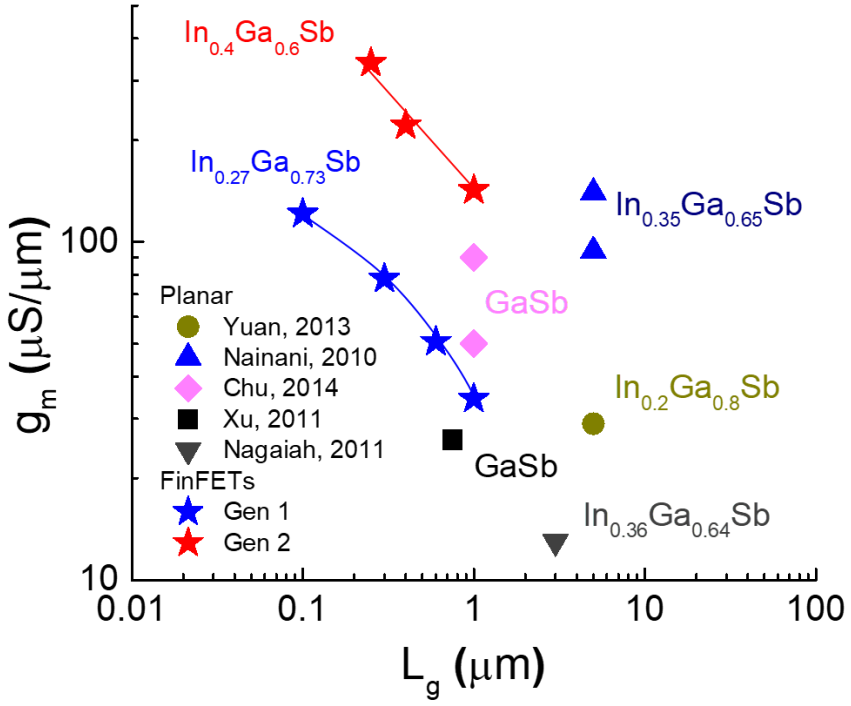


Fig. 4-13: Maximum $g_{m\ vs.}$ L_g for G1 and G2 InGaSb p-channel FinFETs with different channel composition, benchmarked with antimonide planar MOSFETs.

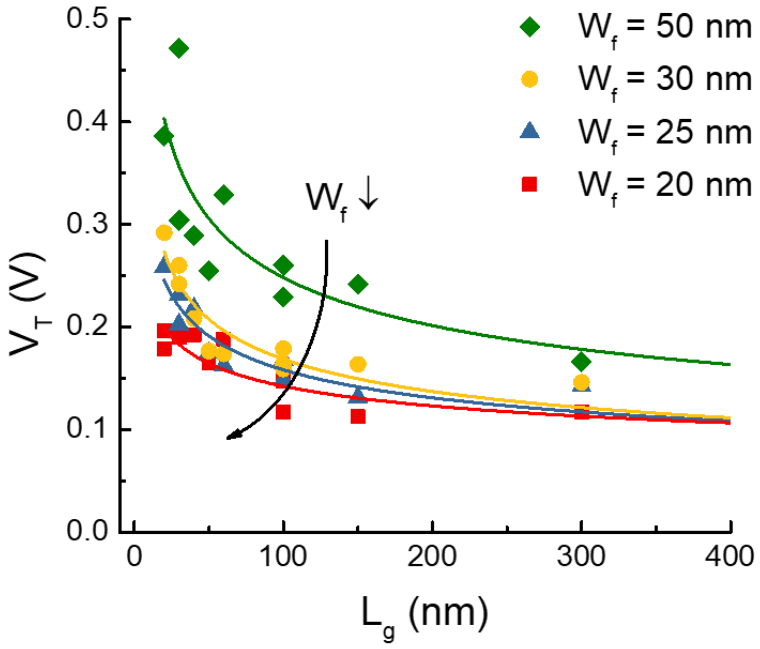


Fig. 4-14: Linear V_T rolloff with gate length of G2 InGaSb FinFETs with various fin widths ranging from 50 to 20 nm.

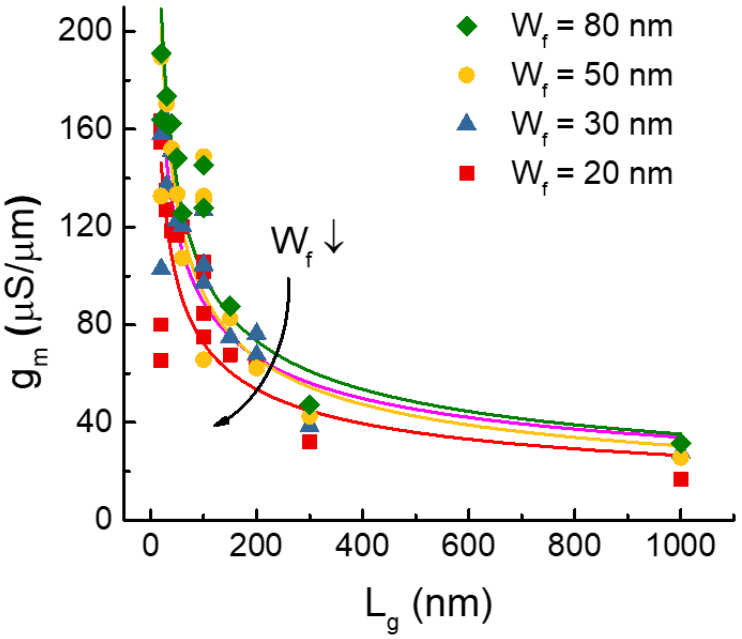


Fig. 4-15: Maximum g_m as a function of L_g for G2 devices with fin widths ranging from 80 to 20 nm. g_m is measured at V_{DS} = -0.5 V.

The scaling behavior of V_T and g_m of G2 devices as a function of fin width and gate length is shown in Fig. 4-14 and 4-15. V_T rolloff is reduced significantly at reduced W_f down to 20 nm. The scaling of g_m follows the classical trend that g_m is enhanced at shorter channel while it is reduced at narrower fin width. All shown devices in G2 are along the $[01\bar{1}]$ direction, as well as in G3, as being discussed next.

4.2.3 Third Generation

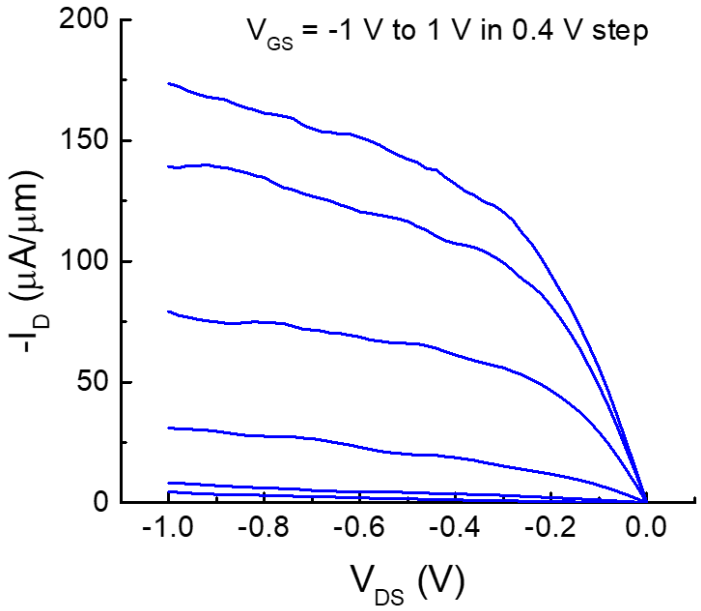


Fig. 4-16: Smoothed output characteristics of G3 InGaSb single-fin FinFET with W_f = 10 nm and L_g = 20 nm.

The G3 FinFETs have the most aggressively scaled devices. As the fabrication technology of InGaSb FinFETs become more mature over the three generations, device yield has also been improved so that working single-fin InGaSb devices are demonstrated successfully for the first time in G3. Fig. 4-16 and 4-17 shows the smoothed output and subthreshold characteristics of the

InGaSb single-fin FinFET with W_f = 10 nm and L_g = 20 nm. A peak smoothed g_m = 160 µS/µm is obtained at V_{DS} = -0.5 V, and a minimum linear S = 260 mV/dec is obtained at V_{DS} = -50 mV. The IV characteristics of single fin devices shows generally higher fluctuations. This low-frequency noise is probably due to the flicker noise from trapping and detrapping events near the channel/oxide interface and is prominent due to the small size of the single-fin device. Similar observations have been previously made in InAs and InGaAs nanowire MOSFETs [107], [108]. The nature and origin of the noise in InGaSb devices have not been studied and it is worth for future detailed investigations.

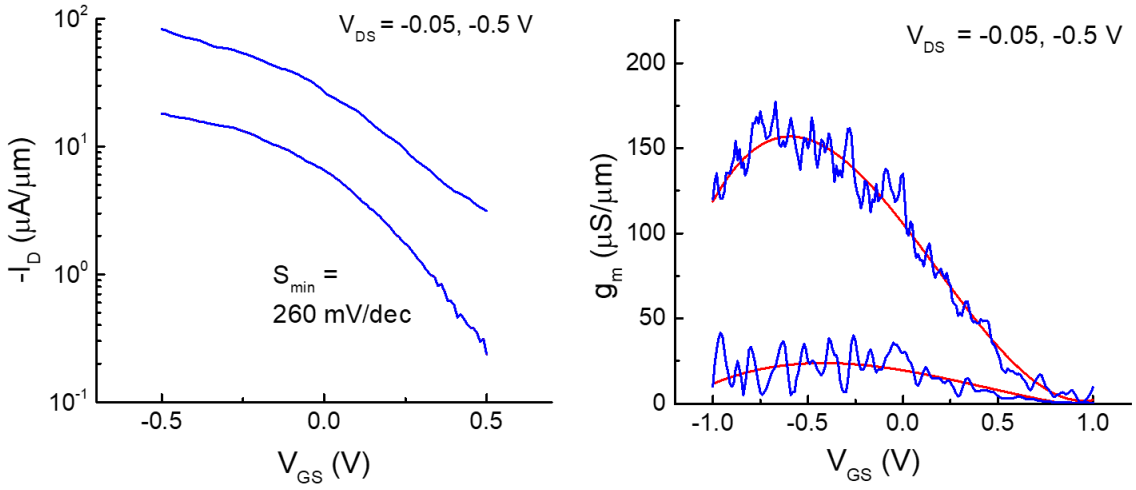


Fig. 4-17: (Left) subthreshold characteristics of G3 InGaSb single-fin FinFET with W_f = 10 nm and L_g = 20 nm. (Right) g_m characteristics of the device, with peak g_m = 160 µS/µm at V_{DS} = -0.5 V.

The output and subthreshold characteristics of a 100 fin array long channel device with W_f = 10 nm and L_g = 1 µm are displayed in Fig. 4-18 and 4-19, showing good saturation behavior. The long channel devices exhibit linear S = 290 mV/dec and peak g_m = 4 µS/µm at V_{DS} = -0.5 V.

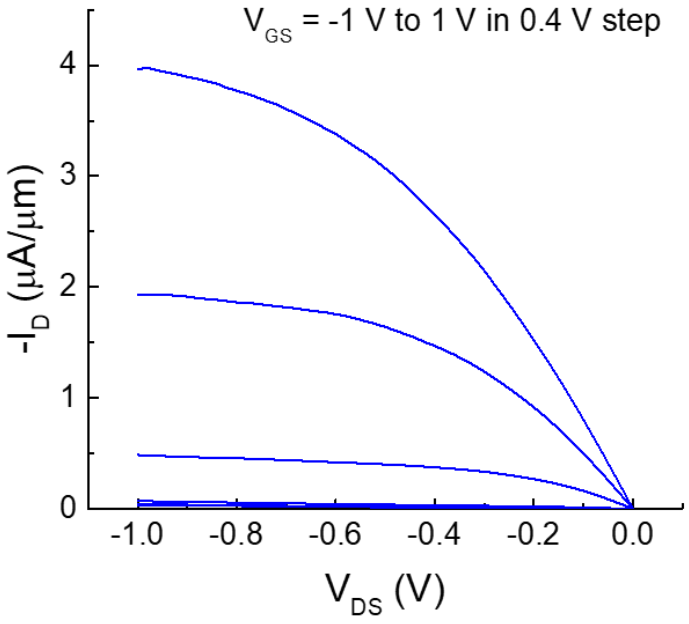


Fig. 4-18: Output characteristics of G3 InGaSb long channel array FinFET with W_f = 10 nm and L_g = 1 μm.

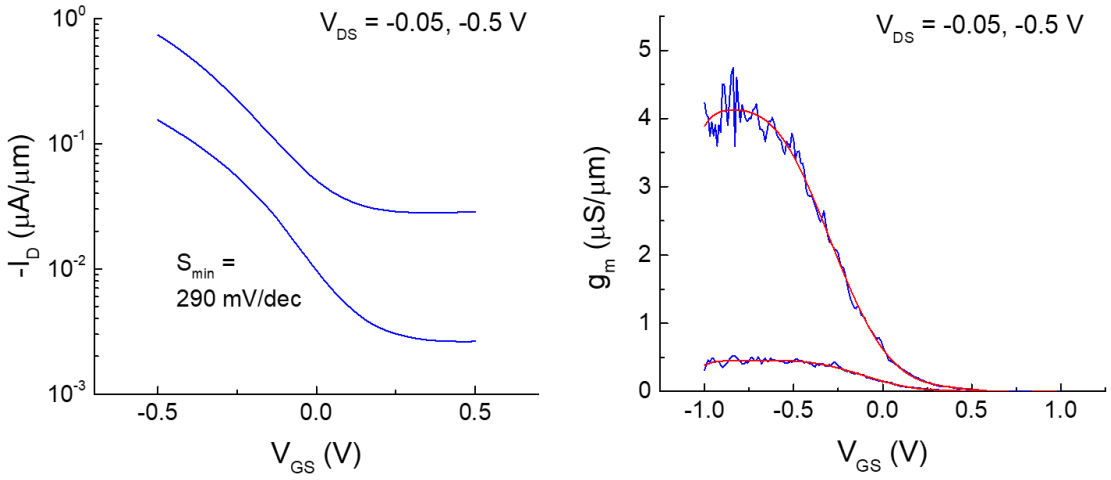


Fig. 4-19: (Left) subthreshold characteristics of G3 InGaSb single-fin FinFET with W_f = 10 nm and L_g = 20 nm. (Right) g_m characteristics of the device.

Figs. 4-20 and 4-21 show the scaling properties of V_T and g_m of G3 devices as a function of fin width down to 10 nm and gate length down to 20 nm. V_T extracted at V_{DS} = -50 mV shows mitigated V_T rolloff with fin width scaling. Maximum transconductance extracted at V_{DS} = -0.5 V shows consistent dependence on L_g and W_f.

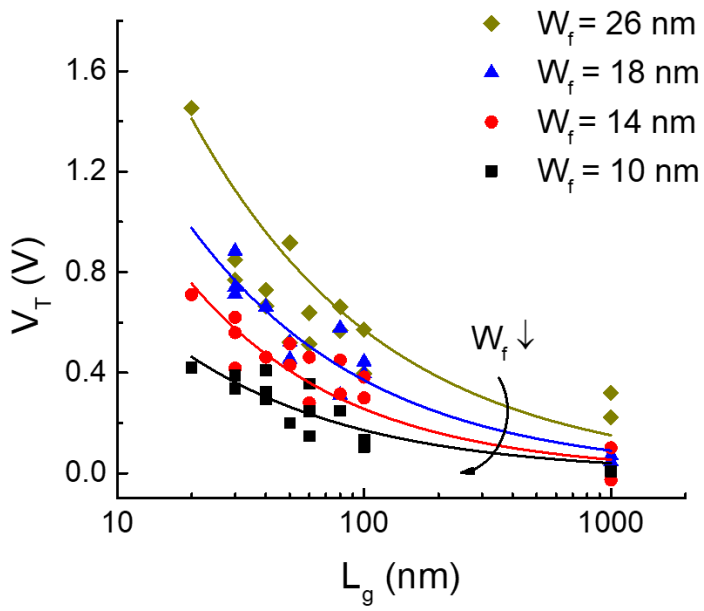


Fig. 4-20: Impact of W_f and L_g scaling on V_T at V_{DS} = -50 mV.

