Department of Electrical and Information Technology

Fabrication and Characterization of Quantum-Well Field-Effect Transistors

Author: Yang Fu

A thesis submitted for the degree of Master of Science

30hp, Sep. 2021-May 2022

Supervisors: Erik Lind and Rainer Timm

Lund University

May 2022

Acknowledgment

I want to thank both of my supervisors, Erik Lind and Rainer Timm who supervised my thesis for the last two semesters. They have kindly helped me in every part of this project work from the beginning to the end. I would also like thank to Patrik Olausson, the Ph.D. student I worked with. He taught me how to fabricate a transistor from the beginning and guided me on how to analyze the data we got from probe measurement. Thanks for you opening the door of a fantastic area for me. It was a wonderful experience that I could tune the electrons flow in the channel using gate bias of FET by different designs and fabrication technologies. I also learnt the attitudes toward scientific research and persistence after many failed attempts from them. What I learned from you will benefit me for the rest of my life.

Content

Abstract

The project aims to optimize the design and fabrication of InGaAs quantum-well field-effect transistor (QW-FET) by investigating transfer and output characteristics of the QW-FET. This work found a lower source/drain contact resistance solution starting with fabricating micrometer-level gate length transistors. It is a multiple-layer design, n-type InGaAs and n-type InP over channel. The stack of metal-oxide-semiconductor (MOS) was studied as well, and the lower subthreshold swing (SS) indicated that oxide layer contact with channel is a better way to control gate electrostatic. To precise the data (contact resistance, sheet resistance and electron mobility) derived from the ON-state resistance and transconductance in transistors, short-gate QW-FET, less than 1 micron, processed using electron-beam lithography technology. Although leakage occurred, it is feasible to attain a promising transistor with high electron mobility and low contact resistance. Moreover, Post Metallization Annealing (PMA) and measurement temperature impacts on transistor performance were also assessed in this work. The evidence shows that the performance parameters of transistors were upgraded after PMA 350℃ in N_2/H_2 for five minutes. The reason is related to a lower density of defect at the interface of channel and high-κ layer after annealing, and thus the chances of electrons scattering by the Coulomb are reduced. However, the resistance results of transmission line model (TLM) did not obviously decrease. By testing transistors at 13K, doubled electron mobility was found, and the dropping of SS also indicated a boost of gate electrostatic control at low temperature.

Acronyms

MOSFET - Metal Oxide Semiconductor Field Effect Transistor
CMOS - complementary metal-oxide-semiconductor
ALD - atomic layer deposition
HEMTS - high-electron-mobility transistors
MOVPE - metalorganic vapor phase epitaxy
PDMA - polydimethylsiloxane
QW-MOSFET - quantum-well metal-oxide-semiconductor field-effect transistor
MOS capacitor - metal oxide semiconductor capacitor
SS - Subthreshold swing
MOSFET - metal oxide semiconductor FET
TLM - transmission line model
FET - Field-effect transistor
TEM - Transmission Electron Microscopes
TMA - trimethylaluminum
PMA - post metallization annealing
PMMA - polymethyl methacrylate based photoresist
EBL - electron beaming lithography
SEM - Scanning Electron Microscope
DOE - design of experiment

Chapter 1 Introduction

1.1 Conventional MOSFET

Metal Oxide Semiconductor Field Effect Transistor (MOSFET), with three terminals source, drain and gate, is a fundamental semiconductor component widely used for switching and amplifying electronic signals in electronic industries. In the late 1960s, the invention of silicon-gate MOSFET IC by Fairchild R&D opened the door to developing the semiconductor integrated circuit industry.

Two complementary MOSFETs can construct a complementary metal-oxide-semiconductor (CMOS) inverter, so the switch is a wide MOSFET application in conventional computing architecture. CMOS dominates the logic family, as one of the transistors always remaining OFF in both logic states leads to reduced power consumption [1]. Most MOSFETs are built on silicon wafers. The advantage of silicon is that SiO_2 naturally forms on the surface of Si, which is called native oxide layer. Consequently, it is easier to produce a MOS structure with a nearly perfect interface between the native oxide layer and silicon [2].

1.2 III-V Compound Semiconductor

Transistor scaling is the goal of scientists and engineers, which requires scaling down the size of transistors and maintaining the power density simultaneously. However, silicon transistor scaling is now facing many challenges, including excessive leakage currents through the SiO_2 gate oxide, exponentially increasing leakage currents between transistor source and drain, increasing source-drain access resistance, carrier-mobility degradation in the transistor channel with electric field increment, as well as increasing device-to-device variation [3]. Besides, the advance in artificial intelligence and machine learning exerts pressure on the demand for computational capabilities. All the above factors are threatening the revolution of nanoelectronics industries.

There are several solutions to address the problem. One of them is turning to a new channel material instead of silicon: group III-V compound semiconductor [4-7]. Although III-V elements cannot produce a native oxide layer with lower defects density like silicon, with the advent of atomic layer deposition (ALD) technology, a high-quality ALD oxide layer can be deposited on the III–V surface. It paves the way for the mass manufacture of III–V MOSFETs. As a result, the mature manufacturing process of III-V is only second to silicon in the semiconductor industry. More Moore report in 2017 also demonstrated III-V compound semiconductor is a promising device in the future as shown in Figure 1.

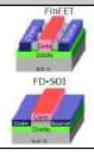
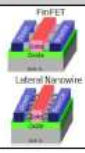
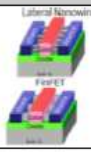
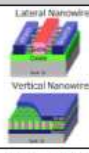
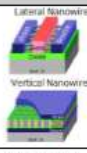
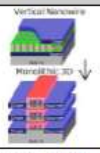
Year	2017	2019	2021	2024	2027	2030
Device Structures						
Channel	Si	SiGe25%	SiGe50%	Ge, III-V, TFET, 2D	Ge, III-V, TFET, 2D mat	Ge, III-V, TFET, 2D mat

Figure 1 projected transistor development trend (courtesy of C. Convertino IBM) [8].

III-V compounds are attractive channel material for future advanced MOSFETs due to their higher electron mobility in contrast with silicon. It makes a continuous reduction in voltage without losing performance possible. From Figure 2, the electron mobility of group III-V compound semiconductors is more than ten times larger than silicon, which indicates III-V compound is a potential candidate for N-channel MOSFET. The excellent transfer performance of these III-V materials will help us achieve many high-speed and high-frequency electronic systems. For instance, some works reported that cutoff frequencies of current gain and power gain could approach 600 GHz and 1 THz on InGaAs-based high-electron-mobility transistors (HEMTs), respectively [9–11]. In addition, researchers have already developed several methods to integrate III-V compound transistor channel materials on the Si platform. For example, III-V nanowire growth on silicon by metalorganic vapor phase epitaxy (MOVPE) [12], a self-aligned gate-last process using molecular beam epitaxy (MBE) [7], heterogeneous assembly through transferring free-standing nanoribbon active layers onto Si with polydimethylsiloxane (PDMS) slab [6].

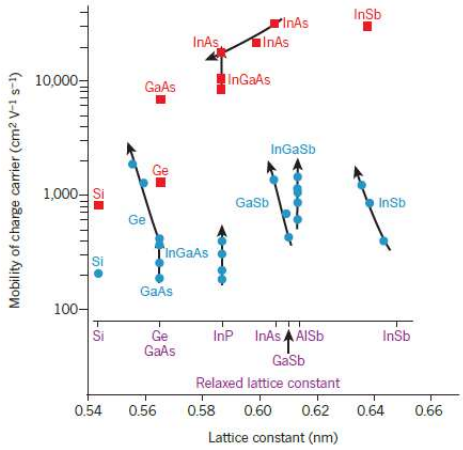


Figure 2 Electron and hole mobility of Si and III–V compound semiconductors to lattice constant electrons (red) and holes (blue) [13].

1.3 InGaAs Quantum-well FET

Quantum-well MOSFET (QW-MOSFET) is a transistor with the channel material of quantum well extending under the raised source and drain regions, and they are cost-saving manufacture and process simplification in comparison to conventional CMOS.

It is also a prospective transistor exhibiting a trade-off electron transport, electrostatic integrity and parasitic resistance together.

$In_xGa_{1-x}As$ QW-MOSFET has high carrier mobility under lower drive voltage (around 0.5V) and can be used in future high-speed and low-power logic applications. In this diploma work, $In_{0.53}Ga_{0.47}As$ QW-MOSFET was fabricated in Nano Lund cleanroom, and the transfer and output characteristics were measured in the Electronics and Information Technology lab. The aim of this project is to optimize fabrication parameters and investigate the contact resistance, sheet resistance, carrier mobility and subthreshold swing of a series of transistors. Our expectation is to find a low power-consumption (low contact resistance and sheet resistance) and high efficiency (high carrier mobility and low subthreshold swing) transistor.

Chapter 2 Theory

This chapter explains the mechanism of quantum-well field-effect transistor from metal oxide semiconductor capacitor. Four parameters studied in the thesis were also introduced in this chapter, which are subthreshold swing, contact resistance, sheet resistance and carrier mobility. However, the four parameters cannot directly be calculated from the transfer characteristic of transistor; they can only be derived from "ON state" resistance and transconductance. Finally, we mentioned how to derive four crucial transistor parameters from g_m and R_{on}.