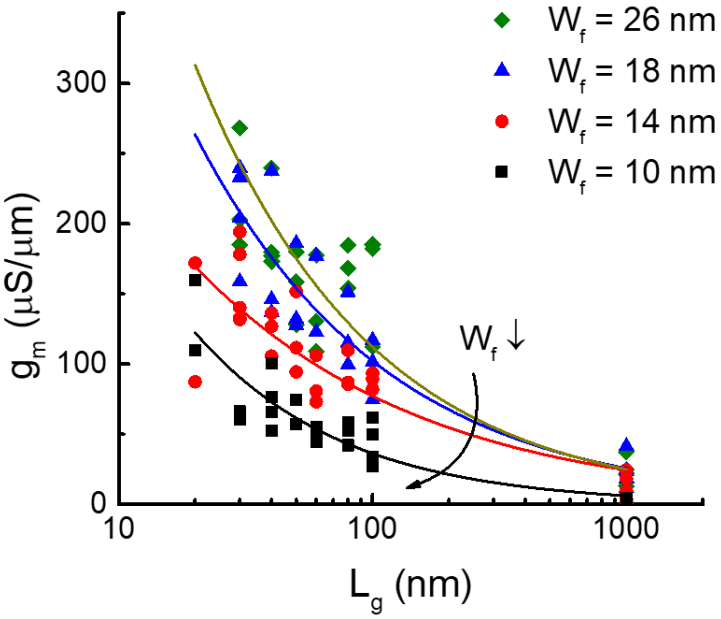


Fig. 4-21: Impact of W_f and L_g scaling on maximum g_m at V_{DS} = -0.5 V.

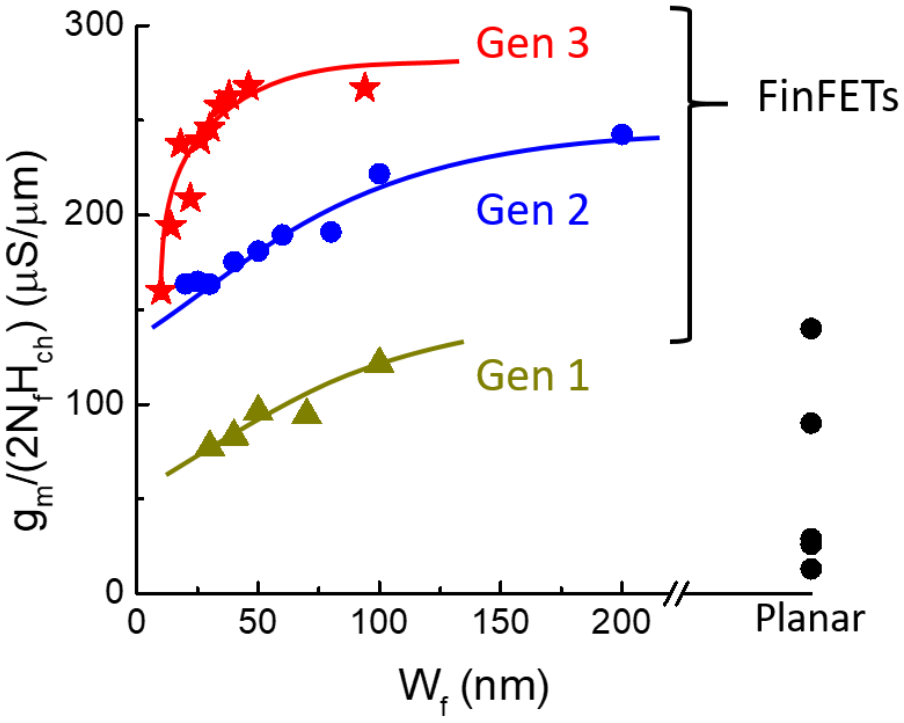


Fig. 4-22: Benchmark of maximum g_m vs. W_f for InGaSb p-channel FinFETs and planar antimonide-based p-channel MOSFETs from the literature. g_m is normalized to the total gate conducting periphery.

In G1 and G2, the maximum transconductance shown in the benchmarks in Fig. 4-13 (g_m vs. L_g) are quoted from FinFETs with very wide fin width (W_f = 100 nm for G1, and 200 nm for G2). In G3, devices with narrow fin width start to perform remarkably better. Fig. 4-22 benchmarks g_m at various fin widths of the three generations of FinFETs and antimonide-based planar MOSFETs. In G3, a record g_m of 268 μS/μm is exhibited at W_f = 46 nm and V_{DS} = 0.5 V. The maximum g_m drops as W_f decreases, as in InGaAs FinFETs [70]. However, for the device with narrow W_f = 10 nm, the maximum g_m is still higher than those of G1 and planar devices.

In the benchmark in Fig. 4-22 and earlier figures, g_m is normalized to the total gate-conducting periphery. A different perspective can be obtained by normalizing g_m by the fin width, as shown in Fig. 4-23. This figure of merit is necessary for multi-gate devices because it highlights the need of high device density. In the era of multi-gate transistors, the current scales in the vertical dimension, and therefore it is desired to achieve high current out of a small fin footprint. When normalized in this way, a significant gap in performance emerges among planar InGaSb PMOS and three generations of FinFETs. For G3 FinFETs, the 10 nm fin width devices demonstrate a high g_m/W_f = 704 μS/μm, which is by nearly a factor of 6X higher than G2 devices and 5X higher than planar devices. This stems from the aggressive scaling of fin width and heterostructure design of higher channel aspect ratio (H_{ch}/W_f) in G3 FinFETs.

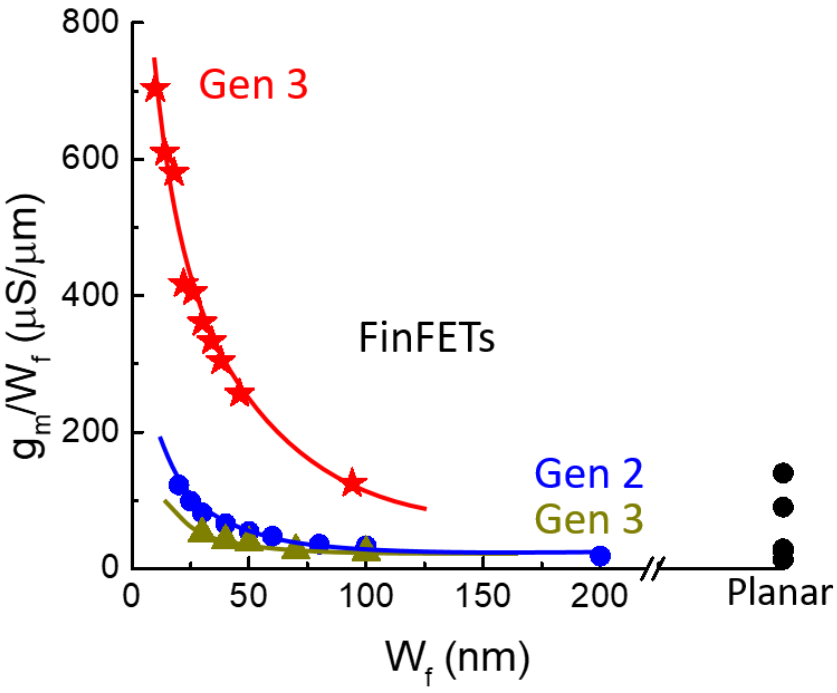


Fig. 4-23: Benchmark of maximum g_m vs. W_f for InGaSb p-channel FinFETs and planar antimonide-based p-channel MOSFETs. g_m is normalized to the fin footprint.

Having reviewed the DC electrical characteristics of the three generations of InGaSb FinFETs, other relevant device properties, such as ON-state resistance, strain effect, and OFF-state leakage current, are be investigated.

4.3 Analysis of ON-State Resistance

4.3.1 First Generation

The ON resistance (R_{on}) of the G1 InGaSb FinFETs is measured from the output characteristics at V_{GS} = -2.5 V for various fin width, normalized by the total conducting periphery. Fig. 4-24 shows the extracted R_{on} as a function of gate length. Fig. 4-25 shows the resistance of the fin channel, R_f, is extracted from the slop of R_{on} vs. L_g, and the access resistance from the source and drain, R_{SD}, is extracted from the extrapolation of R_{on} at zero L_g, plotted in log-log scale. Both R_f and R_{SD} increase as W_f scales down. The increase in the access resistance at small W_f poses a detrimental challenge in FinFET technology, as it has been observed in Si FinFETs [109]–[111] and InGaAs FinFETs [112]. Unlike conventional planar MOSFETs, FinFET lacks the equivalent deep source/drain low resistance silicide contact region, and formation of low resistance 3D junction to the fin is difficult. Therefore, ohmic contacts with low contact resistance are even more critical for multi-gate transistors at scaled dimensions. In G1 FinFETs, average R_{SD} = 372 Ω·μm for W_f = 100 nm, and it increases to an average R_{SD} = 756 Ω·μm for W_f = 30 nm. The increase in R_{SD} also contributes to the g_m degradation as W_f is reduced, as shown in Fig 4-8.

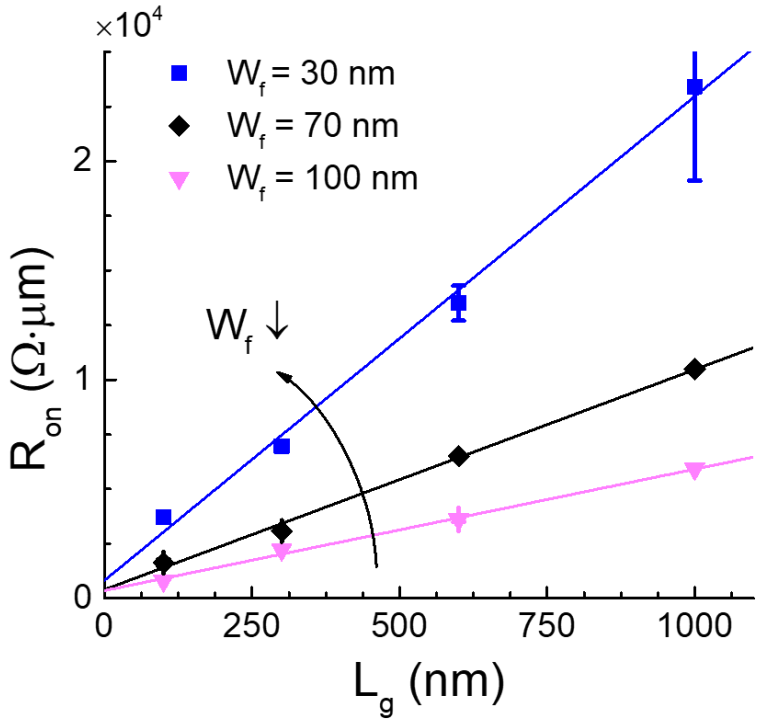


Fig. 4-24: ON resistance extraction of G1 InGaSb FinFETs with various W_f at V_{GS} = -2.5 V.

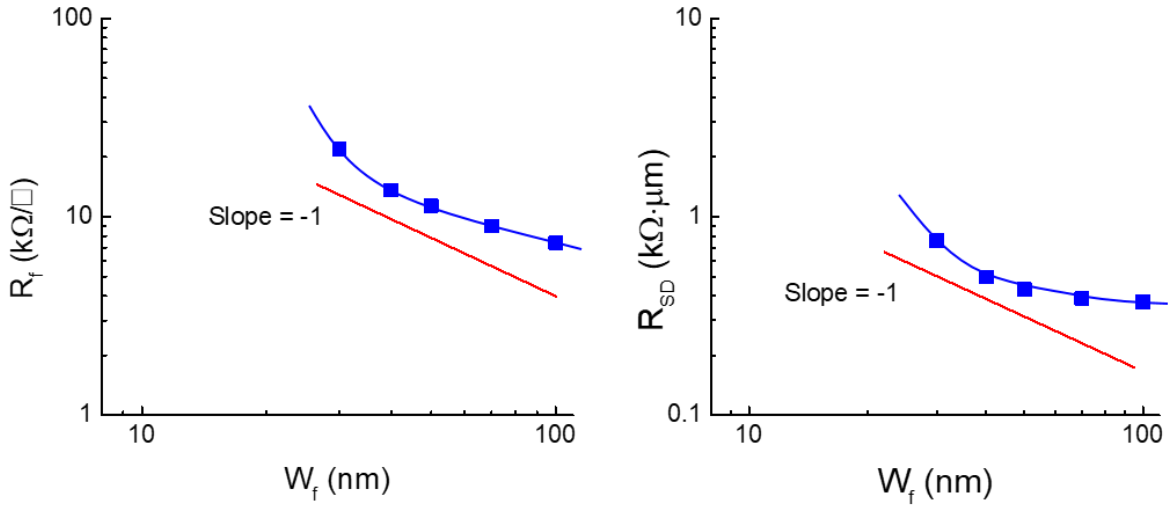


Fig. 4-25: Extracted (left) fin sheet resistance vs. fin width, and (right) access resistance vs. fin width of G1 InGaSb FinFETs, in log-log scale.

In addition, as shown in Fig. 4-25, R_f and R_{SD} increase as the fin width is decreased, while the trend doesn't follow closely as $1/W_f$ (slope = -1 in log-log scale). In narrow fin width, this acute fin width dependence could be attributed to Fermi level pinning at the fin sidewall surface and fin RIE damage, according to the hypothesis of "dead-zone" [112]. A dead-zone could exist as a depletion region at the fin surface, and it does not contribute to the transport. A simple dead-zone model can be described as:

$$\frac{1}{R_{sh}^{f}} = \frac{1}{R_{sh}^{b}}\left(1 - \frac{2x_d}{W_f}\right) \tag{4-9}$$

Where R_{sh}^{f} and R_{sh}^{b} denote the fin and bulk sheet resistance, respectively, and x_d is the dead-zone. From Eq. (1), R_{sh}^{b} and x_d can be extracted from Fig. 4-25, and the extracted x_d can be used to correct the access resistance to eliminate the fin width dependence. For G1 FinFETs, a dead-zone width of 9.3 nm is extracted, and an average corrected access resistance of 282 Ω·μm. For InGaAs FinFETs, the dead-zone width is about 5 nm [112]. The higher x_d in the InGaSb FinFETs suggests

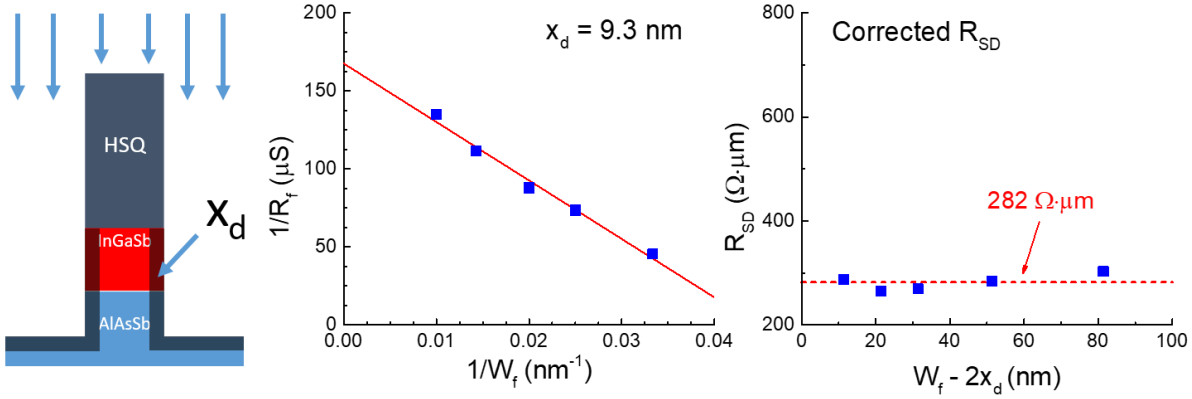


Fig. 4-26: (Left) schematic of the concept of dead-zone analysis on InGaSb FinFETs. (Middle) $1/R_{sh}^{f}$ vs. $1/W_f$ plot for dead-zone estimation. (Right) Access resistance corrected by the effective electrical fin width.

the need for techniques to mitigate sidewall damages and improve surface quality, for instance, using digital etch and proper annealing methods. Fig. 4-26 shows the concept of dead-zone analysis, the parameter extraction, and corrected access resistance.