2.1 Metal-Oxide-Semiconductor (MOS) capacitor

A metal-oxide-semiconductor capacitor is a vital part of MOSFET because it controls the carrier flow in the channel layer. The advantage of this kind of stack provides isolation by preventing the direct flow of charges from gate to channel.
In this project, the semiconductor material in the channel is un-doped $In_xGa_{1-x}As$. Two different wafers were used, one is 10nm $In_{0.53}Ga_{0.47}As$ and the other one is 3nm $In_{0.53}Ga_{0.47}As$+5nm $In_{0.8}Ga_{0.2}As$+4nm $In_{0.53}Ga_{0.47}As$ (the materials are simplified to InGaAs in the following description, the different composition on the two wafters is due to the fact that a first wafer was used up during the hard mask etching process and we bought the second wafer from another supplier). Figure 3 illustrates the band diagram of the ideal MOS capacitor under zero bias between metal and semiconductor. In the figure, $q\varphi$ is work function represents the energy difference between the Fermi level and the vacuum level ($q\varphi_m$ – work function of metal and $q\varphi_s$ work function semiconductor). $q\chi$ is electron affinity defined as the energy difference between the conduction band edge and the vacuum level in the semiconductor, $q\chi_i$ is the oxide electron affinity. The energy barrier between the metal and the oxide is $q\varphi_B$.
When no voltage is applied between metal and semiconductor, the work function difference between metal and semiconductor $q\varphi_{ms}$ is zero in ideal situation,

$$q\varphi_{ms} \equiv q\varphi_m - q\varphi_s = q\varphi_m - \left(q\chi + \frac{E_g}{2}\right) = 0 \quad (1)$$

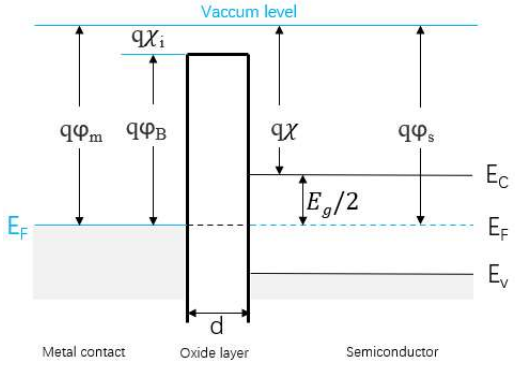
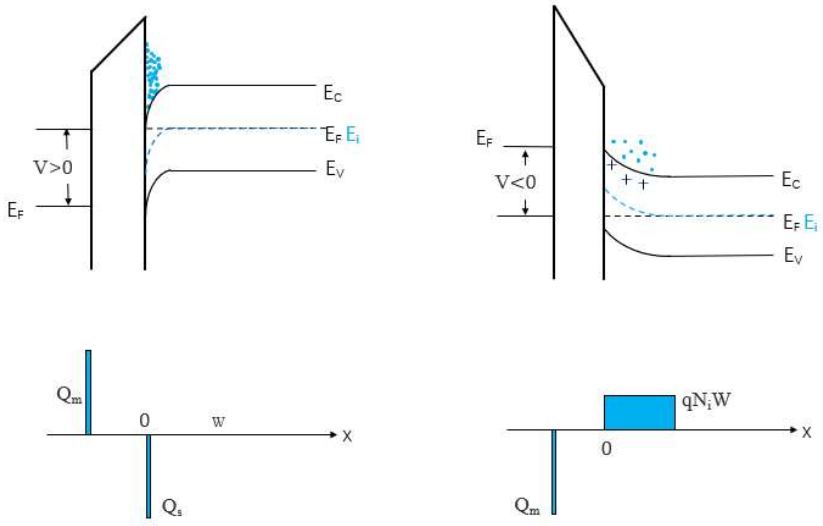


Figure 3 Energy band diagram of an ideal MOS capacitor at V_{bias}=0 V (adapted from [14]).

When a positive bias is applied on the metal contact, negative carriers (electrons) accumulate at the interface of oxide and semiconductor, and the conduction and valence bands near the semiconductor surface are bent downward. If the voltage is higher than the threshold voltage, abundant electrons tend to concentrate near the interface (see the blue dots in Figure 4 (a)) and form a space charge layer filled with electrons. Considering MOS capacitor is ideal, the negative charge distribution per unit area Q_s in the semiconductor is equal to the positive charge per unit area Q_m in the metal ($Q_m = |Q_s|$).

When a small negative voltage ($V<0$) is applied to the metal contact, the energy bands near the semiconductor surface change upward, and the electrons in InGaAs are depleted away from the surface (Figure 4 (b)) and distributed in a space named depletion region. The space charge can be described as qN_iW (N_i-intrinsic carrier concentration), where W is the width of the surface depletion region. Similarly, ($|Q_m| = qN_iW$), but the charge carriers are positive in the semiconductor this time.



(a) (b)

Figure 4 Energy band diagrams and charge distributions of an ideal MOS capacitor (adapted from [14]) (a) electrons accumulation under positive bias (b) electrons depletion under negative bias.

The amount of charge per unit is equal and charge carriers are electrically opposite on metal and semiconductor side, which is based on a presupposition that the MOS capacitor is ideal. Actually, many defects exist, like charge defects in the oxide and trap defects at the oxide and semiconductor interface as shown in Figure 5 (take the example of metal-SiO_2-Si MOS). The performance of the MOS capacitor is significantly affected by the number of defects.

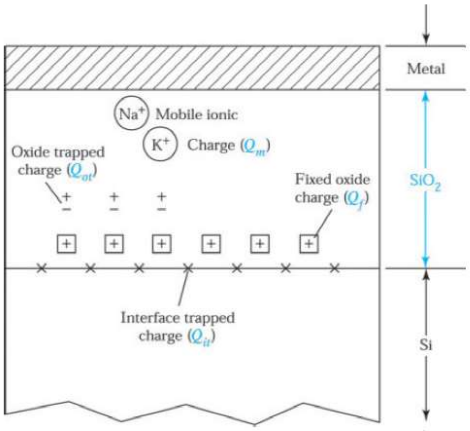


Figure 5 Traps and charges in the metal-SiO_2-Si MOS, Q_{it} the interface trapped charge, Q_i fixed-oxide charge, Q_{ot} oxide-trapped charge, and Q_m mobile ionic charge [14].

Interface traps are a general term for defects at the interface between the semiconductor and the gate oxide. One of the most common interface traps is the dislocation at the interface because the defects interrupt the periodic lattice structure due to different atomic and crystal structures. A schematic picture indicates the interface defects between HfO_2 and InAs in Figure 6. In fact, the MOS capacitor (HfO_2+Al_2O_3+InGaAs) has a more complicated interface and bond structure than Figure 6 in our work. In other words, interface traps are electrically active defects capable of trapping and de-trapping charge carriers at the interface [14], so interface traps play a negative contribution to carrier mobility since the Coulomb scattering increases. In conclusion, forming a high-quality interface with a low density of interface trap (D_{it}) on a channel material is necessary to attain a transistor with steep subthreshold slope, high carrier mobility, and good device reliability.

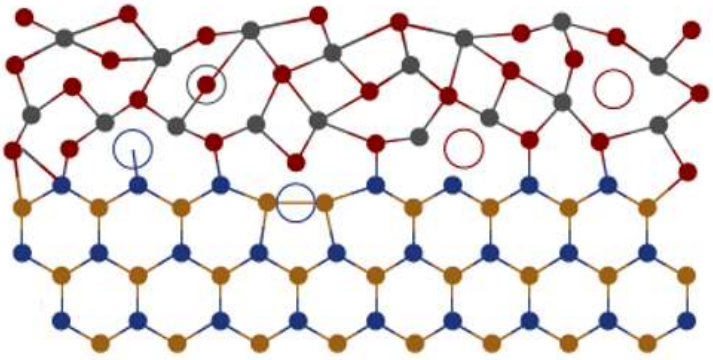


Figure 6 The schematic of the chemical bond structure of HfO_2 and InAs (Hf-gray dot, O-red dot, In-yellow dot and As-blue dot), defects were marked by circles, for instance, oxygen vacancies (red circles), oxygen interstitial (grey circles), and As-As and a dangling bond (blue circles) [15].

A vital factor is the capacitance of the MOS capacitor (gate conductance) for a short-channel transistor. It determined the cutoff frequency of the device, and the higher the gate capacitance, the fast switch on/off in the transistor. The gate capacitance C_g can

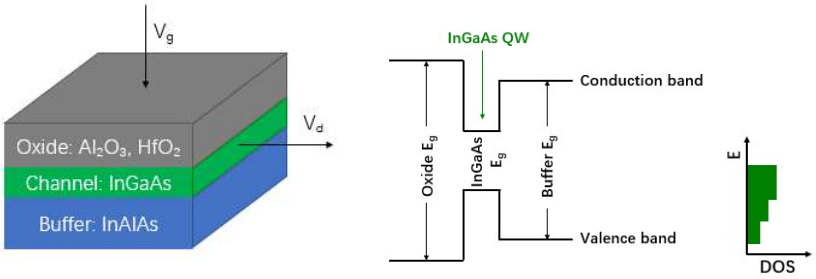
be calculated by

$$C_g = C_{g,intrinsic} + C_{g,parasitic} \approx \frac{\varepsilon_{barrier} L_g}{t_{barrier}} + C_{g,parasitic} \quad (2)$$

Here, $\varepsilon_{barrier}$ is the dielectric constant of the barrier (oxide layer) between channel and gate contact, $t_{barrier}$ is the thickness of the barrier, and L_g is the gate length. A large $\varepsilon_{barrier}/t_{barrier}$ ratio is preferred for reducing the influence of parasitic capacitances [16]. As a result, a high dielectric constant and thin oxide material is necessary.