4.3.2 Third Generation

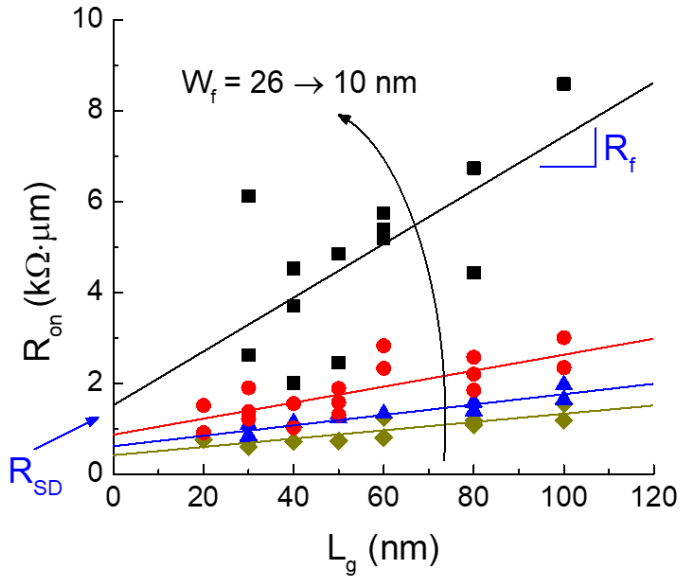


Fig. 4-27: ON resistance extraction of G3 InGaSb FinFETs with various W_f at V_{GS} = -1 V.

In the third generation FinFETs, the p^+-InAs/InAsSb composite cap design is used and the doping level is increased to $3 \cdot 10^{19}$ cm^{-3}. Fig. 4-27 shows the R_{on} in G3 as a function of gate length, for fin width from 26 to 10 nm, measured at V_{GS} = -1 V. Fig. 4-28 shows the extraction of R_f and R_{SD} in log-log scale, following the same as in Fig. 4-25. The expected scaling behavior of R_f and R_{SD} as ~$1/W_f$ is observed here. In G3 FinFETs, average R_{SD} = 253 Ω·μm for W_f = 94 nm, and it increases

to an average R_{SD} = 1.5 kΩ·μm for W_f = 10 nm. Fig. 4-29 plots the access resistance of G1 and G3 in linear scale. For W_f = 30 nm, the R_{on} of G3 FinFETs is about 50% of the R_{on} of G1 FinFETs.

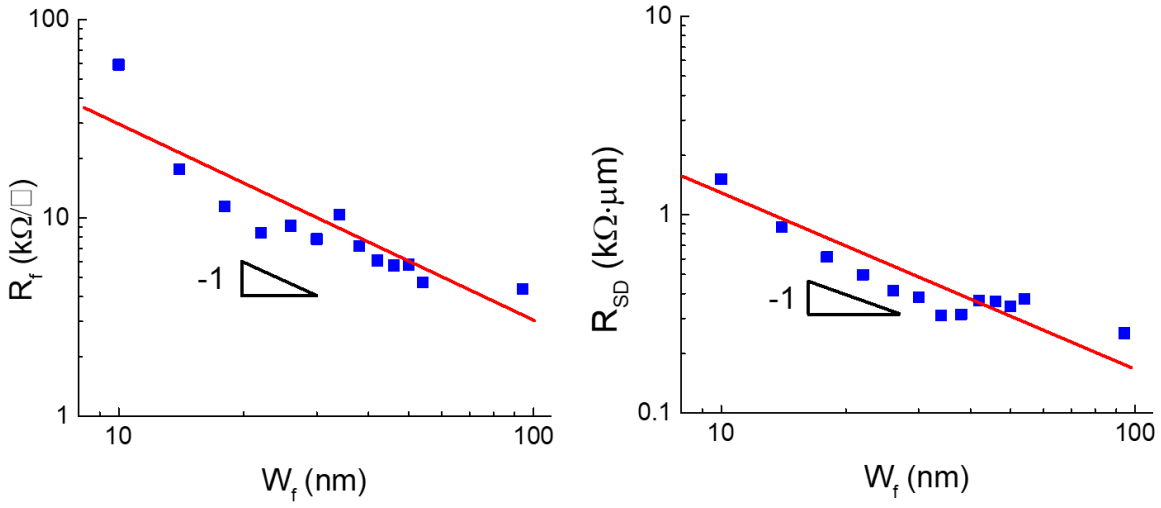


Fig. 4-28: Extracted (left) fin sheet resistance vs. fin width, and (right) access resistance vs. fin width of G3 InGaSb FinFETs, in log-log scale.

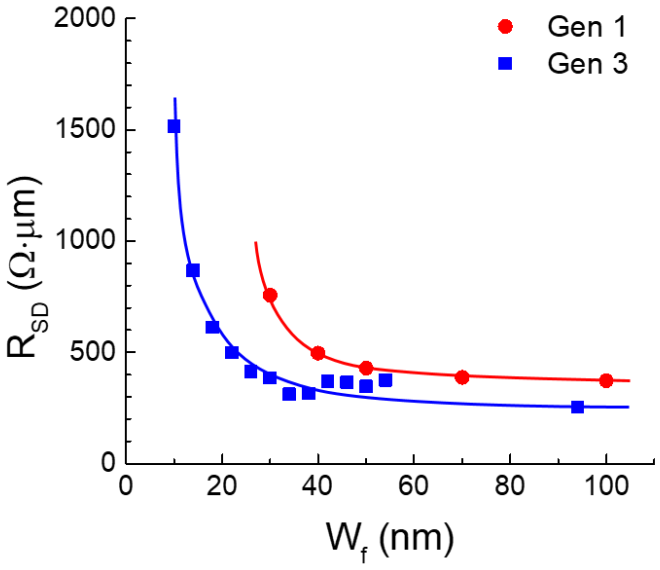


Fig. 4-29: Comparison of access resistance of G1 and G3 InGaSb p-channel FinFETs.

4.4 Impact of Strain

The application of strain has been shown to have favorable enhancement of the transport property in antimonide-based transistors. For $In_xGa_{1-x}Sb$ channel devices, the strain can be incorporated into the heterostructure through epitaxial growth, which results in biaxial strain. It can also be incorporated through processing using stressors, which results in uniaxial strain. Fig. 4-30 shows the percent change of drain current in the linear regime for an $In_{0.41}Ga_{0.59}Sb$ channel, as a function of <110> uniaxial stress [25]. Fig 4-30 also shows the intrinsic g_m as a function of compressive strain for InGaSb p-channel FETs with various indium composition [28], [44], [46], [99], [113]–[117]. Both graph shows the benefit of compressive strain for the antimonide-based transistors.

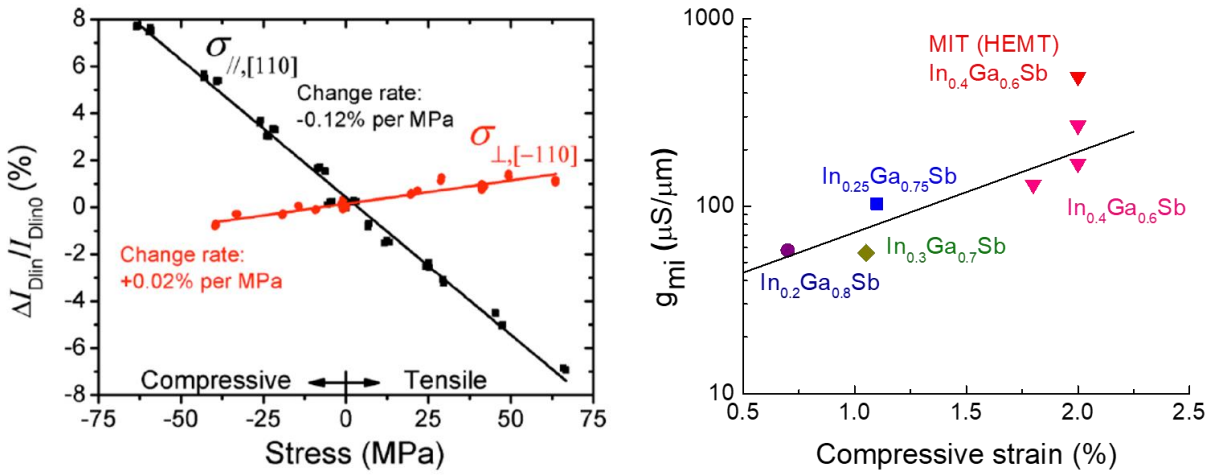


Fig. 4-30: (Left) the percent change in the linear drain current for an $In_{0.41}Ga_{0.59}Sb$ channel, as a function of <110> uniaxial stress. Image adapted from [25]. (Right) intrinsic transconductance as a function of compressive biaxial strain in the $In_xGa_{1-x}Sb$ channel. Image adapted from [99].

The data above are from InGaSb planar FETs. For InGaSb FinFETs, which is biaxially strained naturally by the lattice constant of the buffer layer, the strain will transform from biaxial to uniaxial as the fin width decreases. The impact of strain in p-channel FinFETs have been observed in SiGe FinFETs [13], but it has not been studied in III-V FinFETs.

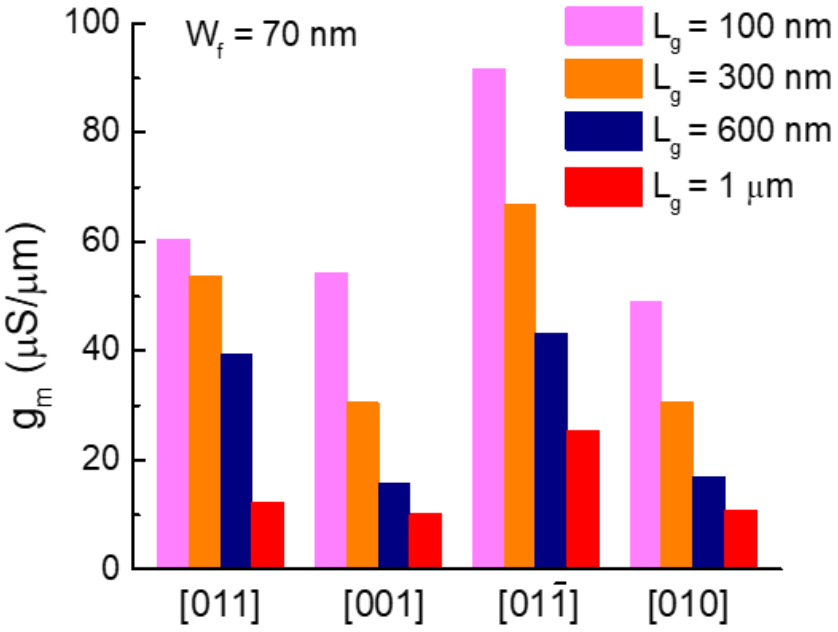


Fig. 4-31: Dependence of maximum g_m on fin orientation for G1 FinFETs with W_f = 70 nm and various L_g.

Fig. 4-31 shows the impact of fin orientation on g_m for G1 FinFETs with W_f = 70 nm. A strong orientation dependence is observed with FinFETs of fin oriented along the [01$\bar{1}$] direction being the best, those along [001] and [010] being the worst, and [011] being somewhere in between. A similar dependence is observed for all L_g and W_f. The orientation dependence is consistent with the 8×8 k·p Schrödinger and Poisson calculations of hole mobility enhancement in GaSb double-gate pMOSFETs under uniaxial stress [118]. The observed impact of fin orientation suggests that

the as-grown biaxial compressive stress in the pseudomorphic InGaSb channel has relaxed along the direction transversal to the fin leaving being a strong anisotropic strain distribution.

4.5 OFF-State Current

One of the greatest challenges in InGaSb FinFETs fabrication is the suppression of the off-state leakage current. As seen from the output characteristics of G1 FinFETs, the poor turn-off behavior is unacceptable for logic applications. The problematic leakage current stems from the highly reactive sidewall of the antimonide fins and the lack of techniques to clean and passivate the sidewall surfaces. As discussed in Ch. 2, much of the work of this thesis has been attended to the development of antimonide-compatible fin-sidewall treatment methods. In this section, we will investigate the behavior and origin of the leakage current based on the devices' electrical characteristics. We will also examine whether the techniques developed in this work are effective in mitigating the issue.

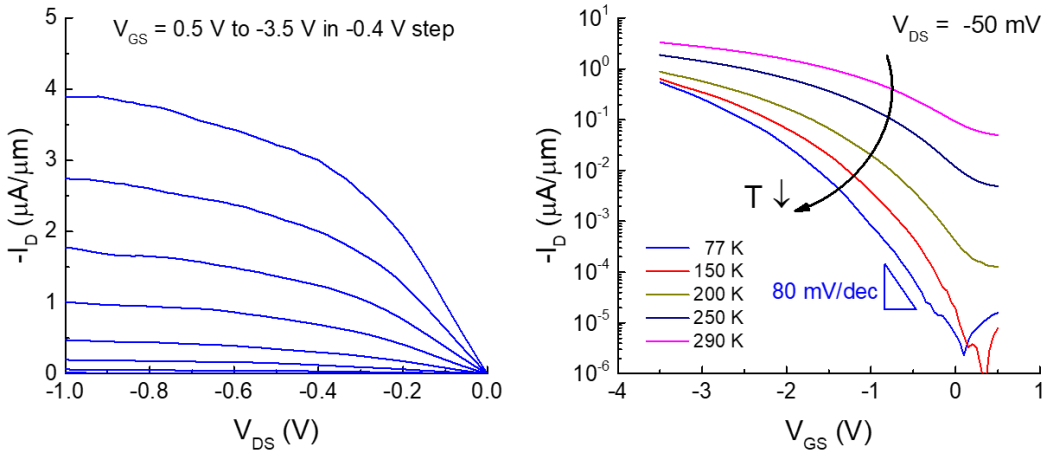


Fig. 4-32: (Left) low temperature (77 K) output characteristics of FinFET with W_f = 30 nm and L_g = 300 nm. (Right) Subthreshold characteristics of the same device at different temperatures (V_{DS} = -50 mV).

4.5.1 First Generation

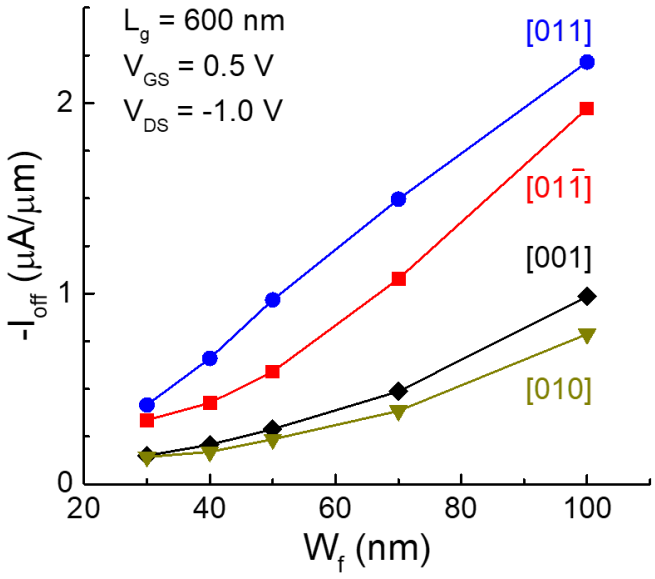


Fig. 4-33: Off-state current as a function of fin width and fin orientation, for FinFETs with L_g = 600 nm, measured at V_{GS} = 0.5 V and V_{DS} = -1.0 V.