2.2 Quantum-well Field Effect Transistor

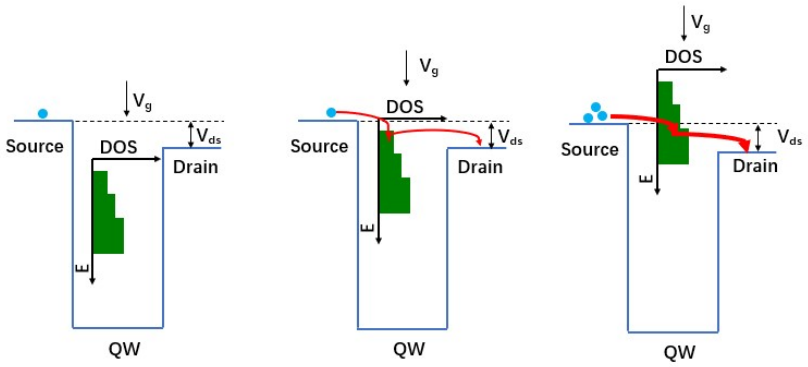
This study utilized 10-12nm un-doped InGaAs as a channel material between the gate oxide and buffer layer as shown in Figure 7 (a). Energy barriers are formed because the bandgap is higher in the oxide and buffer layer than in InGaAs. Due to quantum effects on the thin channel layer, quantum-well forms in the two-dimensional InGaAs sheet. The electron density of state (DOS) does not exponentially increase with energy like in bulk matter, but follows the step shape with the energy increase (see Figure 7 (b)) [17].



(a) (b)

Figure 7 Schematic of the quantum well in a channel (a) MOS capacitor layout in transistor (b) band diagram and density of state in channel.

The gate voltage can tune the electron DOS in the quantum-well, so that transistor switching function can be realized when applying bias on gate, source and drain. Figure 8 describes the working principle of current control by gate voltage under a certain V_{ds} in a transistor. Figure 8 (a) no current condition: although potential difference exists between source and drain, the electron on the source side does not have states to fill in, so electrons cannot transfer to drain side. Figure 8 (b) and Figure 8 (c) are the conditions of low current and high current, which shows the amount of electron DOS within the range of source potential and drain potential. The DOS can be tuned by gate voltage, such that there are states in quantum well for electrons to fill in. As a result, the electrons pass through the quantum well and are transferred from source to drain to generate current. In another word, the magnitude of current is determined by how many electrons DOS in the quantum well can be used for electron transfer. There are two ways to control the magnitude of current, by tuning V_g under a certain V_{ds} or tuning V_{ds} under a certain V_g.



(a) (b) (c)

Figure 8 Schematic of electrons transport between source and drain under a certain V_{ds} condition. (a) no current (b) low current (c) high current.

2.3 Subthreshold Swing

Subthreshold swing (SS) is a critical parameter to reflect electrostatic control of the channel. It is defined as the change in gate voltage that must be applied to create a one-decade increase in the output current. The parameter of subthreshold slope is also mentioned in the study, which is reciprocal of subthreshold swing. Moreover, in metal-oxide-semiconductor FET (MOSFET), the current-switching process is limited by thermal injection of electrons over an energy barrier because the carriers obey Fermi-Dirac statistic distribution so that SS can be written as

$$SS = \underbrace{\frac{dV_g}{d\psi_s}}_{m}\underbrace{\frac{d\psi_s}{d(log_{10}I_d)}}_{n} \cong \left(1+\frac{C_d}{C_{ox}}\right)\frac{kT}{q}ln10 \rightarrow \frac{kT}{q}ln10 \quad (3)$$

Where the term m is the transistor body factor, and n is a factor that characterizes the change of the drain current with the surface potential, Ψ_S, reflecting the conduction mechanism in the channel. V_g is the gate voltage, I_d is the drain current, kT/q is the thermal voltage, and C_d and C_{ox} are the depletion and the oxide capacitances, respectively [18]. In the ideal condition, $C_d \rightarrow 0$ and $C_{ox} \rightarrow \infty$, the SS result of Equation (3) is proportional to temperature, and the limitation of SS under room temperature (300K) is 60mV/decade. Considering short-channel effects, it is impossible to obtain a transistor with subthreshold swing of 60mV/decade [19]. The studies on $In_xGa_{1-x}As$ QW-MOSFET showed the subthreshold swing ranges from 80 to 200 mV/decade at room temperature [20-23].

2.4 Contact Resistance R_c and Sheet Resistance R_{sh}

In this work, we calculated contact resistance (R_c) and sheet resistance (R_{sh}) from the transfer characteristic of transistor devices and transmission line model (TLM).
An excellent ohmic connection on the source/drain side can pass the required current

with a small voltage drop compared to the drop across the channel region of the transistor [14], which make sure no significant degradation of the transistor performance. Especially in a short-channel MOSFET, the source/drain series resistance cannot be negligible [14].
Figure 9 exhibits the simplified transistor schematic with metal source/drain directly contact with channel. The contact resistance can be written as:

$$R_c = \frac{R_{total} - R_{ch}}{2} \tag{4}$$

Here R_c is source or drain contact resistance, R_{ch} is channel resistance, so it is known that the value of R_c can be calculated through $R_{total}/2$ when channel length is infinitely close to zero, since $R_{ch} \approx 0\Omega$.

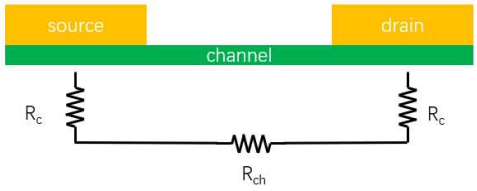


Figure 9: Schematic of series resistance in a simplified transistor.

Sheet resistance (R_{sh}) is generally used to characterize the resistance of two-dimensional material (thin film) made by doped semiconductor, metal deposition, resistive paste printing, and glass coating [24]. R_{sh} equation can be derived from the resistance Equation (5) of bulk materials.
In bulk conductors, resistance can be written as:

$$R = \rho \frac{L}{A} = \rho \frac{L}{Wt} \tag{5}$$

Here ρ is the resistivity and A is the cross-sectional area. L, W and t are the length, width and thickness of the conductor respectively.
Imaging the material is a sheet (two-dimensional material) as shown in Figure 10, Equation (5) can be written as:

$$R = \frac{\rho}{t}\frac{L}{W} = R_{sh}\frac{L}{W} \Rightarrow R_{sh} = \frac{R}{L/W} \tag{6}$$

According to Equation (6), R_{sh} can be derived from the slope of a linear regression fitting of R and L/W. In our work, the two-dimensional sheet is InGaAs channel and R_{sh} is approximately equivalent to R_{ch}.

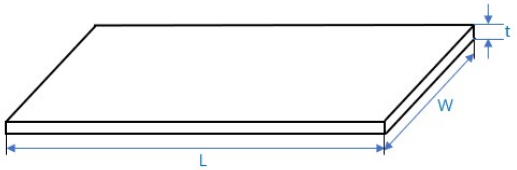


Figure 10 Schematical view of a two-dimensional (sheet) conductor.

2.5 Carrier Mobility

Carrier (electron) mobility determines the transfer performance of a transistor, such as operation speed and response frequency of a transistor.
Under thermal equilibrium at room temperature 300K, the electrons move randomly and promptly in semiconductor material in all three dimensions. Applying an electric field ε on channel, the thermal equilibrium condition is broken, and each electron will move along with the opposite direction of field direction under the force qε. The extra velocity driven by the electric field is called drift velocity. During the electrons drift in lattice, all momentum gained between all collision events is lost to the lattice during collisions in a steady state [14]. Based on the theorem of momentum, we have

$$q\varepsilon\tau_c = m_e v_{drift} \tag{7}$$

$$v_{drift} = \left(\frac{q\tau_c}{m_e}\right)\varepsilon = \mu_n \cdot \varepsilon \tag{8}$$

Here, m_e is effective mass of an electron and τ_c is the average time between one collision (called mean free time). Equation (7) states that the electron drift velocity is proportional to the applied electric field. The proportionality factor is defined as electron mobility μ_n and presents how strongly the motion of an electron is influenced by an applied electric field, the unit is with units of cm^2/Vs [14], so μ_n can be described as

$$\mu_n = \frac{q\tau_c}{m_e} \propto \tau_c \tag{9}$$

According to Equation (9), the electron mobility μ_n is proportional to the mean free time τ_c. Two major scattering mechanisms, lattice scattering and impurity scattering, in turn impact on τ_c. According to the equipartition theorem, each vibration atom in the crystal of semiconductor has an average energy in x-direction [25]:

$$E_x = \frac{p_x^2}{2M} + \frac{M}{2}\omega^2 x^2 = k_B T \tag{10}$$

where p is momentum, ω is oscillation angular frequency, and k_B is Boltzmann's constant, M is the mass of the atom. In Equation (10), the first term of the addition represents the kinetic energy of the atom and the second term of the addition represents the potential energy of the atom.
Setting the atomic vibration with amplitude x as the model of Figure 11 and combining with the Equation (10), it can be derived that the atomic vibration cover area is proportional to T ($S = \pi x^2 \propto T$). Additionally, the electrons have a thermal velocity $v_{th} \propto T^{1/2}$(under non-generate condition) according to average electron kinetic energy is given by

$$\frac{1}{2} m_e v_{th}^2 = \frac{3}{2} k_B T \tag{11}$$

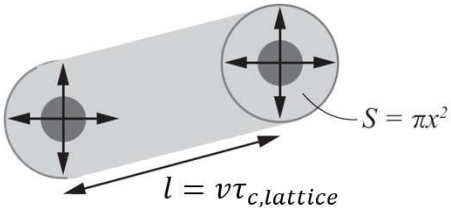


Figure 11: The path of an electron between two lattice scattering events [25].