Fig. 4-32 shows the low temperature output characteristics of a G1 FinFET with W_f = 30 nm and L_g = 300 nm, measured at 77 K in liquid nitrogen ambient, and the linear subthreshold characteristics (V_{DS} = - 50 mV) of the same device at temperature from 77 K to room temperature. The turn-off behavior is improved at low temperature, though it is not ideal, because the subthreshold slope is reduced to only 80 mV/dec at 77 K. This suggests that the high leakage current in the off-state should be attributed to both high defect state density at the InGaSb/high-k dielectric interface as well as thermal leakage current flowing within the device.

Fig. 4-33 shows the off-state current of devices with L_g = 600 nm, measured at V_{GS} = 0.5 V and V_{DS} = -1.0 V, as a function of fin width from 30 to 100 nm, as groups of four fin orientation. The off-state current shows strong dependence on the fin width and the fin orientation. As fin width decreseas, the off-current decreases. Higher leakage current is measured in [011] and [01$\bar{1}$] fins than the other two orientations. Such dependance suggests that a significant part of the off-state leakage current flows inside the fins, in addition to the possible leakage in the body of the device underneath the fins.

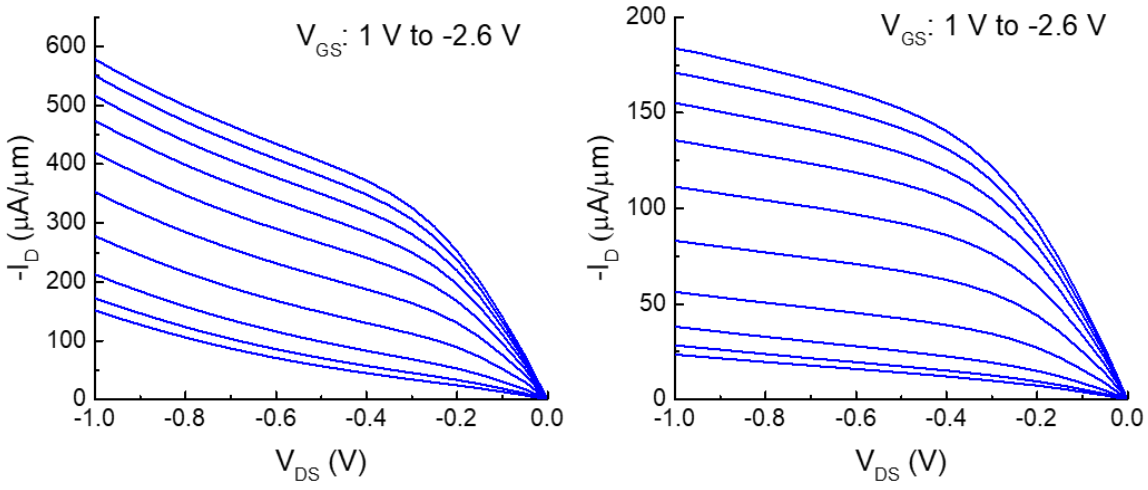


Fig. 4-34: Output characteristics of G1 FinFETs (Left) without any post-RIE surface treatment, and (right) with 1% HCl:H_2O dip for 30 seconds. The FinFETs have W_f = 70 nm and L_g = 250 nm, measured at V_{GS} = 1 V to -2.6 V in -0.4 V steps.

It is important to examine the impact of sidewall treatment methods at this point. In G1 FinFETs, conventional HCl:H_2O surface clean [75] has been attempted on a wafer which is processed in parallel. Fig. 4-34 shows the effect of a 1% HCl:H_2O treatment for 30 seconds, on two devices with the same dimensions of W_f = 70 nm and L_g = 250 nm. It shows improved output characteristics, in both on-state and off-state, though the drain current is decreased. Fig. 4-35

shows the subthreshold characteristics of the two devices with and without HCl:H_2O treatment. It shows clear improvement in the subthreshold swing, though the improvement is not sufficient. The SEM images in Fig. 4-35 reveal more details of what takes place from the HCl treatment. It shows vertical nanowire test structures on the same sample. Without HCl treatment, the etched nanowire has a diameter of 116 nm. After the 30 seconds HCl treatment, the diameter is reduced to 106 nm. In addition, the profile of the vertical nanowire is changed, and the surface of the sidewall turns out to be rougher. This observation is consistent with the ones discussed in Ch. 2, which shows that the water-based HCl treatment etches the antimonide sidewalls and is not compatible for FinFET fabrication.

4.5.2 Second Generation

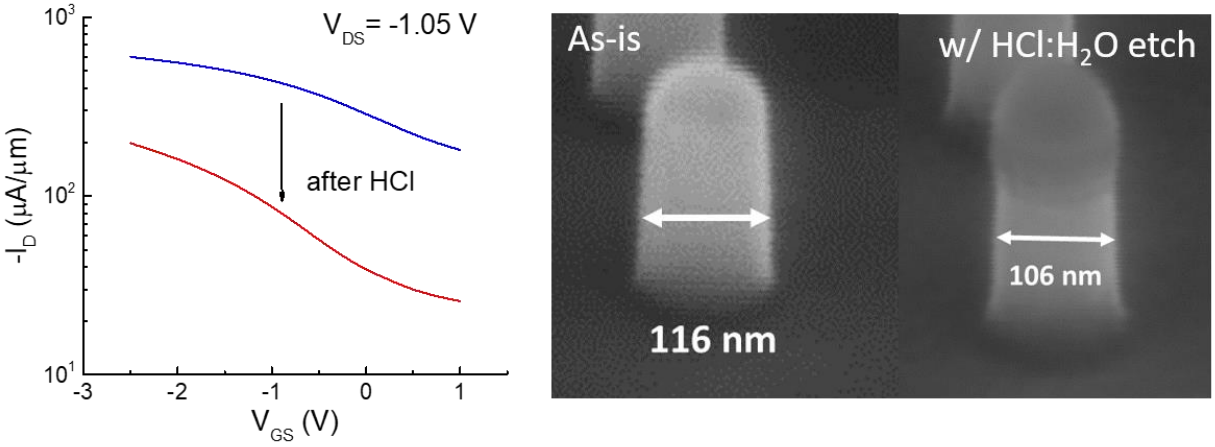


Fig. 4-35: (Left) subthreshold characteristics of the same devices in Fig. 4-34 showing the impact of HCl:H_2O treatment. (Right) SEM images of vertical nanowire test structures with and without the 1% HCl:H_2O 30 seconds treatment.

In second-generation FinFETs, the fin sidewall is treated with HCl diluted in isopropanol (IPA). The water-free alcohol-based HCl treatment is found to be self-limiting on the antimonide sidewall, as discussed in details in Section 2.3. Fig. 4-36 shows the output and subthreshold characteristics of two G2 FinFETs with and without treatment of 10% HCl:IPA for 30 seconds (W_f = 10 nm and L_g = 100 nm). HCl:IPA treatment results in a significant reduction of leakage current and an improvement on the saturation and turn-off behavior. A longer HCl:IPA treatment does not seem to further improve the turn-off behavior, and the minimum S saturates for W_f less than 40 nm. This suggests the presence of additional leakage paths in the device, such as undesired conduction within the fin buffer and at the buffer surface. This remains one of the major challenges in InGaSb FinFET fabrication, and more efforts is devoted in the third-generation devices to identify the origin the leakage paths and suppress them.

4.5.3 Third Generation

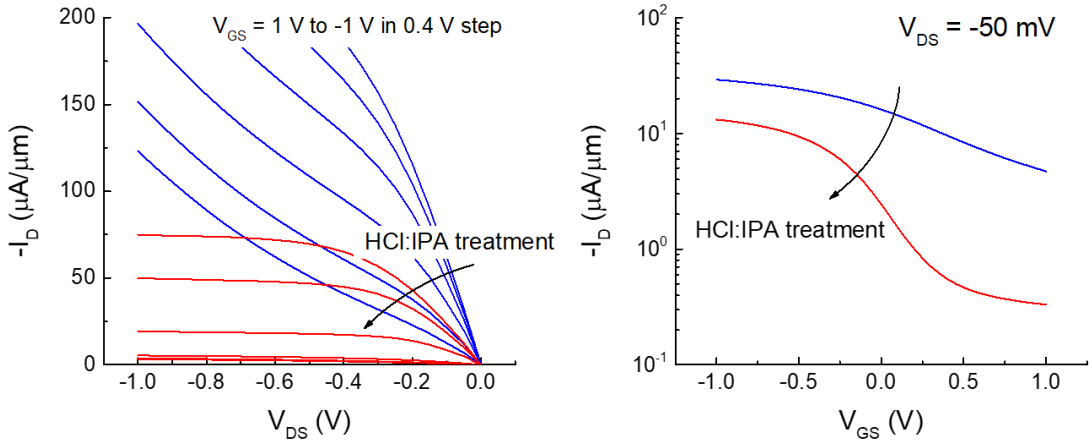


Fig. 4-36: (Left) output characteristics and (right) subthreshold characteristics of G2 FinFETs depicting the impact of 10% HCl:IPA treatment for 30 seconds. The FinFETs have W_f = 20 nm and L_g = 100 nm, measured at V_{GS} = 1 V to -1 V in -0.4 V steps.

In the third generation FinFETs, the antimonide-compatible digital etch is introduced for the first time, using oxidation in oxygen atmosphere combined with HCl:IPA treatment. Although the demonstrated devices, even at long channel, show insufficient turn-off characteristics, as seen in Fig. 4-17 and 4-18, the results nevertheless represent a significant improvement over the previous generations of InGaSb FinFETs due to the new digital etch technique. To illustrate its role, Fig. 4-37 shows the impact of digital etch on the subthreshold characteristics of identical long channel transistors (W_f = 20 nm, L_g = 1 μm), after 0, 1, and 4 cycles of digital etch. Without digital etch, turn-off is very poor. A single digital etch cycle yields a drastic improvement with the residual current dropping more than two decades. This illustrates the ability of digital etch to reduce sidewall RIE damage.

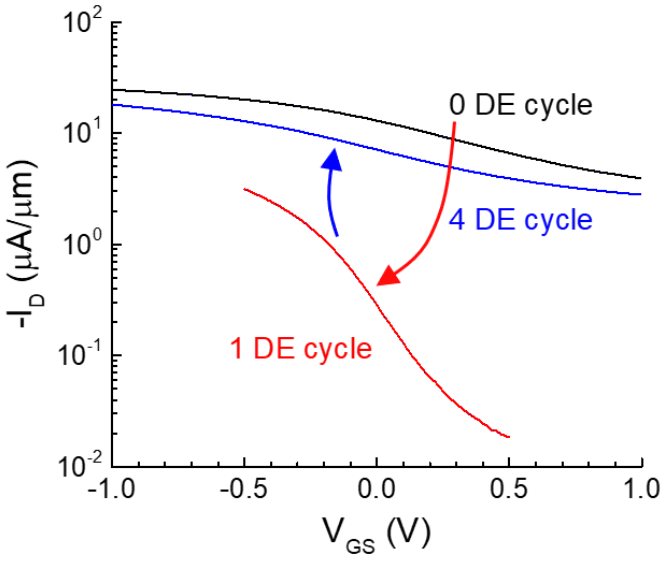


Fig. 4-37: Subthreshold characteristics of FinFETs with the same final W_f = 20 nm, after 0, 1, and 4 digital etch cycles. L_g = 1 μm for all the devices, and V_{DS} = - 50 mV.

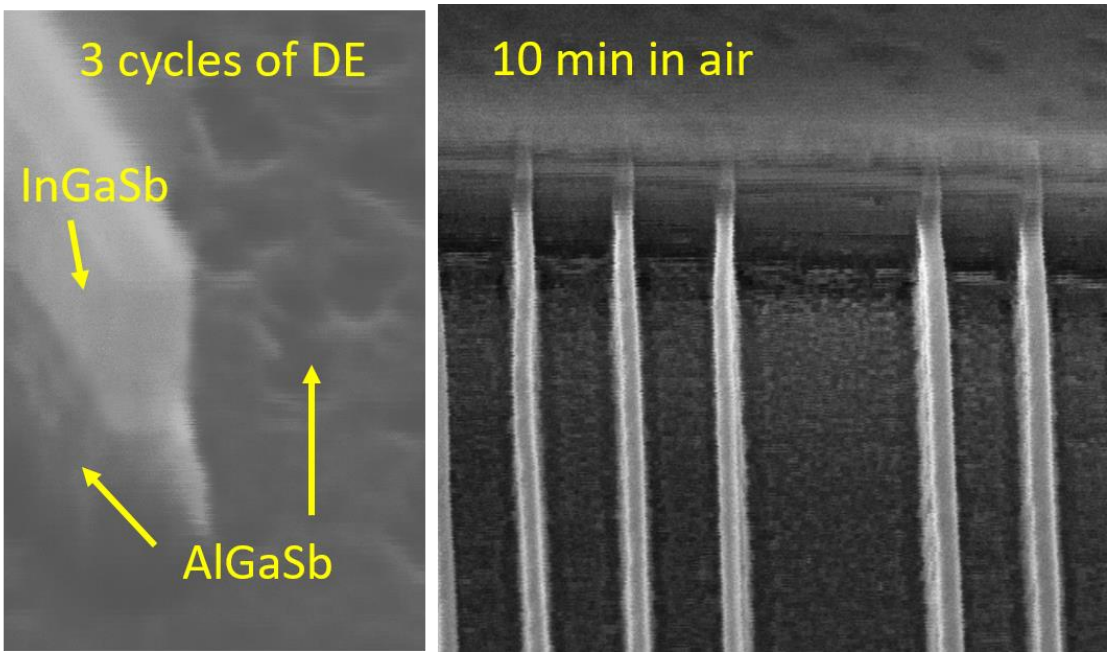


Fig. 4-38: (Left) SEM image of fin structure after 3 cycles of digital etch showing non-uniform sidewall and surface etching of the $Al_{0.97}Ga_{0.03}Sb$ portion of the structure. (Right) SEM images of fins after 10 min exposure in air after fin RIE.

On the other hand, the sample with an additional digital etch of 3 cycles shows a worse off-state behavior, as shown in Fig. 4-37. This could be attributed to the fact that the $Al_{0.93}Ga_{0.07}Sb$ buffer is extremely easy to be oxidized and damaged for its high aluminum composition. The oxidation of the buffer could be severe damaged, during the oxidation steps of the digital etch and the inevitable exposure of air and moisture in between digital etch cycles. In Ch. 2, this issue is not observed possibly due to lower Al composition in the buffer, which is less than 70%. The left SEM image in Fig. 4-38 shows the antimonide fin sidewall after 3 cycles of digital etch. It shows that the AlGaSb part of the fin sidewall and wafer surface is attacked and severely roughened, while the InGaSb portion of fin remains relatively smooth. The right SEM image in Fig. 4-38 shows the antimonide fins after long exposure (10 minutes) in air after fin RIE. The surface of the etched

device become very rough, which can be even observed by naked eyes via color change. It shows the challenges in the antimonide heterostructure design, and it requires further studies to prevent the FinFETs from excessive oxidation during digital etch.

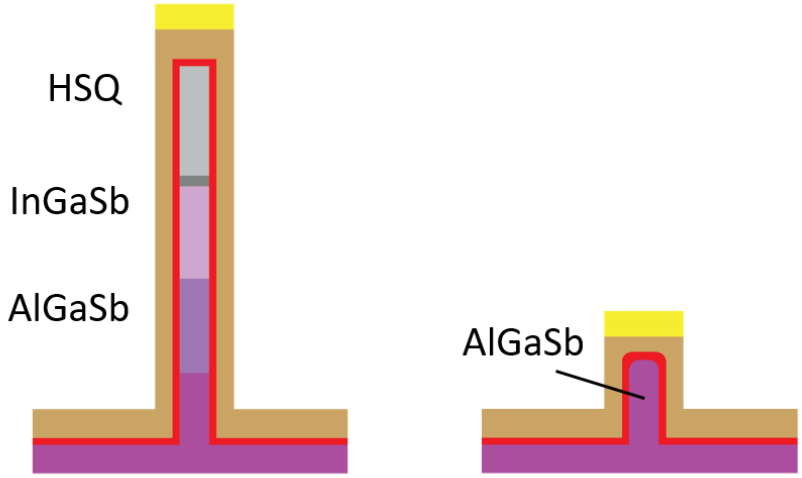


Fig. 4-39: (Left) Schematic of the cross-sections of (left) a normal FinFET, and (right) a test device with no InGaSb channel but gated AlGaSb fins, fabricated on the same wafer.