Supposing the concentration of scatters is N_s and an electron transfer distance (mean free pass) $l = v_{th}\tau_{c,lattice}$ between two lattice scattering events, so one scattering occurs within the volume of the cylinder in Figure 11 is $Sv_{th}\tau_{c,lattice}N_s$ [25], so we can write,

$$\tau_{c,lattice} = \frac{1}{Sv_{th}N_s} \propto \frac{1}{T \times T^{\frac{1}{2}}} = T^{-3/2} \tag{12}$$

According to Equations (9) and (12), we have $\mu_{n,lattice}$ is proportional to $T^{-3/2}$. Figure 12 shows the general trend of carrier mobility with temperature T due to lattice scattering. Because the atoms in the crystal vibrate more intensively with temperature increase, the electrons have more chance to be scattered by the atoms during drifting, so that $\mu_{n,lattice}$ will decrease dramatically under higher temperature.

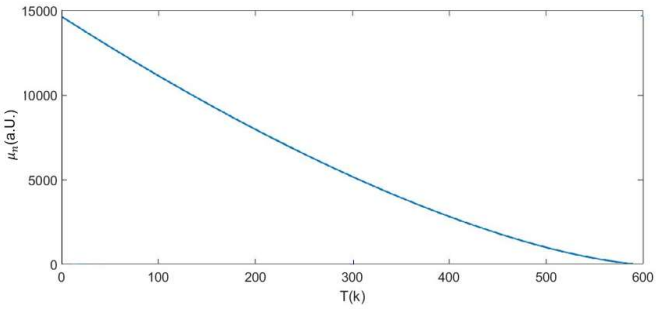


Figure12 $\mu_{n,lattice}$ and T curve

About the other mechanism, impurity scattering, a model is created like Figure 13 that the trajectory of an electron bends towards a positive ion. It can be seen that impurity scattering happens when the kinetic energy is smaller or equal to the potential energy.

$$E_{kin} \leq E_{pot} \Leftrightarrow \frac{3}{2}k_B T = \frac{e^2}{4\pi\epsilon_0\epsilon_r r_c} \tag{13}$$

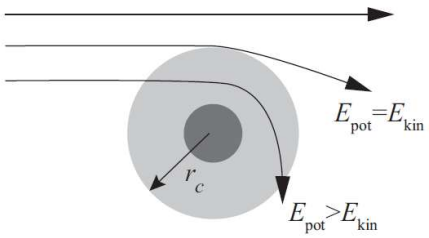


Figure 13 Scattering occurs if the electron is within a distance r_c from the ion. r_c is the critical distance at which the kinetic and potential energies are equal [25].

Since, the cross-section of impurity scattering is $S = \pi r_c^2 \propto T^{-2}$. Similarly for the lattice scattering mechanism, we assume that the impurity concentration is $N_{impurity}$ and an electron transfer distance $l = v_{th}\tau_{impurity}$ between two scattering events dues to impurities, in analogy to Equation (12), we have:

$$\tau_{impurity} = \frac{1}{S v_{th} N_{impurity}} \propto \frac{1}{T^{-2} T^{\frac{1}{2}}} = T^{3/2} \tag{14}$$

where $N_{impurity}$ is the concentration of ionized impurities. According to Equation (9), we know $\mu_{n,impurity} = \propto \tau_{impurity}$, consequently, $\mu_{n,impurity}$ is proportional with $T^{3/2}$ (see Figure 14), which indicates that, under higher temperature, the track of electrons with higher kinetic energy is difficult to be curved by Coulomb force exerted by impurities or dopants.

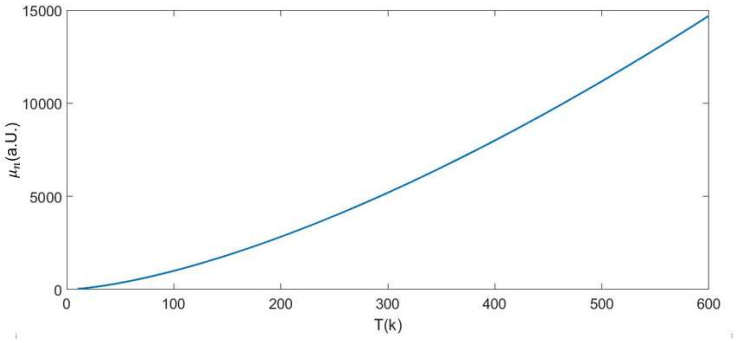


Figure 14 $\mu_{n,impurity}$ and T curve

Approximately, the sum of the number of collisions due to the two main scattering mechanisms is the total number of collisions events occurring in a unit of time [14]. It is given by two equations:

$$\frac{1}{\tau_c} = \frac{1}{\tau_{c,lattice}} + \frac{1}{\tau_{c,impurity}} \tag{15}$$

$$\text{and } \frac{1}{\mu_n} = \frac{1}{\mu_{lattice}} + \frac{1}{\mu_{impurity}} \tag{16}$$

Finally, whether the electron mobility is positively or negatively correlated to the temperature depends on which scattering mechanism dominated in the system, lattice scattering or impurity scattering.

2.6 Transconductance g_m and R_{on}

In this work, two methods were adopted to estimate the value of contact resistance, sheet resistance and carrier mobility, through transconductance g_m under V_{ds}=500mV and R_{on} under V_{ds}=50mV, because g_m should be calculated from the saturation region of transistor transfer characteristic (500mV) and R_{on} should be calculated from the linear region of transistor transfer characteristic (50mV).

Transconductance is the ratio of change in output current (between source and drain) and the change in gate voltage under certain V_{ds}. In general, the larger the

transconductance, the greater the gain (amplification) [26]. The formula can be written as

$$g_m = \frac{\Delta I_d}{\Delta V_g} \tag{17}$$

The contact resistance is relevant to transconductance and carrier mobility for a long channel gate FET. Through simulation, the g_m equation can be described as [27]

$$g_m = \frac{1}{R_c} - \frac{1}{R_c\sqrt{\frac{4K_n}{L_g}R_c(V_g - V_T)+1}} \tag{18}$$

where $K_n = \frac{C_{gg}\mu_n}{2}$ (C_{gg}=gate capacitance×width of gate).

A transistor as a switch has a "on/off" state. Although it is not a perfect short and open circuit. The resistance in linear area of transistor transfer characteristic is named R_{on}.

$$\mathrm{R_{on}} = \frac{\mathrm{V_d}}{\mathrm{I_d}}|_{V_d \to 0} \tag{19}$$

Contact resistance and carrier mobility can be also extracted from R_{on}, as there is a relation given by [28]:

$$\mathrm{R_{on}} \approx \frac{\mathrm{L}}{\mu_n W C_g(\mathrm{V_g} - \mathrm{V_{th}})} + 2R_c \tag{20}$$

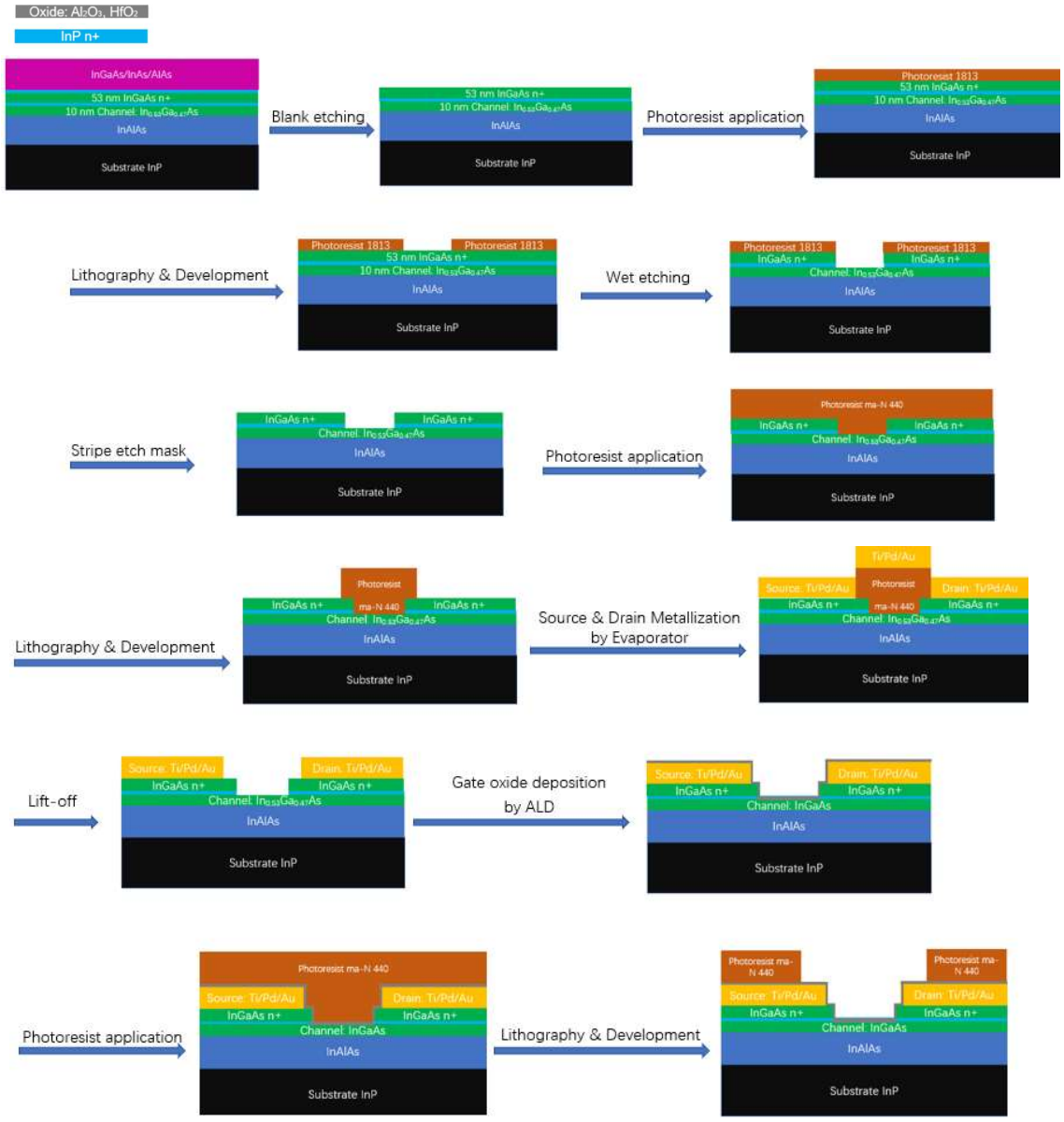
According to Equation 19, the parameters μ_n and R_c can be extracted from a linear regression fitting of experimental R_{on} and L (gate length).