A siginificant portion of the residual off-state current that is left after one digital etch cycle in Fig. 4-37 is due to buffer leakage. This is illustrated in Fig. 4-39 and 4-40. Fig. 4-39 shows schematics of a normal InGaSb FinFET and a test FinFET without the InGaSb channel. The test device is fabricated on the same wafer as the G3 FinFETs. The HSQ fin mask of the test device is exposed with underdosed electron-beam lithography, so that the HSQ is consumed fast in the fin RIE, leaving the InGaSb portion of the fin unprotected and etched away. Therefore, it is in essence a gated AlGaSb buffer device.

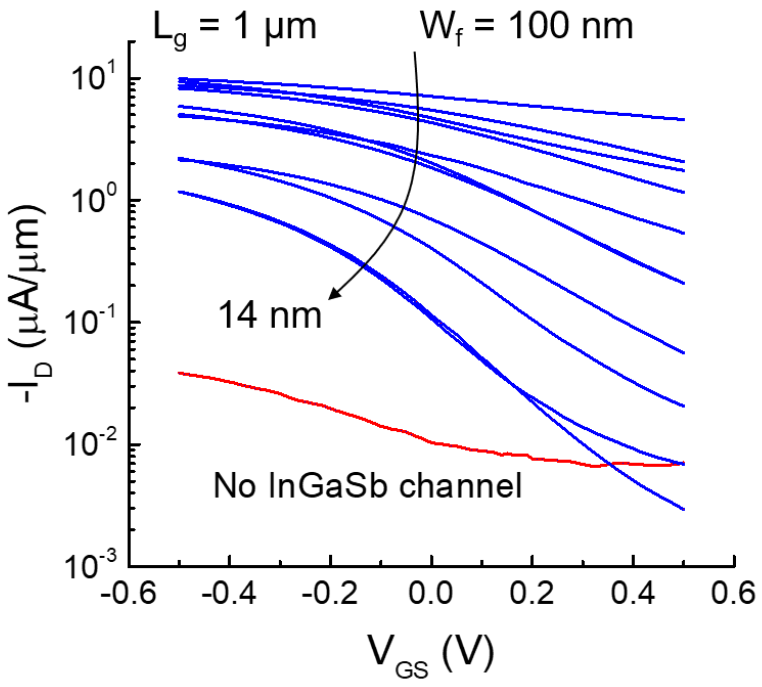


Fig. 4-40: Subthreshold characteristics of FinFETs with various W_f and L_g = 1 μm, and test device with identical geometry but no InGaSb channel. All devices are treated with one cycle of digital etch.

Fig. 4-40 shows the subthreshold characteristics of long-channel devices with fin width ranging from 100 nm to 14 nm. All devices were treated with one digital etch cycle. Below a fin width of 22 nm, the turn-off behavior remains unchanged. The test device with no InGaSb channel shows that there is a residual leakage path underneath the channel which is responsible for the improper off-state behavior. This is likely due to the etched and exposed AlGaSb buffer which is easier to be oxidized than InGaSb, forming a surface conduction leakage path. Alleivating this will require structural device innovations, such as lateral suspended or vertical device structures.

4.5.4 Leakage Analysis and Benchmark

In 2015, the first InGaSb p-channel FinFETs are demonstrated, and those devices suffered from poor turn-off behavior. In 2017, the third generation FinFETs have partially addressed this issue showing much improved turn-off behavior, by demonstrating an antimonide-compatible digital etch with alcohol-based treatment. As discussed by the previous sections, it seems that there exist at least four conducting paths contributing to the off-state current, which can be described as the following equations:

$$I_{off} = I_{ch} + I_{bf,fin} + I_{bf,surf} + I_{bf,body} \quad (4\text{-}10)$$

In Eq. (2), I_{ch} is the leakage flowing through the InGaSb fin channel, caused by partial Fermi-level pinning as a result of interface defects. $I_{bf,fin}$ is the leakage flowing through the AlGaSb buffer part of the fin body. $I_{bf,surf}$ is the leakage flowing through the surface of the buffer, including the vertical surface of the fin buffer and the horizontal mesa surface. $I_{bf,body}$ is the lekage flowing underneath the buffer body of the device. I_{ch}, $I_{bf,fin}$, and $I_{bf,surf}$ are assumed to be surface conductions. Each leakage path is expected to scale with various device dimensions. In the linear regime, to the first order:

$$I_{ch} \propto 2\frac{N_f H_{fc} W_f}{L_f} = 2H_{fc}\frac{W_{f,tot}}{L_f} \quad (4\text{-}11)$$

$$I_{bf,fin} \propto 2\frac{N_f W_f H_{fb}}{L_f} = 2H_{fb}\frac{W_{f,tot}}{L_f} \quad (4\text{-}12)$$

$$I_{bf,surf} \propto \frac{2N_f H_{fb} + \left(W_{ct} - W_{f,tot}\right)}{L_f} \approx \frac{2N_f\left(H_{fb} + L_d\right)}{L_f}, if\ N_f \gg 1 \tag{4-13}$$

$$I_{bf,body} \propto \frac{W_{ct}}{L_f} \tag{4-14}$$

The dimnesions in the above equations are sketched in Fig. 4-41. N_f is the number of fins, H_{fc} is the height of the channel portion of the fin, H_{fb} is the height of the buffer portion of the fin, $W_{f,tot}$ is the total fin width, W_{ct} is the total device width (which is equivalent to the mesa width), L_f is the fin length (which is equivalent to the gate length), and L_d is the spacing between the fins. The dimensions above are illustrated in Fig. 4-41.

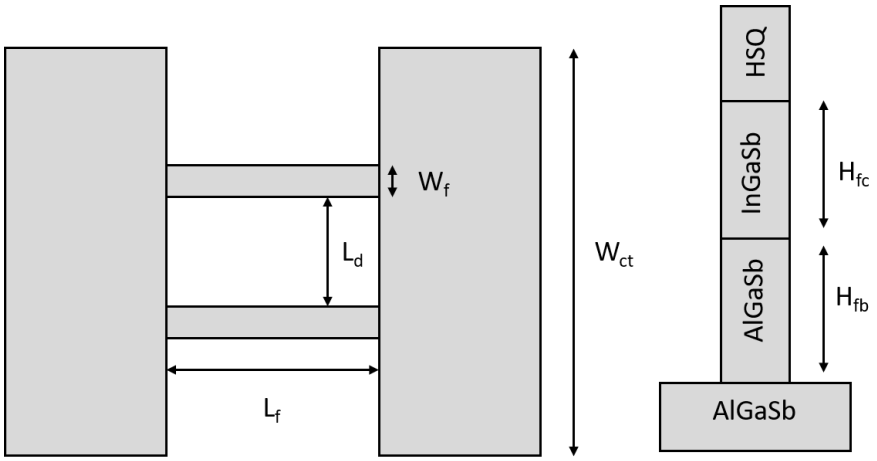


Fig. 4-41: (Left) top-view and (left) cross-sectional view (right) schematics of the FinFET structure with the definition of the dimensions.

As seen in Eqs. (3) and (4), unfortunately, I_{ch} and $I_{bf,fin}$ has the same scaling property, so it is difficult to distinguish them from scaling analysis. Other measurements such as low temperature measurement can be helpful since the leakage in the channel and buffer should have different temperature dependance and activation energy. Nevertheless, we try to look at the scaling properties of the InGaSb FinFETs and develop a qualitative understanding of the leakage current.

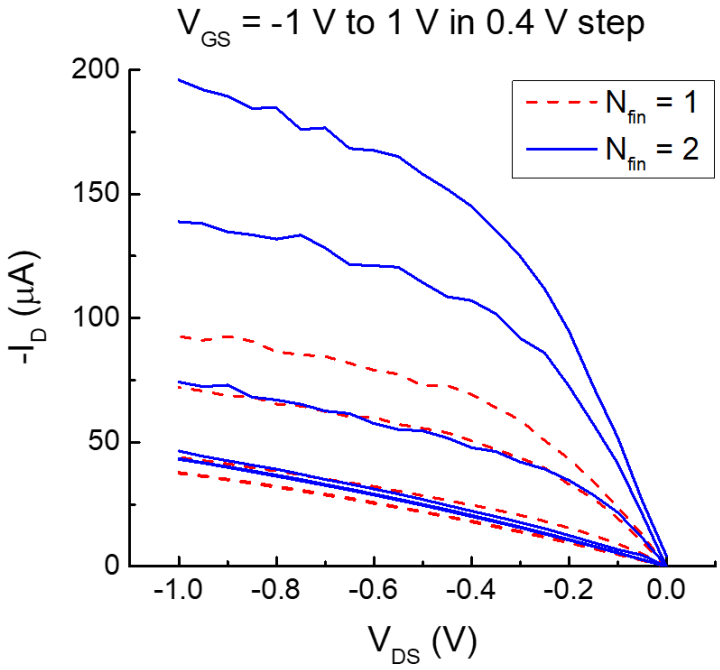


Fig. 4-42: Output characteristics of G2 InGaSb FinFETs with one or two fins per device. Fins are treated with HCl:IPA for one minute. The current is not normalized.

Fig. 4-42 shows the un-normalized output characteristics for G2 FinFETs with one or two fins (N_f = 1 and 2), and the same mesa width (W_{ct}), with the same W_f = 20 nm and L_g = 100 nm. When the device is off (V_{GS} = 1 V), the off-state drain current is approximately the same, while the on-state current scales linearly with N_f. This suggests different conducting paths in the on and off-state, and the presence of leakage paths outside the fins. Fig. 4-43 shows a more complete scaling

behavior of the off-state current in G2 FinFETs, extracted at V_{DS} = -50 mV and constant gate overdrive V_{GT} = 0.6 V. In the upper subfigure of Fig. 4-43, the unnormalized off-current is plotted as a function of total device width and gate length, in semi-log scale. In the lower subfigure, the off-current normalized by the total device width, W_{ct}, is plotted as a function of fin width and gate length, in log-log scale. All devices in Fig. 43 are array devices with $N_f \gg 1$. The un-normalized off-current is shown to increase as the total device width increases, and increases as the gate length decreases. This strong device width dependance suggests the existence of buffer leakage, in the form of $I_{bf,surf}$ or $I_{bf,body}$. When the off-current is normalized to the total device width, it shows strong dependance on the fin width. However, the suppression of the leakage current slows down

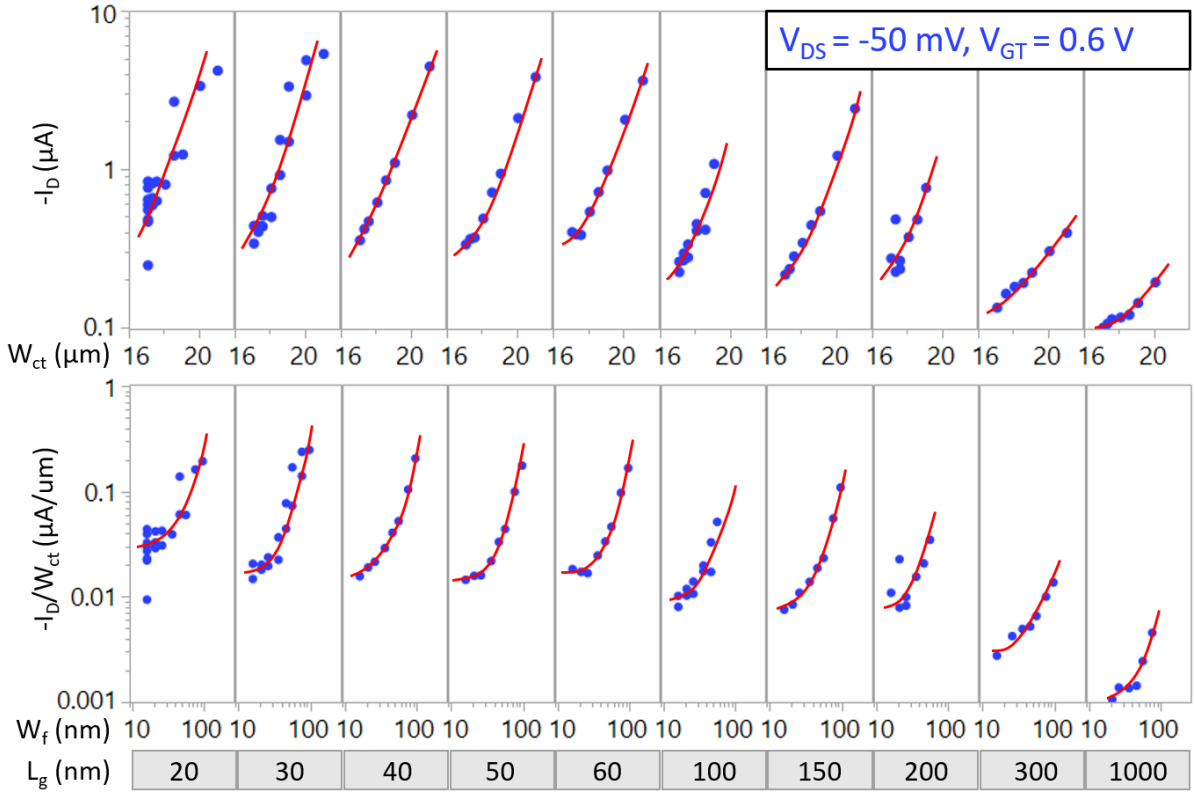


Fig. 4-43: (Top) un-normalized off-state current in G2 FinFETs vs. total device width and gate length. (Bottom) off-state current normalized to the total device width vs. fin width and gate length. I_D is extracted at V_{DS} = -50 mV and V_{GT} = 0.6 V.

when W_f is below 40 nm. This suggests the improperly closed InGaSb channel or AlGaSb fin, due to large D_{it}. This is determined by the etched fin sidewall quality.

Fig. 4-44 shows the scaling behavior of the drain current, normalized by the gate length, for G2 FinFETs, which are treated with HCl:IPA for 1 minute, and G3 FinFETs, which are treated with an addition digital etch cycle, measured in the linear regime (V_{DS} = -50 mV), at $V_{GT} = V_{GS} - V_T$ = 0.6 V. For the G2 devices (in blue), I_{off} shows a super-linear dependency on W_f and becomes less sensitive to W_f < 40 nm. On the other hand, for the G3 devices (in red), it shows a linear dependency on W_f. The difference in I_{off} may be attributed to the different heterostructure and

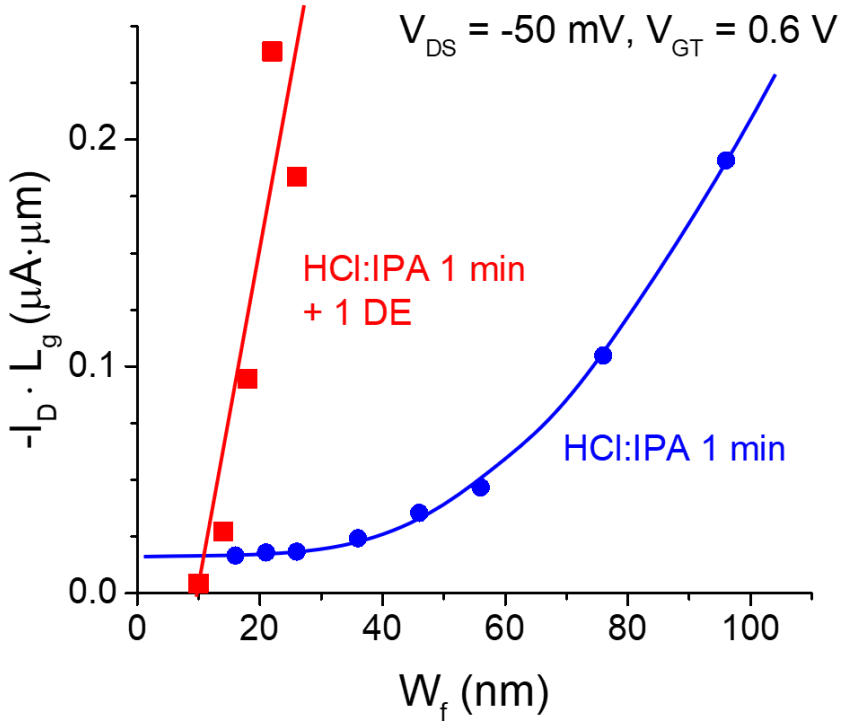


Fig. 4-44: Drain current normalized by the fin length vs. fin width, measured at V_{DS} = -50 mV and $V_{GS} - V_T$ = 0.6 V for G2 and G3 FinFETs.

processes of G2 and G3 devices. It shows that, in G3, the FinFETs are more effective in turning off the channel at scaled fin width.

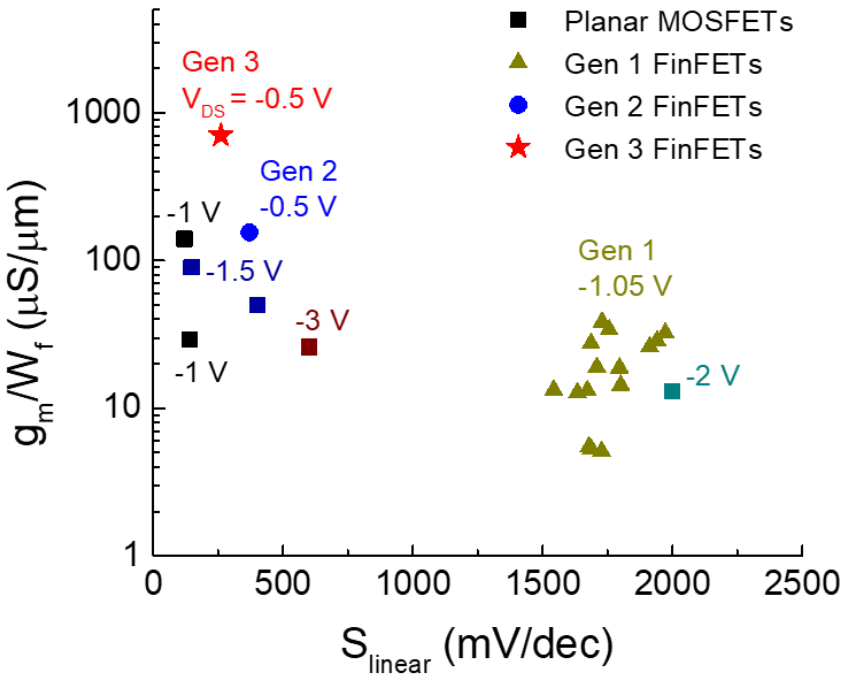


Fig. 4-45: Benchmark of maximum g_m vs. linear subthreshold slope for InGaSb p-channel FinFETs and planar antimonide-based p-channel MOSFETs. V_{DS} bias points for g_m extraction are noted next to each data points.

Finally, Fig. 4-45 shows a benchmark of the three generations of InGaSb p-channel FinFETs, together with antimonide-based planar MOSFETs in the literature. It benchmarks the peak transconductance, normalized by the device footprint, versus the minimum subthreshold slope in the linear regime. The drain voltages at which the maximum g_m is extracted is noted next to the data points. It shows significant improvement in both the transconducance and subthreshold slope

in the three generations of FinFETs, The G3 FinFETs demonstrate maximum g_m surpassing the best of the planar MOSFETs while obtaining a subthreshold slope approaching the lowest of them.

4.6 Chapter Summary

In this chapter the results of the three generations of InGaSb p-channel FinFETs were presented and analyzed. The scaling properties of the FinFETs were shown, and their performance was benchmarked with other published antimonide-based III-V p-channel MOSFETs. Analysis of the ON-state resistance was carried out, which shows reduced access resistance in the G3 devices. Dead-zone analysis was performed on the ON-state resistance, and an estimation of 9.3 nm dead-zone width is obtained. Then, the impact of uniaxial strain in InGaSb FinFETs was demonstrated for the first time. The rest of the chapter focuses on the investigation of the off-state current in each device generation. Scaling study of the devices suggests multiple leakage paths within the fin and body of the FinFETs. The alcohol-based surface treatment in G2 and the antimonide-based digital etch in G3 were found to be able to suppress the leakage current to some extent. In summary, the device electrical characteristics presented in this chapter shows that the technologies developed in the previous chapters have advanced the state-of-the-art of InGaSb MOSFETs.

CHAPTER 5. Summary and Future Work

5.1 Summary

This thesis has presented a pioneering work of achieving the first InGaSb p-channel FinFETs via a self-aligned top-down approach, preparing the antimonide-based compound semiconductors for their use in future logic and low power computing applications. It has advanced the technology of antimonide-based devices, an extremely important yet underdeveloped family of III-V compound semiconductors.

It first identifies the critical technological challenges that must be addressed to realize an antimonide-based multi-gate transistor. Three building blocks of the InGaSb FinFETs have been investigated. First is the dry etching technology of the antimonide-based heterostructure. A RIE process has been developed, and it is able to etch fin and vertical nanowire structures down to 15 nm fin width or diameter, with high aspect ratio, vertical profile and smooth sidewall. The RIE process is compatible to heterostructure containing both antimonides and arsenides. Second, in order to clean the etched fin sidewalls and further scaled down the device dimension, a novel alcohol-based antimonide-compatible digital etch technique has been invented. It demonstrates a consistent etch rate of 2 nm/cycle, and InGaSb vertical nanowire with 9 nm diameter has been fabricated. The alcohol-based digital etch also shows drastically improved mechanical yield on

InGaAs vertical nanowires, by demonstrating record vertical nanowires with diameter as narrow as 5 nm and aspect ratio over 45. The new digital etch serves as a critical step before the formation of high-dielectric/metal gate stack. Thirdly, the ohmic contact technology has been investigated and optimized to reduce the contact resistance. A composite p+ capping layer structure and a nickelide contact scheme have been developed, showing an ultra-low contact resistance of 22 Ω·μm.

The key technological components developed in this work have been integrated into a process flow that has been used to demonstrate the first InGaSb p-channel FinFET. The heterostructure design and epitaxy growth process have been discussed. Three generations of InGaSb FinFETs have been demonstrated. This antimonide-based FinFET process is contact-first, gate-last, self-aligned, Si-compatible, and has low very thermal budget, with an emphasis on scalability and manufacturability. In the third generation, aggressively scaled FinFETs with minimum fin width of 10 nm, channel height of 23 nm, and gate length of 20 nm has been successfully demonstrated.

The electrical performance of the InGaSb p-channel FINFETs have been presented and analyzed for the three generations of processes. The scaling properties of the FinFETs shows well-behaved FinFET characteristics. The performance has been benchmarked with other antimonide-based p-channel MOSFETs in the literature. The first generation FinFET shows comparable performance to the best of planar InGaSb pMOSFETs. The performance of the second and third generation devices exhibits record maximum transconductance. The maximum g_m per footprint in the third generation reaches over 700 μS/μm for devices with 10 nm fin width. Analysis of the ON-state resistance, uniaxial strain and OFF-state leakage current have been carried out. Scaling study of

the devices suggests multiple leakage paths within the fin and body of the transistor. The alcohol-based surface treatment and the antimonide-compatible digital etch have been found to be partially effective in improving the turn-off characteristics of the FinFETs.

In conclusion, core processing technologies for InGaSb p-channel FinFETs have been developed in this thesis, enabling the demonstration of the first of such devices. This thesis has not only highlighted the potential of InGaSb p-channel multigate MOSFETs, but also has pushed significantly the state-of-the-art of antimonide fabrication technology for more general applications in which the antimonide-based compound semiconductors can shine.

5.2 Suggestion for Future Research

In this thesis, the process technologies of InGaSb multigate transistor have been developed from scratch. Although the devices show promising characteristics, there are many imperfections. For example, the OFF-state current is still insufficient for proper logic applications. The ON-state performance is also not as good as the more mature technologies, such as Si_xGe_{1-x} FinFETs. Much further research could be explored, based on the technologies and studies in this thesis. These include

- Optimize or re-design the heterostructure design and epitaxy to avoid highly reactive buffer layers, such as AlGaSb with high Al composition.
- Optimize of the high-k dielectric deposition and the interface quality. This includes investigation of other possible ALD dielectric films, such La_2O_3, $LaAlO_3$, ZrO_2, etc.

- Study the impact of forming gas annealing and optimize the annealing condition to further improve the MOS interface.
- Develop InGaSb vertical nanowire MOSFETs to achieve higher scalability, gate control, and eliminate leakages from the buffer.
- Integrate the InGaSb FinFET technology into other types of devices, such as InAs/GaSb tunnel FETs (TFETs).
- Develop thermal atomic layer etching (ALE) technology for the InGaSb heterostructure, to obtain etching with sub-atomic layer precision, as well as the possibility of a fully in-situ ALE-ALD process of the gate stack.
- Also, single channel InGaSb based CMOS.

Appendix: Sample Process Flow of InGaSb FinFETs

Process flow of the third generation InGaSb FinFETs:

Module	Step	Specification	Recipe	Tool
Ohmic	Clean	DI water, solvent	3 min each	
	EBL	PMMA A8	3500 rpm, 1 min	TRL-coater
		Bake	180°C, 2 min	Hot-plate
		Exposure	10 nA, 1200-1400 $\mu C/cm^2$	Elionix
		Dev in MIBK:IPA (1:3)	90 s	
	Clean	$HCl:H_2O$ (1:3)	1 min	Acidhood
	Metal deposition	Ni	15 nm, 1 Å/s	EbeamFP
		Pt	10 nm, 1 Å/s	EbeamFP
		Au	20 nm, 1 Å/s	EbeamFP
	Lift-off	Acetone	> 2 hrs	
	Inspect	SEM		SEM
Gate-foot	Contact spacer	30 nm SiO_2	HFWJSIO 24 s, 300°C	STS-CVD
	EBL	ZEP:Anisol (1:3)	3000 rpm, 1 min	TRL-coater
		Bake	180°C, 2 min	Hot-plate

		Exposure	1 nA	Elionix
		Dev in xylene	1 min	
	Etching	SiO_2 etch	JQSIO2/CF4-AV	PQ/Oxford-100
	Inspect	SEM		SEM
	Resist strip	Hot NMP	> 2 hrs	
	Resist descum	Ashing	1 kW, 10 min	TRL-asher
Mesa	EBL	Pre-bake	120C°C 10 min	TRL-oven
		MaN-2403	3000 rpm, 1 min	TRL-coater
		Bake	80°C, 1 min	Hot-plate
		Exposure	10 nA, 400 $\mu C/cm^2$	Elionix
		Dev in MF CD-26	1 min	
	Etching	SiO_2 etch	JQSIO2/CF4-AV	PQ/Oxford-100
	Inspect	SEM		SEM
	Resist strip	Hot NMP	> 2 hrs	
	Resist descum	Ashing	1 kW, 15 min	TRL-asher
Cap recess	Dry cap recess	ICP RIE	Recipe 67	SAMCO
	Wet cap recess	Citric acid:H_2O_2 (10:1)	~30 s	Acidhood
	Inspect	SEM		SEM
Fin	Adhesion layer	Si_3N_4 deposition	HFSINXZ, 2 nm	STS-CVD

	EBL	6% HSQ	3500 rpm, 1 min	TRL-coater
		Exposure	1 nA	Elionix
		Dev in 25% TMAH	1 min	Acidhood
	Fin RIE	Sb RIE	Recipe 54	SAMCO
	Inspect	SEM	Test samples only	SEM
	Sidewall clean	10% HCl:IPA	1 min	Acidhood
	Digital etch	Oxidation	O_2 purge, 2 min	TRL-asher
		Oxide etch	10% HCl:IPA 1 min	Acidhood
	Inspect	SEM	Test samples only	SEM
Gate stack	Gate dielectric	Al_2O_3 ALD	35 cycles, 175°C	ICL-ALD
	Gate metal	Al sputtering	60 nm, 1 Å,/s	TRL-AJA
Gate head	EBL	PMMA A6	3500 rpm, 1 min	TRL-coater
		Bake	180°C, 2 min	Hot-plate
		Exposure	10 nA	Elionix
		Dev in MIBK:IPA (1:3)	90 s	
	Metal deposition	Ti	10 nm, 1 Å/s	EbeamFP
		Au	250 nm, 2 Å/s	EbeamFP
	Lift-off	Acetone	> 2 hrs	
	Gate metal etch	Al etch	Recipe 83	SAMCO
	Inspect	SEM		SEM
Via	Via spacer	TEOS deposition	TEOS_zero stress	Oxford-100

	EBL	ZEP-520A	3000 rpm, 1 min	TRL-coater
		Bake	180°C, 2 min	Hot-plate
		Exposure	10 nA, 560 uC/cm^2	Elionix
		Dev in xylene	60 s	
	Via Etch	TEOS etch	CF4-AV	Oxford-100
		Dielectric etch	Recipe 59	SAMCO
		SiO_2 etch	CF4-AV	Oxford-100
	Inspect	SEM	Inspect each etch	SEM
	Resist strip	Hot NMP	> 2 hrs	
	Resist descum	Ashing	15 min	TRL-asher
Pads	Photolithography	HMDS		TRL-HMDS
		AZ5214	3500 rpm, 30s	TRL-coater
		Bake	80°C, 5 min	Hot-plate
		Exposure	375 nm, 16 mJ	MLA-150
		Image reversal bake	105°C, 65 s	Hot-plate
		Flood exposure	35 s	OAI-Flood
		Dev in AZ-422	90 s	
	Metal deposition	Ti	10 nm, 1 Å/s	EbeamFP
		Au	350 nm, 2 Å/s	EbeamFP
	Lift-off	Acetone	> 2 hrs	
Annealing	FGA	5.5% H in 95% N_2	Various Temp	RTA-pieces

Recipes in the process

Deposition recipes:

SiO_2 CVD Tool: STS-CVD Recipe: HFSIO_WL Stabilization: 4 min 30 s Pressure: 900 mTorr Gas flow: 1420/392/10 sccm $N_2O/N_2/SiH_4$ 13.56 MHz on Showerhead Power: 50 W Platen temperature: 300°C Showerhead temperature: 250°C	**Si_3N_4 Adhesion Layer for HSQ** Tool: STS-CVD Recipe: HFSINXIN Stabilization: 1 min 30 s Pressure: 550 mTorr Gas flow: 1960/40/80 sccm $N_2/SiH_4/NH_3$ Power: 25 W Platen temperature: 300°C Showerhead temperature: 250°C Rate: 18-20 s for 2-3 nm
TEOS CVD Tool: Oxford-100 Recipe: Zero-stress TEOS Gas: 300/50 sccm O_2/ArTOES Pressure: 500 mTorr Temperature: 300°C LF power/pulse time: 40 W/8 s HF power/pulse time: 40 W/12 s	**Mo Gate Sputtering** Tool: AJA-TRL Recipe: Wenjie_Mo_30 nm Loadlock pressure < $3 \cdot 10^{-6}$ Torr Main chamber pressure < $3 \cdot 10^{-6}$ Torr Gas: 15 sccm Ar Deposition pressure: 2 mTorr Deposition power: 100 W Rate: 0.8 Å/s
Al Gate Sputtering Tool: AJA-TRL Recipe: Wenjie_Al_100 nm Loadlock pressure < $3 \cdot 10^{-6}$ Torr Main chamber pressure < $3 \cdot 10^{-6}$ Torr Gas: 35 sccm Ar Deposition pressure: 3 mTorr Deposition power: 150 W Rate: 0.8 Å/s	**Ni E-beam Evaporation** Tool: ebeamFP Chamber pressure: < 10^{-6} Torr Deposition power: 15% Deposition rate: 1 Å/s
HfO_2 ALD Tool: ICL-ALD Substrate temperature: 175°C TDMAH temperature: 97°C TDMAH pulse: 50 ms H_2O pulse: 20 ms Rate: 1 Å/s	**Al_2O_3 ALD** Tool: ICL-ALD Substrate temperature: 250°C Gas flow: 20 sccm N_2 TMA pulse: 15 ms H_2O pulse: 20 ms Rate: 1 Å/s

Etching recipes:

InGaSb Fin RIE Tool: SAMCO Recipe 74: Gas flow: $BCl_3/Ar/SiCl_4$ = 3/11/0.4 sccm (Adjustment of gas flow may be needed by condition of the chamber) Substrate temperature: 250°C ICP/Bias power: 20/280 W Pressure: 0.2 Pa	**Digital Etch** Pre-DE clean: 10% HCl:IPA 1 min Oxidation: Tool: Asher O_2 purge, 2 min Oxide etch: Tool: Acidhood 10% HCl:IPA 1 min
SiO_2/TEOS RIE Tool: Oxford-100 Recipe: CF4-AV Gas flow: 50 sccm CF_4 Pressure: 10 mTorr Power: 250 W	**SiO_2/TEOS RIE** Tool: Plasmaquest Recipe: JQSIO2 Gas flow: CF_4/H_2 = 32/3 sccm ECR power: 200 W DC bias: 100 V Pressure: 20 mTorr Rate: ~10 nm/min
Al_2O_3/HfO_2 RIE Tool: SAMCO Recipe: 59 Gas flow: BCl_3/Ar = 9/2 sccm Substrate temperature: 40°C ICP/Bias power: 75/75 W Pressure: 0.3 Pa	**III-V Cap Dry Recess** Tool: SAMCO Recipe: 67 Gas flow: Cl_2/N_2 = 10/3 sccm Substrate temperature: 120°C ICP/Bias power: 20/50 W Pressure: 0.2 Pa
Al RIE Tool: SAMCO Recipe: 83 Gas flow: BCl_3/Ar = 6/2 sccm Substrate temperature: 80°C ICP/Bias power: 75/75 W Pressure: 0.2 Pa	**Mo RIE** Tool: Oxford-100 Recipe: RIE-SF6-WJ-50W Gas flow: 45/5 sccm SF_6/O_2 Pressure: 20 mTorr Power: 50 W Rate: ~ 20 nm/min

Bibliography

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

[2] I. Ferain, C. A. Colinge, and J.-P. Colinge, "Multigate transistors as the future of classical metal-oxide-semiconductor field-effect transistors," *Nature*, vol. 479, no. 7373, pp. 310–316, Nov. 2011.