Chapter 3 Design, Fabrication and Measurement

This chapter describes the process flow, main steps of the quantum-well FET fabrication process, measurement methods as well.

3.1 Process Flow

Figure 15 shows the process flow of FET44 sample C. The details of main steps will be disclosed in the following sections.



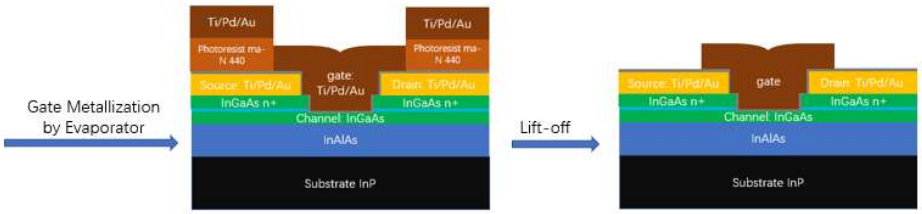


Figure 15 Schematic process flow of FET44 sample C

3.2 Blank Etching

In this work, two different wafters, IET3 and MOS1, were used to fabricate FET44 and FET53, on which epitaxial layers were grown in advance as shown in Figure 16. Only the IET3 substrate needs to do blank etching to remove the layer above InGaAs n+ (the purple and first blue layer in Figure 16 (a)), because the wafer has already pre-grown source and drain layers for another research. A H_3PO_4: H_2O_2: H_2O (1:1:25) solution and a HCl: H_2O (1:1) solution are used to etch the InGaAs/InAs/AlAs layers, with an etch rate of approximately 2 nm/s and InP with an etch rate of approximately 4 nm/s.

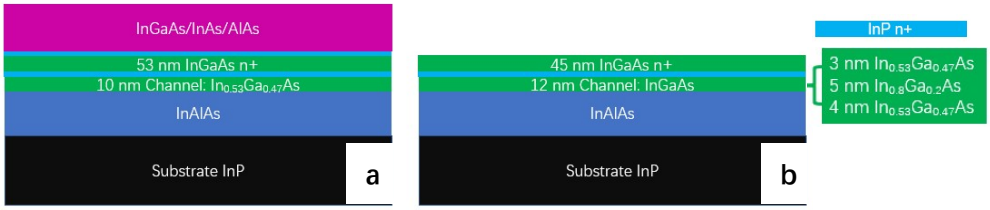


Figure 16 Schematic of cross-sectional wafer layout (a) IET3 (b) MOS1.

3.3 Mesa Etching

Generally, mesa etching is the first step of transistor fabrication. Locally semiconductor layers on the substrate, InP in this work, should be removed. The purpose is to remove all the layers on the substrate as the layout we designed, so that all the samples, transistor devices, TLM, reference and hall bar, could be insulated with each other. A positive photoresist S1813 was spun on and post-baked at 115℃ 90 seconds, forming a 1.2-1.4um photoresist film with good adhesion on the sample surface before exposure. The etchants are the same as previous blank etching step.

3.4 Metal Contact Pattern Definition

The negative photoresist ma-N 440 was utilized to define the metal contact on source, drain and gate areas. This photoresist can create a vertical profile and also has high thermal stability to endure in metal evaporation step. Post-baking at 95℃ and 3mins is needed before lithography for strengthen the adhesion between photoresist and sample surface. As a consequence, lateral etching creeping can be avoided during wet etching.

In the lithography steps, we selected a mask aligner equipment MJB4 using Hg lamp as UV source at 365nm wavelength, and the resolution of this machine is 2μm. A quartz mask must be used during exposure under MJB4, so the limited mask patterns and cleanliness on the quartz mask will affect the quality of generated patterns. However, the shortages of MJB4 are complemented by another aligner equipment MLA150. MLA150 is a maskless aligner that provides for flexible change of pattern, distortion compensation and other software corrections. The samples can be exposed under a laser source with 375nm and 405nm. The solution of the system can be down to 1 μm with an accuracy below 250 nm under optimized conditions [29].

3.5 Gate Oxide Deposition

The quality of the gate oxide layer determines the performance of a transistor. Finding a satisfactory oxide layer that helps to decrease charge trapping is crucial. Many research work is focused on Al_2O_3, HfO_2 material, and found that the oxide could largely remove the imperfect native oxides on III-V compounds [30, 31]. Al_2O_3 inter-layer is essential. TEM (Transmission Electron Microscopes) images demonstrated the roughness of the HfO_2/InGaAs interface is significantly improved when Al_2O_3 layer is added between HfO_2 and InGaAs[32]. The high-κ material used in this project is a bilayer Al_2O_3 and HfO_2 deposited by atomic layer deposition (ALD) at 350 ℃. ALD is a vapor phase technique used to grow thin films with atomic layer precision onto a substrate. The surface of a substrate is exposed to alternating precursors sequentially. Water and trimethylaluminum (TMA) as precursors were pulsed in the chamber and formed 1.2nm Al_2O_3 layer over the channel InGaAs, then 10nm HfO_2 layer was deposited using TDAHf and H_2O as precursors.

Before the deposition, passivation is a necessary step. It is reported sulphur surface bonds creation helped inhibit the native oxide formation during ammonium polysulfide treatment [33]. As an indium diffusion/segregation barrier, sulfur stabilized InGaAs surface and reduced the oxide trap density in Al_2O_3 deposition process as well [34]. In our work, the samples should soak in $(NH4)_2S_x$:H_2O 1:1 solution for 20 minutes, before that, the surface of InGaAs channel should be oxidized in a UV-Ozone cleaning system UVOH 150 for 10 minutes.

3.6 Pad Metallization

The metal selection in this step requires a strong bond with InGaAs and low contact resistance. Ti is the layer have excellent adhesion to InGaAs, and it is also reported that the contact resistance between Ti/Pd/Au and InGaAs can reach to 10^{-7} Ωcm^2 after annealing [35]. Finally, 5nm/5nm/300nm Ti/Pd/Au layers were deposited as metal contact on the pads of source, drain and gate by electron-beam evaporator.

Gate metallization can be deposited directly after high-κ formation. Nevertheless, source and drain pads need a pre-treatment, as InGaAs surface rapidly oxidizes when exposed to air. A native oxidization removal step-oxygen plasma ashing and an ex-situ

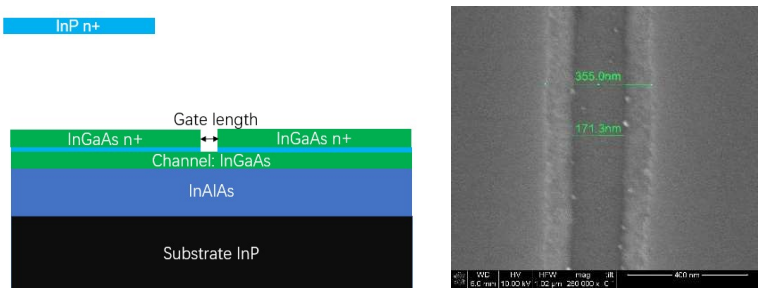
pre-deposition oxide cleaning procedure using HCl: H_2O 1:20 were deployed before thermal evaporation.

3.7 Post Metallization Annealing (PMA)

PMA is an effective way to improve SS and electron mobility because of defects reduction at the interface between oxide and semiconductor. The investigations of PMA atmosphere Ar, N_2, H_2 impact on the interface of Al_2O_3/HfO_2 have been reported [36, 37]. Over 400℃ annealing temperature probably leads to interdiffusion issue at the interface of InGaAs/Al_2O_3 and metal contact with semiconductor [16]. Our PMA condition was selected as 350℃ in N_2/H_2 for five minutes using RTP 1200-100 in the work.