[3] H. Iwai, "Si MOSFET roadmap for 22nm and beyond," in *2009 4th International Conference on Computers and Devices for Communication (CODEC)*, 2009, pp. 1–4.

[4] A. Khakifirooz and D. A. Antoniadis, "Transistor Performance Scaling: The Role of Virtual Source Velocity and Its Mobility Dependence," in *2006 International Electron Devices Meeting*, 2006, pp. 1–4.

[5] "Technological Progress," *Our World in Data*. [Online]. Available: https://ourworldindata.org/technological-progress. [Accessed: 12-Apr-2018].

[6] S. Kumar, "Fundamental Limits to Moore's Law," *ArXiv151105956 Cond-Mat*, Nov. 2015.

[7] M. M. Waldrop, "The chips are down for Moore's law," *Nat. News*, vol. 530, no. 7589, p. 144, Feb. 2016.

[8] D. J. Frank, R. H. Dennard, E. Nowak, P. M. Solomon, Y. Taur, and H.-S. P. Wong, "Device scaling limits of Si MOSFETs and their application dependencies," *Proc. IEEE*, vol. 89, no. 3, pp. 259–288, Mar. 2001.

[9] J. del Alamo, *Integrated Microelectronic Devices: Physics and Modeling*. Pearson, 2018.

[10] D. J. Frank, "Power-constrained CMOS scaling limits," *IBM J. Res. Dev.*, vol. 46, no. 2.3, pp. 235–244, Mar. 2002.

[11] M. C. Lemme, L.-J. Li, T. Palacios, and F. Schwierz, "Two-dimensional materials for electronic applications," *MRS Bull.*, vol. 39, no. 08, pp. 711–718, 2014.

[12] M. M. Shulaker *et al.*, “Carbon nanotube computer,” *Nature*, vol. 501, no. 7468, pp. 526–530, Sep. 2013.

[13] P. Hashemi *et al.*, “First demonstration of high-Ge-content strained-$Si_{1-x}Ge_x$ (x=0.5) on insulator PMOS FinFETs with high hole mobility and aggressively scaled fin dimensions and gate lengths for high-performance applications,” in *Electron Devices Meeting (IEDM), 2014 IEEE International*, 2014, pp. 16.1.1-16.1.4.

[14] J. A. D. Alamo, D. A. Antoniadis, J. Lin, W. Lu, A. Vardi, and X. Zhao, “Nanometer-Scale III-V MOSFETs,” *IEEE J. Electron Devices Soc.*, vol. 4, no. 5, pp. 205–214, Sep. 2016.

[15] D. A. Antoniadis, I. Aberg, C. N. Chleirigh, O. M. Nayfeh, A. Khakifirooz, and J. L. Hoyt, “Continuous MOSFET performance increase with device scaling: The role of strain and channel material innovations,” *IBM J. Res. Dev.*, vol. 50, no. 4.5, pp. 363–376, Jul. 2006.

[16] J. Lin, D. A. Antoniadis, and J. A. del Alamo, “Novel intrinsic and extrinsic engineering for high-performance high-density self-aligned InGaAs MOSFETs: Precise channel thickness control and sub-40-nm metal contacts,” in *2014 IEEE International Electron Devices Meeting*, 2014, pp. 25.1.1-25.1.4.

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

[18] C. Auth *et al.*, “A 10nm high performance and low-power CMOS technology featuring 3rd generation FinFET transistors, Self-Aligned Quad Patterning, contact over active gate and cobalt local interconnects,” in *2017 IEEE International Electron Devices Meeting (IEDM)*, 2017, pp. 29.1.1-29.1.4.

[19] M. D. Giles *et al.*, “Understanding stress enhanced performance in Intel 90nm CMOS technology,” in *Digest of Technical Papers. 2004 Symposium on VLSI Technology, 2004.*, 2004, pp. 118–119.

[20] C. Auth *et al.*, “45nm High-k + metal gate strain-enhanced transistors,” in *2008 Symposium on VLSI Technology*, 2008, pp. 128–129.

[21] P. Packan *et al.*, “High performance 32nm logic technology featuring 2nd generation high-k + metal gate transistors,” in *2009 IEEE International Electron Devices Meeting (IEDM)*, 2009, pp. 1–4.

[22] C. Auth *et al.*, “A 22nm high performance and low-power CMOS technology featuring fully-depleted tri-gate transistors, self-aligned contacts and high density MIM capacitors,” in *2012 Symposium on VLSI Technology (VLSIT)*, 2012, pp. 131–132.

[23] P. Hashemi *et al.*, “High performance and reliable strained SiGe PMOS FinFETs enabled by advanced gate stack engineering,” in *2017 IEEE International Electron Devices Meeting (IEDM)*, 2017, pp. 37.3.1-37.3.4.

[24] D. Lin *et al.*, “Enabling the high-performance InGaAs/Ge CMOS: a common gate stack solution,” in *2009 IEEE International Electron Devices Meeting (IEDM)*, 2009, pp. 1–4.

[25] L. Xia, J. B. Boos, B. R. Bennett, M. G. Ancona, and J. A. del Alamo, “Hole mobility enhancement in In0.41Ga0.59Sb quantum-well field-effect transistors,” *Appl. Phys. Lett.*, vol. 98, no. 5, p. 053505, Jan. 2011.

[26] B. R. Bennett, M. G. Ancona, J. B. Boos, and B. V. Shanabrook, “Mobility enhancement in strained p-InGaSb quantum wells,” *Appl. Phys. Lett.*, vol. 91, no. 4, p. 042104, Jul. 2007.

[27] B. R. Bennett, M. G. Ancona, J. B. Boos, C. B. Canedy, and S. A. Khan, “Strained GaSb/AlAsSb quantum wells for p-channel field-effect transistors,” *J. Cryst. Growth*, vol. 311, no. 1, pp. 47–53, Dec. 2008.

[28] M. Radosavljevic *et al.*, “High-performance 40nm gate length InSb p-channel compressively strained quantum well field effect transistors for low-power (V_{CC}=0.5V) logic applications,” in *Electron Devices Meeting, 2008. IEDM 2008. IEEE International*, 2008, pp. 1–4.

[29] T.-W. Kim, D.-H. Kim, and J. A. del Alamo, “60 nm self-aligned-gate InGaAs HEMTs with record high-frequency characteristics,” in *Electron Devices Meeting (IEDM), 2010 IEEE International*, 2010, pp. 30.7.1-30.7.4.

[30] N. Waldron, D.-H. Kim, and J. A. del Alamo, “A self-aligned InGaAs HEMT architecture for logic applications,” *IEEE Trans. Electron Devices*, vol. 57, no. 1, pp. 297–304, Jan. 2010.

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

[32] A. Vardi, J. Lin, W. Lu, X. Zhao, and J. A. del Alamo, “High aspect ratio InGaAs FinFETs with sub-20 nm fin width,” in *2016 IEEE Symposium on VLSI Technology*, 2016, pp. 1–2.

[33] N. Waldron *et al.*, “InGaAs Gate-All-Around Nanowire Devices on 300mm Si Substrates,” *IEEE Electron Device Lett.*, vol. 35, no. 11, pp. 1097–1099, Nov. 2014.

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

[35] K. H. Goh *et al.*, “Gate-all-around CMOS (InAs n-FET and GaSb p-FET) based on vertically-stacked nanowires on a Si platform, enabled by extremely-thin buffer layer technology and common gate stack and contact modules,” in *2015 IEEE International Electron Devices Meeting (IEDM)*, 2015, pp. 15.4.1-15.4.4.

[36] M. Yokoyama, “InAs/GaSb-on-insulator single channel complementary metal-oxide-semiconductor transistors on Si structure: Applied Physics Letters: Vol 109, No 21,” 2016. [Online]. Available: http://aip.scitation.org/doi/abs/10.1063/1.4968187. [Accessed: 02-Aug-2017].

[37] A. W. Dey, J. Svensson, B. M. Borg, M. Ek, and L.-E. Wernersson, “Single InAs/GaSb Nanowire Low-Power CMOS Inverter,” *Nano Lett.*, vol. 12, no. 11, pp. 5593–5597, Nov. 2012.

[38] G. Zhou *et al.*, “Novel gate-recessed vertical InAs/GaSb TFETs with record high I_{ON} of 180 μA/μm at V_{DS} = 0.5 V,” in *Electron Devices Meeting (IEDM), 2012 IEEE International*, 2012, pp. 32.6.1-32.6.4.

[39] E. Memišević, J. Svensson, M. Hellenbrand, E. Lind, and L. E. Wernersson, "Scaling of Vertical InAs-GaSb Nanowire Tunneling Field-Effect Transistors on Si," *IEEE Electron Device Lett.*, vol. 37, no. 5, pp. 549–552, May 2016.

[40] 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 Ion = 10 μA/μm for I_{off} =1 nA/ μm at V_{ds} = 0.3 V," in *2016 IEEE International Electron Devices Meeting (IEDM)*, 2016, pp. 19.1.1-19.1.4.

[41] A. S. Babadi, J. Svensson, E. Lind, and L.-E. Wernersson, "Impact of doping and diameter on the electrical properties of GaSb nanowires," *Appl. Phys. Lett.*, vol. 110, no. 5, p. 053502, Jan. 2017.

[42] K. Nishi *et al.*, "High hole mobility front-gate InAs/InGaSb-OI single structure CMOS on Si," in *2015 Symposium on VLSI Technology (VLSI Technology)*, 2015, pp. T174–T175.

[43] R. L. Chu *et al.*, "Passivation of GaSb using molecular beam epitaxy Y2O3 to achieve low interfacial trap density and high-performance self-aligned inversion-channel p-metal-oxide-semiconductor field-effect-transistors," *Appl. Phys. Lett.*, vol. 105, no. 18, p. 182106, Nov. 2014.

[44] A. Nainani *et al.*, "Development of high-k dielectric for antimonides and a sub 350 C III-V pMOSFET outperforming Germanium," in *Electron Devices Meeting (IEDM), 2010 IEEE International*, 2010, pp. 6.4.1-6.4.4.

[45] B. R. Bennett, A. S. Bracker, R. Magno, J. B. Boos, R. Bass, and D. Park, "Monolithic integration of resonant interband tunneling diodes and high electron mobility transistors in the InAs/GaSb/AlSb material system," *J. Vac. Sci. Technol. B Microelectron. Nanometer Struct. Process. Meas. Phenom.*, vol. 18, no. 3, pp. 1650–1652, May 2000.

[46] Z. Yuan *et al.*, "Antimonide-Based Heterostructure p-Channel MOSFETs With Ni-Alloy Source/Drain," *IEEE Electron Device Lett.*, vol. 34, no. 11, pp. 1367–1369, 2013.

[47] M. Xu, R. Wang, and P. D. Ye, "GaSb Inversion-Mode PMOSFETs With Atomic-Layer-Deposited as Gate Dielectric," *IEEE Electron Device Lett.*, vol. 32, no. 7, pp. 883–885, Jul. 2011.

[48] P. Nagaiah *et al.*, “In Situ Deposited HfO2 with Amorphous-Si Passivation as a Potential Gate Stack for High Mobility (In)GaSb- Based P-MOSFETs,” *ECS Trans.*, vol. 41, no. 3, pp. 223–230, Oct. 2011.

[49] R. Magno *et al.*, “High speed, low power electronics using Sb-based semiconductors,” in *International Symposium on Compound Semiconductors, 2003*, 2003, pp. 175–176.

[50] A. Nainani *et al.*, “Optimization of the Al2O3/GaSb Interface and a High-Mobility GaSb pMOSFET,” Oct. 2011.

[51] M. Radosavljevic *et al.*, “Non-planar, multi-gate InGaAs quantum well field effect transistors with high-K gate dielectric and ultra-scaled gate-to-drain/gate-to-source separation for low power logic applications,” in *Electron Devices Meeting (IEDM), 2010 IEEE International*, 2010, pp. 6.1.1-6.1.4.

[52] S.-H. Kim *et al.*, “High Performance Tri-Gate Extremely Thin-Body InAs-On-Insulator MOSFETs With High Short Channel Effect Immunity and Tunability,” *IEEE Trans. Electron Devices*, vol. 61, no. 5, pp. 1354–1360, May 2014.

[53] X. Zhou *et al.*, “Scalability of InGaAs gate-all-around FET integrated on 300mm Si platform: Demonstration of channel width down to 7nm and Lg down to 36nm,” in *2016 IEEE Symposium on VLSI Technology*, 2016, pp. 1–2.

[54] O. P. Kilpi, J. Wu, J. Svensson, E. Lind, and L. E. Wernersson, “Vertical heterojunction InAs/InGaAs nanowire MOSFETs on Si with I_{on} = 330 µA/µm at Ioff = 100 nA/µm and V_D = 0.5 V,” in *2017 Symposium on VLSI Technology*, 2017, pp. T36–T37.

[55] A. Vardi, X. Zhao, and J. A. del Alamo, “Quantum-size effects in sub 10-nm fin width InGaAs FinFETs,” in *2015 IEEE International Electron Devices Meeting (IEDM)*, 2015, pp. 31.3.1-31.3.4.

[56] “Dry Etching of Electronic Oxides, Polymers, and Semiconductors - Pearton - 2005 - Plasma Processes and Polymers - Wiley Online Library.” [Online]. Available: https://onlinelibrary.wiley.com/doi/full/10.1002/ppap.200400035. [Accessed: 17-Apr-2018].

[57] E. W. Berg and S. W. Pang, “Electrical and optical characteristics of etch induced damage in InGaAs,” *J. Vac. Sci. Technol. B Microelectron. Nanometer Struct. Process. Meas. Phenom.*, vol. 16, no. 6, pp. 3359–3363, Nov. 1998.

[58] S. W. Pang, “Surface damage on GaAs induced by reactive ion etching and sputter etching,” *J. Electrochem. Soc.*, vol. 133, no. 4, pp. 784–787, Apr. 1986.

[59] S. Murad, M. Rahman, N. Johnson, S. Thoms, S. P. Beaumont, and C. D. W. Wilkinson, “Dry etching damage in III–V semiconductors,” *J. Vac. Sci. Technol. B Microelectron. Nanometer Struct. Process. Meas. Phenom.*, vol. 14, no. 6, pp. 3658–3662, Nov. 1996.

[60] P. Hashemi, J. T. Teherani, and J. L. Hoyt, “Investigation of hole mobility in gate-all-around Si nanowire p-MOSFETs with high-k/metal-gate: Effects of hydrogen thermal annealing and nanowire shape,” in *2010 International Electron Devices Meeting*, 2010, pp. 34.5.1-34.5.4.

[61] J. Yamamoto, T. Kawasaki, H. Sakaue, S. Shingubara, and Y. Horiike, “Digital etching study and fabrication of fine Si lines and dots,” *Thin Solid Films*, vol. 225, no. 1, pp. 124–129, Mar. 1993.

[62] K. J. Kanarik *et al.*, “Overview of atomic layer etching in the semiconductor industry,” *J. Vac. Sci. Technol. A*, vol. 33, no. 2, p. 020802, Mar. 2015.

[63] G. C. DeSalvo *et al.*, “Wet chemical digital etching of GaAs at room temperature,” *J. Electrochem. Soc.*, vol. 143, no. 11, pp. 3652–3656, Nov. 1996.

[64] A. Alian, C. Merckling, G. Brammertz, M. Meuris, M. Heyns, and K. D. Meyer, “InGaAs MOS transistors fabricated through a digital-etch gate-recess process and the influence of forming gas anneal on their electrical behavior,” *ECS J. Solid State Sci. Technol.*, vol. 1, no. 6, pp. P310–P314, Jan. 2012.

[65] S. Lee *et al.*, “High transconductance surface channel In0.53Ga0.47As MOSFETs using MBE source-drain regrowth and surface digital etching,” in *2013 International Conference on Indium Phosphide and Related Materials (IPRM)*, 2013, pp. 1–2.

[66] S. Lee *et al.*, “High performance raised source/drain InAs/In0.53Ga0.47As channel metal-oxide-semiconductor field-effect-transistors with reduced leakage using a vertical spacer,” *Appl. Phys. Lett.*, vol. 103, no. 23, p. 233503, Dec. 2013.

[67] J. Lin, X. Zhao, D. A. Antoniadis, and J. A. del Alamo, “A novel digital etch technique for deeply scaled III-V MOSFETs,” *IEEE Electron Device Lett.*, vol. 35, no. 4, pp. 440–442, Apr. 2014.

[68] C. Y. Huang *et al.*, “12 nm-gate-length ultrathin-body InGaAs/InAs MOSFETs with $8.3 \cdot 10^5$ I_{ON}/I_{OFF},” in *2015 73rd Annual Device Research Conference (DRC)*, 2015, pp. 260–260.

[69] X. Zhao and J. A. d. Alamo, “Nanometer-scale vertical-sidewall reactive ion etching of InGaAs for 3-D III-V MOSFETs,” *IEEE Electron Device Lett.*, vol. Early Access Online, 2014.

[70] A. Vardi and J. A. del Alamo, “Sub-10-nm fin-width self-aligned InGaAs FinFETs,” *IEEE Electron Device Lett.*, vol. 37, no. 9, pp. 1104–1107, Sep. 2016.

[71] N. Waldron *et al.*, “Gate-all-around InGaAs nanowire FETS with peak transconductance of 2200 µS/µm at 50nm L_g using a replacement Fin RMG flow,” in *2015 IEEE International Electron Devices Meeting (IEDM)*, 2015, pp. 31.1.1-31.1.4.

[72] X. Zhao, A. Vardi, and J. A. del Alamo, “InGaAs/InAs heterojunction vertical nanowire tunnel FETs fabricated by a top-down approach,” in *2014 IEEE International Electron Devices Meeting*, 2014, pp. 25.5.1-25.5.4.

[73] F. Lukeš, “Oxidation of Si and GaAs in air at room temperature,” *Surf. Sci.*, vol. 30, no. 1, pp. 91–100, Mar. 1972.

[74] Z. Y. Liu, T. F. Kuech, and D. A. Saulys, “A comparative study of GaSb (100) surface passivation by aqueous and nonaqueous solutions,” *Appl. Phys. Lett.*, vol. 83, no. 13, pp. 2587–2589, Sep. 2003.

[75] A. Nainani *et al.*, “Device quality Sb-based compound semiconductor surface: A comparative study of chemical cleaning,” *J. Appl. Phys.*, vol. 109, no. 11, p. 114908, Jun. 2011.

[76] Z. Y. Liu, B. Hawkins, and T. F. Kuech, “Chemical and structural characterization of GaSb(100) surfaces treated by HCl-based solutions and annealed in vacuum,” *J. Vac. Sci. Technol. B*, vol. 21, no. 1, pp. 71–77, Jan. 2003.

[77] C. J. Vineis, C. A. Wang, and K. F. Jensen, “In-situ reflectance monitoring of GaSb substrate oxide desorption,” *J. Cryst. Growth*, vol. 225, no. 2–4, pp. 420–425, May 2001.

[78] A. Ali *et al.*, “Fermi level unpinning of GaSb (100) using plasma enhanced atomic layer deposition of Al2O3,” *Appl. Phys. Lett.*, vol. 97, no. 14, p. 143502, Oct. 2010.

[79] M. V. Lebedev, E. V. Kunitsyna, W. Calvet, T. Mayer, and W. Jaegermann, “Sulfur Passivation of GaSb(100) Surfaces: Comparison of Aqueous and Alcoholic Sulfide Solutions Using Synchrotron Radiation Photoemission Spectroscopy,” *J. Phys. Chem. C*, vol. 117, no. 31, pp. 15996–16004, Aug. 2013.

[80] L. Zhao, Z. Tan, J. Wang, and J. Xu, “Effects of ozone post deposition treatment on interfacial and electrical characteristics of atomic-layer-deposited Al_2O_3 and HfO_2 films on GaSb substrates,” *Appl. Surf. Sci.*, vol. 289, pp. 601–605, Jan. 2014.

[81] C. Thelander *et al.*, “Development of a Vertical Wrap-Gated InAs FET,” *IEEE Trans. Electron Devices*, vol. 55, no. 11, pp. 3030–3036, Nov. 2008.

[82] M. Heyns and W. Tsai, “Ultimate Scaling of CMOS Logic Devices with Ge and III–V Materials,” *MRS Bull.*, vol. 34, no. 07, pp. 485–492, 2009.

[83] “International Roadmap for Semiconductors,” 2012.

[84] Z. Zhang *et al.*, “Ultra Low Contact Resistivities for CMOS Beyond 10-nm Node,” *IEEE Electron Device Lett.*, vol. 34, no. 6, pp. 723–725, Jun. 2013.

[85] H. Yu *et al.*, “$1.5 \cdot 10^9$ $\Omega \cdot cm^2$ contact resistivity on highly doped Si:P using Ge pre-amorphization and Ti silicidation,” in *2015 IEEE International Electron Devices Meeting (IEDM)*, 2015, pp. 21.7.1-21.7.4.

[86] J. C. Lin, S. Y. Yu, and S. E. Mohney, “Characterization of low-resistance ohmic contacts to n- and p-type InGaAs,” *J. Appl. Phys.*, vol. 114, no. 4, pp. 044504-044504–8, 2013.

[87] R. Dormaier and S. E. Mohney, “Factors controlling the resistance of Ohmic contacts to n-InGaAs,” *J. Vac. Sci. Technol. B Microelectron. Nanometer Struct.*, vol. 30, no. 3, p. 031209, 2012.

[88] A. Baraskar *et al.*, “Ex situ Ohmic contacts to n-InGaAs,” *J. Vac. Sci. Technol. B Microelectron. Nanometer Struct.*, vol. 28, no. 4, p. C5I7-C5I9, 2010.

[89] A. Baraskar, A. C. Gossard, and M. J. W. Rodwell, “Lower limits to metal-semiconductor contact resistance: Theoretical models and experimental data,” *J. Appl. Phys.*, vol. 114, no. 15, p. 154516, Oct. 2013.

[90] W. Lu, “A test structure to characterize nano-scale ohmic contacts in III-V MOSFETs,” *Electron. Device Lett.*, vol. 35, no. 2, 2014.

[91] E. M. Lysczek, S. E. Mohney, and T. N. Wittberg, “Shallow ohmic contacts to p-type InAs,” *Electron. Lett.*, vol. 39, no. 25, pp. 1866–1868, Dec. 2003.

[92] R. Dormaier *et al.*, “Thermal stability of Pd/Pt/Au Ohmic contacts to InAlSb/InAs heterostructures for high electron mobility transistors,” *J. Appl. Phys.*, vol. 105, no. 4, pp. 044505-044505–8, Feb. 2009.

[93] J. A. del Alamo and T. Mizutani, “Rapid thermal annealing of InP using GaAs and InP proximity caps,” *J. Appl. Phys.*, vol. 62, no. 8, pp. 3456–3458, Oct. 1987.

[94] L. A. Walsh *et al.*, “Ni-(In,Ga)As Alloy Formation Investigated by Hard-X-Ray Photoelectron Spectroscopy and X-Ray Absorption Spectroscopy,” *Phys. Rev. Appl.*, vol. 2, no. 6, p. 064010, Dec. 2014.

[95] L. Czornomaz *et al.*, “CMOS compatible self-aligned S/D regions for implant-free InGaAs MOSFETs,” *Solid-State Electron.*, vol. 74, pp. 71–76, Aug. 2012.

[96] R. Oxland *et al.*, “An ultralow-resistance ultrashallow metallic source/drain contact scheme for III-V NMOS,” *IEEE Electron Device Lett.*, vol. 33, no. 4, pp. 501–503, Apr. 2012.

[97] W. Lu, J. K. Kim, J. F. Klem, S. D. Hawkins, and J. A. del Alamo, “An InGaSb p-channel FinFET,” in *2015 IEEE International Electron Devices Meeting (IEDM)*, 2015, pp. 31.6.1-31.6.4.

[98] J. Lin, D. A. Antoniadis, and J. A. del Alamo, “Sub-30 nm InAs quantum-well MOSFETs with self-aligned metal contacts and Sub-1 nm EOT HfO_2 insulator,” in *Electron Devices Meeting (IEDM), 2012 IEEE International*, 2012, pp. 32.1.1-32.1.4.

[99] L. W. Guo, L. Xia, B. R. Bennett, J. B. Boos, M. G. Ancona, and J. A. del Alamo, “Enhancing p-channel InGaSb QW-FETs via process-induced compressive uniaxial Strain,” *IEEE Electron Device Lett.*, vol. 35, no. 11, pp. 1088–1090, Nov. 2014.

[100] C. Nguyen, B. Brar, and H. Kroemer, “Surface-layer modulation of electron concentrations in InAs–AlSb quantum wells,” *J. Vac. Sci. Technol. B Microelectron. Nanometer Struct. Process. Meas. Phenom.*, vol. 11, no. 4, pp. 1706–1709, Jul. 1993.

[101] B. R. Bennett, A. A. Podpirka, J. B. Boos, and S. L. Kumar, “Strained InGaSb/AlGa(As)Sb Quantum Wells for p-Channel Transistors,” *J. Electron. Mater.*, vol. 45, no. 6, pp. 2757–2762, Jun. 2016.

[102] M. Losurdo, P. Capezzuto, G. Bruno, A. S. Brown, T. Brown, and G. May, “Fundamental reactions controlling anion exchange during mixed anion heterojunction formation: Chemistry of As-for-Sb and Sb-for-As exchange reactions,” *J. Appl. Phys.*, vol. 100, no. 1, p. 013531, Jul. 2006.

[103] B. R. Bennett, B. V. Shanabrook, and M. E. Twigg, “Anion control in molecular beam epitaxy of mixed As/Sb III-V heterostructures,” *J. Appl. Phys.*, vol. 85, no. 4, pp. 2157–2161, Feb. 1999.

[104] W. S. Hwang, J. Chen, W. J. Yoo, and V. Bliznetsov, “Investigation of etching properties of metal nitride/high-k gate stacks using inductively coupled plasma,” *J. Vac. Sci. Technol. Vac. Surf. Films*, vol. 23, no. 4, pp. 964–970, Jun. 2005.

[105] J. Chen, W. J. Yoo, Z. Y. Tan, Y. Wang, and D. S. H. Chan, “Investigation of etching properties of HfO based high-K dielectrics using inductively coupled plasma,” *J. Vac. Sci. Technol. Vac. Surf. Films*, vol. 22, no. 4, pp. 1552–1558, Jul. 2004.

[106] A. Ortiz-Conde, F. J. García Sánchez, J. J. Liou, A. Cerdeira, M. Estrada, and Y. Yue, “A review of recent MOSFET threshold voltage extraction methods,” *Microelectron. Reliab.*, vol. 42, no. 4, pp. 583–596, Apr. 2002.

[107] K. M. Persson, E. Lind, A. W. Dey, C. Thelander, H. Sjöland, and L. E. Wernersson, "Low-frequency noise in vertical InAs nanowire FETs," *IEEE Electron Device Lett.*, vol. 31, no. 5, pp. 428–430, May 2010.

[108] N. Conrad *et al.*, "Low-frequency noise and RTN on near-ballistic III-V GAA nanowire MOSFETs," in *2014 IEEE International Electron Devices Meeting*, 2014, pp. 20.1.1-20.1.4.

[109] J. Kedzierski *et al.*, "Extension and source/drain design for high-performance FinFET devices," *IEEE Trans. Electron Devices*, vol. 50, no. 4, pp. 952–958, Apr. 2003.

[110] M. Jurczak, N. Collaert, A. Veloso, T. Hoffmann, and S. Biesemans, "Review of FINFET technology," in *2009 IEEE International SOI Conference*, 2009, pp. 1–4.

[111] M. J. H. van Dal *et al.*, "Highly manufacturable FinFETs with sub-10nm fin width and high aspect ratio fabricated with immersion lithography," in *2007 IEEE Symposium on VLSI Technology*, 2007, pp. 110–111.

[112] A. Vardi, W. Lu, X. Zhao, and J. A. del Alamo, "Nanoscale Mo ohmic contacts to III-V fins," *IEEE Electron Device Lett.*, vol. 36, no. 2, pp. 126–128, Feb. 2015.

[113] J. B. Boos *et al.*, "High mobility p-channel HFETs using strained Sb-based materials," *Electron. Lett.*, vol. 43, no. 15, pp. 834–835, Jul. 2007.

[114] K. Takei *et al.*, "Nanoscale InGaSb heterostructure membranes on Si substrates for high hole mobility transistors," *Nano Lett.*, vol. 12, no. 4, pp. 2060–2066, Apr. 2012.

[115] J. B. Boos *et al.*, "Sb-based n- and p-channel heterostructure FETs for high-speed, low-power applications," *IEICE Trans. Electron.*, vol. E91.C, no. 7, pp. 1050–1057, 2008.

[116] J. F. Klem, J. A. Lott, J. E. Schirber, S. R. Kurtz, and S. Y. Lin, "Strained quantum well modulation-doped InGaSb/AlGaSb structures grown by molecular beam epitaxy," *J. Electron. Mater.*, vol. 22, no. 3, pp. 315–318, Mar. 1993.

[117] C.-H. Yu, H.-T. Hsu, C.-Y. Chiang, C.-I. Kuo, Y. Miyamoto, and E. Y. Chang, "Performance evaluation of InGaSb/AlSb p-channel high-hole-mobility transistor fabricated using BCl_3 dry etching," *Jpn. J. Appl. Phys.*, vol. 52, no. 2R, p. 020203, Feb. 2013.

[118] P. Chang, X. Liu, G. Du, and X. Zhang, “Assessment of hole mobility in strained InSb, GaSb and InGaSb based ultra-thin body pMOSFETs with different surface orientations,” in *Electron Devices Meeting (IEDM), 2014 IEEE International*, 2014, pp. 7.7.1-7.7.4.