3.8 Refined Pattern Design (Short-gate and Refined-SD)

FET53 is designed for the investigation of short-gate transistors. Unlike the photoresist mentioned in section 3.2 and 3.3, a sort of polymethyl methacrylate-based photoresist PMMA 950 A4 were utilized in the short-gate definition. It is a high-resolution positive resist that can be written by electron beam as well as x-ray and deep UV microlithographic processes. In FET53 fabrication process, the spinner speed was set up as 4500 rpm with 1500rmp/s acceleration and post-baking at 180℃ for 135 seconds to make a super-thin film of a thickness 180nm firmly adhere with the wafer. Then electron beam lithography (EBL) technology was adopted in lithography step. Raith EBL use an accelerated beam of electrons as source and the guaranteed minimum line width is 8 nm [38]. As a result, the trenches with 45 or 52nm depth (the depth depends on if InP n+ layer was kept) and gate length of nanometer magnitude level were successfully fabricated above the channel layer. The narrowest gate length of FET53 is 171nm (see Figure 17).



(a) (b)

Figure 17 (a)gate trench schematic (b) 171 nm gate-length under SEM (courtesy of Patrik Olausson).

Metallization on the pads of source and drain were divided into two steps. A thinner Ti/Pd/Au with thickness 5nm/5nm/40nm was deposited firstly for short gate-length design (see Figure 18 (a)), and then 5nm/5nm/300nm Ti/Pd/Au were deposited as the

same with mention in section 3.5 (see Figure 18 (b)) on the same sample. The SEM (Scanning Electron Microcopy) image in Figure 19 shows the intact metal contact was successfully fabricated to connect channel with source and drain after the first metallization step.

(a) (b)

Figure 18 Metallization steps of refined pattern on source and drain (a) after 5nm/5nm/40nm Ti/Pd/Au deposition (b) after 5nm/5nm/300nm Ti/Pd/Au.

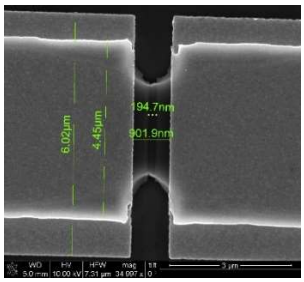


Figure 19 Metallization after 5nm/5nm/40nm Ti/Pd/Au deposition under SEM (courtesy of Patrik Olausson).

3.8 DC-IV Measurement

The performance of transistors and TLM was measured by an automatic probe system at room temperature and 13K. Transfer characteristics I_d vs V_g and output characteristic I_d vs V_d can be measured by probe station through applying probes on the three pads of transistors. Four-point probes method was used to test TLM contact resistance, which is a more accurate resistance measurement technology using separate pairs of current-carrying and voltage-sensing electrodes, so that the effect of probe resistance can be eliminated in the circuit [39] (see Figure 20). It is investigated that this method is also used to measure sheet resistance of thin films (particularly semiconductor thin films)[40].

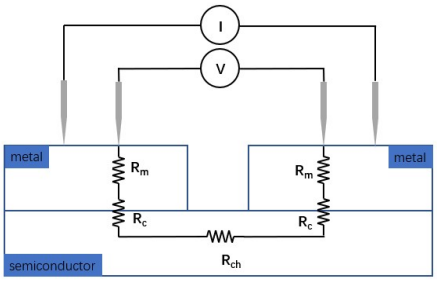


Figure 20 Schematic of the four-point probe method.

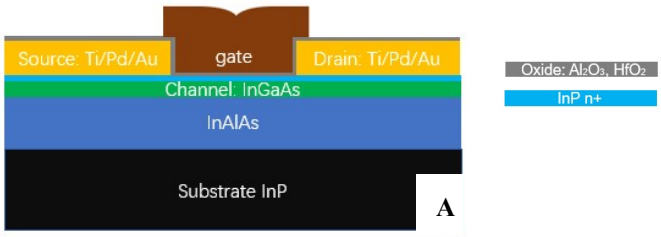
Chapter 4 Results and Discussion

This chapter will discuss the performance of the transistor devices and TLMs in Batch FET44 and Batch FET53. In FET44, three samples with different transistor structures were fabricated on three IET3 wafers. We found a lower contact resistance structure for the InGaAS QW MOSFET in this trial. However, we were unable to obtain more accurate transistor performance data, because we lacked the measurement data for short-gate transistors (gate length lower than 1μm) to do the data fitting work. After several failed experiments and running out of IET3 wafers, short-gate transistor devices were successfully fabricated on the MOS1 wafer (another wafer) in FET53. In the end, we derived more accurate performance parameters through measuring the short-gate transistors. Another benefit of FET53 is that the influence of blank etching processing was deleted because of using the structure-optimized wafer MOS1.

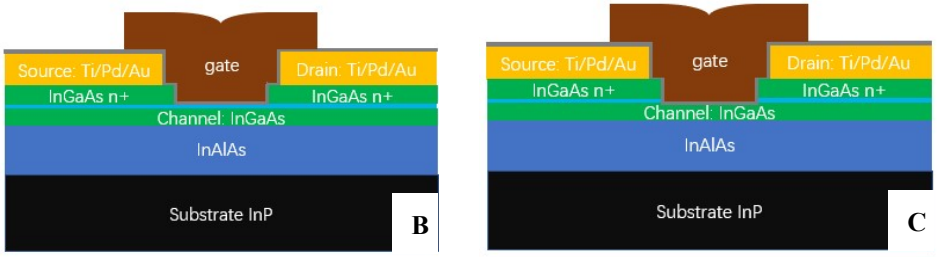
4.1 Results of FET44

4.1.1 Transistors

Samples of different structures, named sample A, sample B and sample C, were prepared in FET44. The aim is to study the contact resistance by comparing the transistor performance on samples A and B (different layouts on source/drain side) and assess the gate electrostatic control by comparing the transistor performance on samples B and C (different structures above channel layer). The metal contact on all samples is Ti/Pd/Au with thickness 5nm/5nm/300nm. Sample A fabricated the metal contact on n-type InP layer on source/drain side. Sample B kept one more layer, n-type InGaAs, between metal and n-type InP. Sample C has the same source/drain layout to sample B, which is metal/n-type InGaAs/n-type InP. The only difference is sample B has one more n type InP layer between gate oxide and channel at MOS stack. The detailed layout of the three samples is shown in Figure 21.
Every sample consists of ten transistors of gate-width parameters as 6μm and 70μm (see Figure 22) and five different gate lengths. The gate length of transistors on sample A is 6,12,18, 24 and 30μm, while the gate length on the other two samples was designed narrower than sample A, they are 2,4,10,16 and 22μm. Ultimately, the gate length with 2μm was not successfully developed during dummy gate definition step, so the shortest transistor data was missing on sample B and C.



(InGaAs n+ layer removed and metal contact with InP n+ layer).



(InP n+ layer kept between oxide and channel) (InP n+ layer etched above channel)

Figure 21 The layout of transistors on FET44.

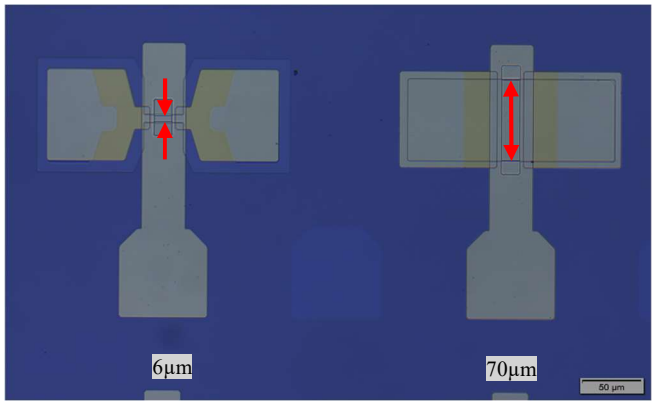


Figure 22 Transistor image of different gate width.

4.1.2 TLM

TLM areas were also planned on FET 44, with metal contact on n type InGaAs for analysis of contact resistance and sheet resistance. The distance is 2, 3, 4, 5, 6, 20 μm on piece one and 7, 8, 9, 10, 11, 20 μm on piece two. Figure 23 displayed the TLM schematic structure and optical microscope photo.

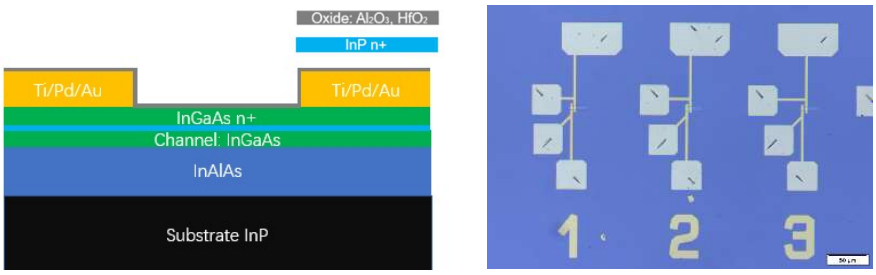


Figure 23 (a)TLM schematic structure. (b) TLM photo under optical microscope.

4.1.3 Transistor: SS, R_c R_{sh} and μ_n

The drain bias (V_d)of the transistors was set as standard value 50mV and 500mV. After sweeping the gate voltage (V_g) from -3V to 0.5V, the transfer characteristic normalized by L_g/w (L_g-gate length, w-gate width) is shown in Figure 24. From simple FET theory,

in linear regime [28]: $I_d = \frac{w \cdot \mu_n \cdot C_g \left[(V_g - V_{th}) V_d - \frac{V_d^2}{2} \right]}{L_g}$

$$\Rightarrow \frac{I_d L_g}{w} = C_g \mu_n \left[(V_g - V_{th}) V_d - \frac{V_d^2}{2} \right] \quad (21)$$

in saturation regime [28]: $I_d = \frac{w \cdot \mu_n \cdot C_g (V_g - V_{th})^2}{2 L_g}$

$$\Rightarrow \frac{I_d L_g}{w} = \frac{\mu_n \cdot C_g (V_g - V_{th})^2}{2} \quad (22)$$

Where, I_d is drain current, C_g is gate capacitance, V_{th} is threshold voltage, μ_n is electron mobility. According to the Equation (21) and (22), the value of $\frac{I_d L_g}{w}$ is irrelevant to the size of the gate area, so normalizing the transfer characteristic help exclude dimensional factors of the gate from the same sample. Ideally, the transfer characteristic of ten transistors from each group (the same color curves) should overlap.

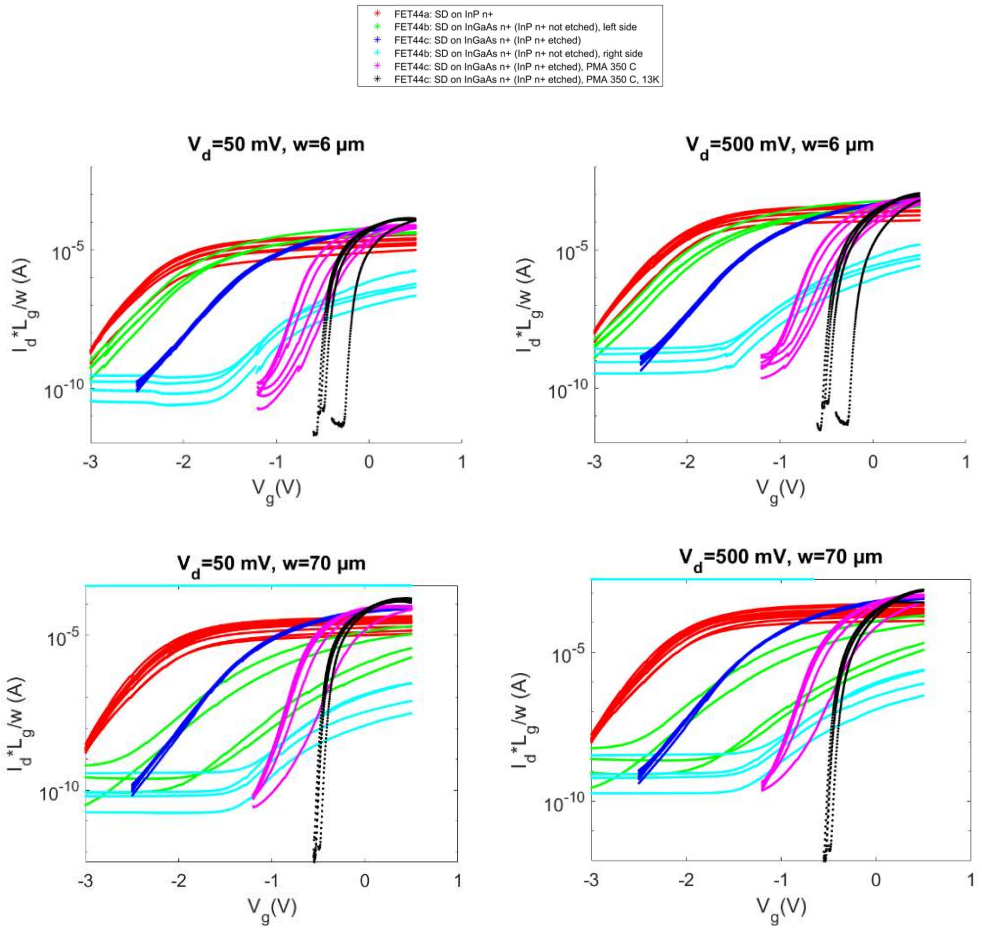


Figure 24 Transfer characteristic of FET44.

In this work, the slight scattering of transfer performance from the same group probably related to the small difference of resistances and mobility in transistor, whereas the data of sample C (blue curves) presented a good consistency, which indicated the uniform processing of sample C. It is noticed that the transfer performance of sample B has a larger variance on different pieces (see the green curves and light blue curves). This inconsistency coincides with the finding under optical microscope shown in Figure 25- the difference in color contrast is visible on the hall bar of sample B. The saturation current $I_{saturation}$ reached the range of 10^{-5} to 10^{-4} A under 50mV drain bias, and approached the range of 10^{-4} to 10^{-3} A under 500mV drain bias on all the transistors. The saturation current is higher than previous batch sample implies excellent carrier mobility. Furthermore, in consideration of the best transfer performance of sample C, the sample was treated by PMA and tested I-V under room temperature and 13 K low temperature (see purple curves and black curves). Obviously, the slope changed in the liner area of transfer characteristic after PMA.

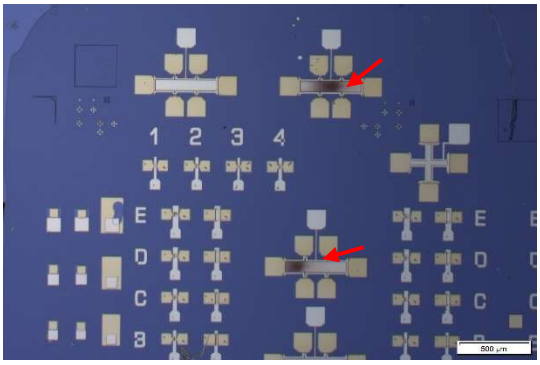


Figure 25 Difference of color contrast on sample B under optical microscope (see red arrows).

The transistor subthreshold swing (SS) was extracted from transfer data of each sample as shown in Figure 26. The SS of sample B will not be disscussed, as the curve of transfer performance significantly spread out on the same struture transistors. SS data of Sample A and sample C is similar, distributing near 300 mV/dec. After PMA at 350°C in H_2/N_2 atmosphere, the SS of sample C was reduced to 100 mV/dec at room temperature and further reduced to 20 mV/dec at 13K. That is a significant improvement of electrostatic gate control after annealing.

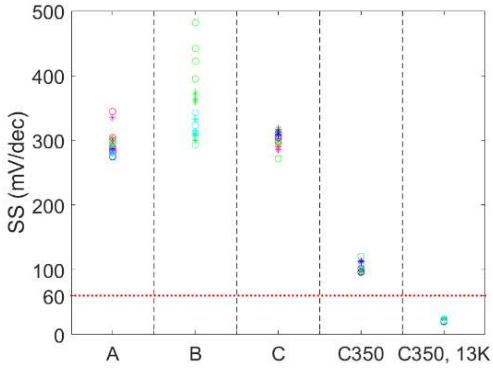


Figure 26 SS data of transistors of FET44.

Based on Equation (18), contact resistance R_c and carrier mobility μ_n were calculated from the fitted curve of g_m against the gate length; here, g_m is extracted at V_{ds}=500 mV. The fitting curves are shown in Figure 27, and the results are listed in Table 1. The curve of sample A was not fitted as well as sample C, especially on 6μm gate-width transistors. The reason for the scattered data of sample A may be caused by the gate-length difference between the actual transistors and our design. Generally, it fits well on the straight part of the curves with the contact distance change from 1μm to 20μm, which means the derived μ_n results are reasonable. However, the bending tendency of the curves with the contact distance lower than 1μm is uncertain because of lacking the short contact distance data, so the total contact resistance values on source/drain may have a big error are shown for reference. To get more accurate contact resistance data, it requires to fabricate short gate-length transistors (narrow to 100nm gate-length is necessary).

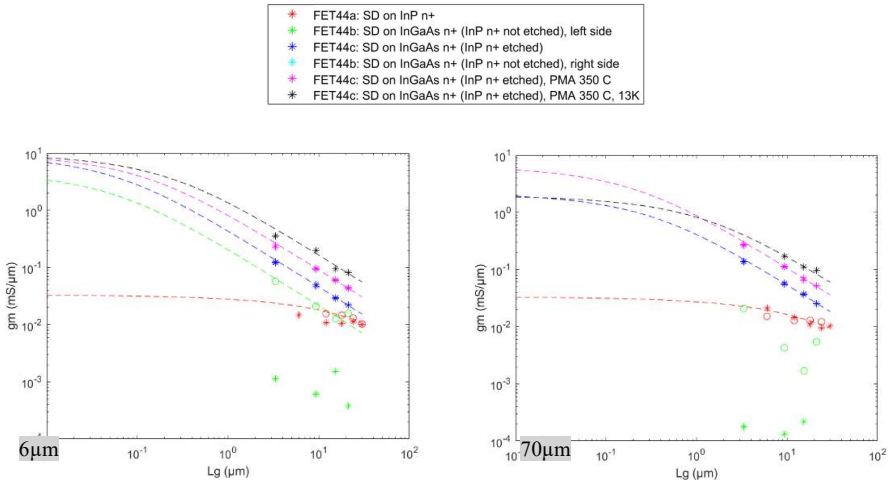


Figure 27 fitting curve of contact resistance R_c and Carrier Mobility μ_n based on g_m &L_g.

Table 1: R_c (Ωμm) / μ_n (cm^2/Vs) fitting results

Sample	g_m of 6 μm transistors	g_m of 70 μm transistors
FET44A	15000/1950	15000/1500
FET44C	50/1450	225/1750
FET44C PMA350°C	50/2900	75/3400
FET44C PMA350°C 13K	50/5300	250/5900

From Table 1, μ_n is doubled after PMA and increased further under test temperature of 13K. R_c of sample C is significantly less than Sample A (the benchmark group). Figure 28 shows the series resistance circuit of both samples. The fact is that the lower R_c sample has one more series contact resistance because of the existence of n-type InGaAs. The n-type InGaAs layer probably help reduce the barrier height or thickness at the interface between two different layers, leading to lower contact resistance results on sample C. It is concluded that sample C has the best contact design on source/drain side in FET44, metal contact / n-type InGaAs / n-type InP.

$$R_{C3} + R_{C4} \ll R_{C1}$$

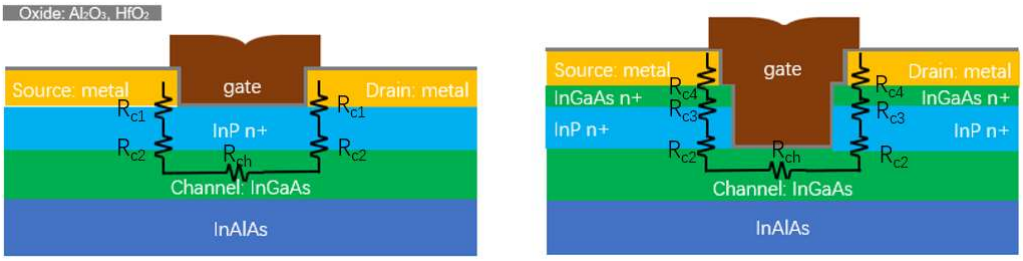


Figure 28 The circuit schematic of sample A (left) and sample C (right).

Using R_{on} at V_{ds}=50 mV, we obtained R_c, μ_n (according to Equation (19)), as well as R_{sh} (according to Equation (6)) of sample C. Figure 29 shows the fitting data of R_c and μ_n from R_{on} with gate-length, and Figure 30 and 31 represents fitting data of R_{sh} from R_{on} with gate length/width. The results are summarized in Table 2.

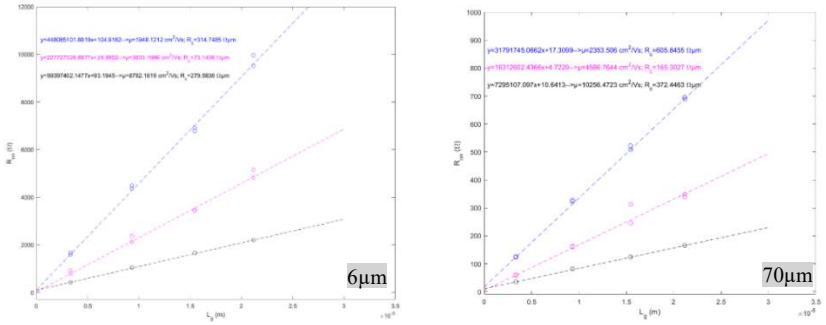


Figure 29 R_c and μ_n derive from R_{on} of sample C (V_{ov}= V_{gs}-V_{th}=0.3 V)
(Before PMA- blue, after PMA- purple and after PMA 13K-black)

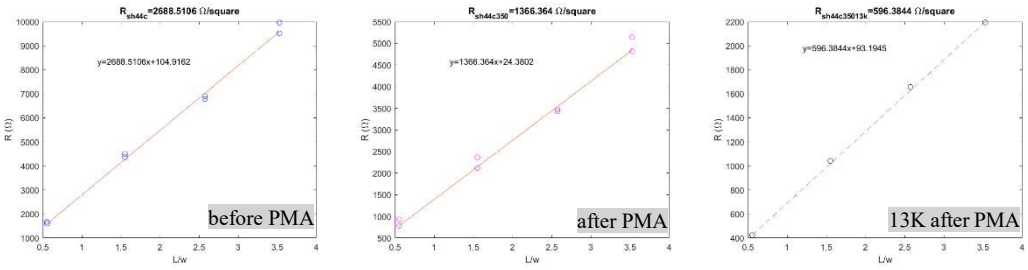


Figure 30 R_{sh} of sample C from R_{on} vs L_g on 6μm-width transistors.

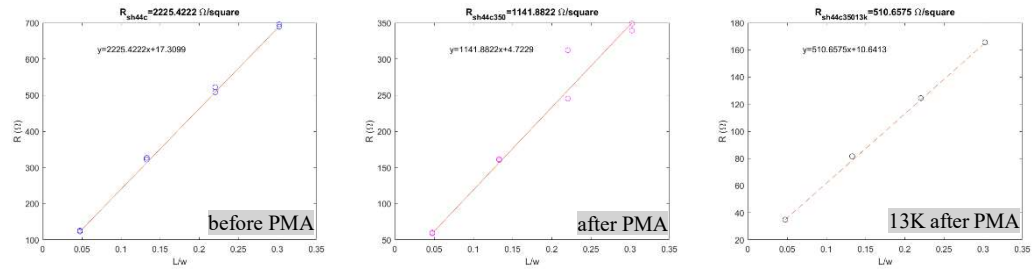


Figure 31 R_{sh} of sample C from R_{on} vs L_g on 70μm-width transistors.

Table 2: R_c (Ωμm) / R_{sh} (Ω/square) and μ_n (cm^2/Vs) result based on R_{on}
(The negative values of resistance probably because of big error of sample A data)

Sample	R_{on} 6 μm transistors	R_{on} of 70 μm transistors
FET44A	112975/-846/-6193	58645/1626/3221
FET44C	315/2689/1948	606/2225/2354
FET44C PMA350°C	73/1366/3833	165/1141/4587
FET44C PMA350°C 13K	280/596/8782	372/511/10256

R_c is dramatically low on sample C, which matches the finding derived from g_m and further proves the sample C contact design on source/drain is the best design in FET44. μ_n enhances more significantly testing under 13k. The reasons will be discussed in the discussion section.

4.1.4 TLM: R_c and R_{sh}

TLM R_c and R_{sh} are derived from the same original resistance data obtained by probe station, but fitting with respect to distance and distance/width, respectively. Figure 32 displays the fitting result of R_c (the fitting results of R_{sh} are shown in Figure A1 in the Appendix). The twelve original resistance data, blue dots in Figure 32 before PMA, are sparse along with two different straight lines, so the fitting them is not reliable. Although R_c = -9.3085Ωμm is unreasonable, based on the confidence interval -52Ωμm to 33Ωμm, it indicates that the contact resistance is very low as we expect. The TLM of distance parameters lower than 6 μm were treated with PMA 350°C. The resistance data measured by probe station nearly did not change after PMA. Therefore, annealing may not help improve the quality of the contact interface between metal and n type InGaAs layer in this case. From Table 3, R_{sh} results nearly remain the same before and after PMA. The significant improvement of R_{sh} at 13K are probably only related to the electron mobility increase in the channel layer at low temperature.

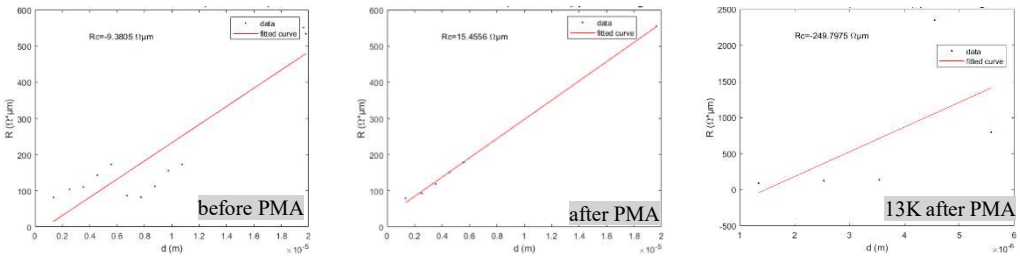


Figure 32 R_c of TLM.

Table 3: R_c (Ωμm) / R_{sh} (Ω/square)
(Note: the negative values of resistance because of data scattering or more data needed to do accurate fitting)

sample	TLM InGaAs n+
FET44C	-9/25
FET44C, PMA350°C	15/27
FET44C, PMA350°C, 13K	-249/342

4.2 Results of FET53

4.2.1 Transistors

The structure of short-gate sample A is nearly the same as sample C in FET44 except for the source and drain design as shown in Figure 33. Except short-gate length, the metallization step is divided into two steps (add one thinner and smaller metal layer as mentioned in chapter 2).

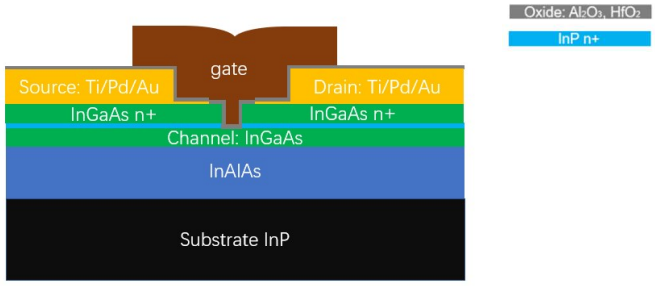


Figure 33 The layout of transistors on sample A.

2.5μm gate-width and 63.5μm gate-width transistors were fabricated. The gate dimension is listed in Table 4. All the length of gate was measured by SEM for collect precise data. Seven out of tenth transistors have the gate length lower than 1μm. The narrowest gate-length is 171.3nm (see SEM photo from Figure 17 (b) in chapter 3), and the longest gate-length is 5.84μm. Figure 34 shows an optical microscope photo of the transistors.

Table 4: the gate length (μm) of transistors on sample A

Transistor sample	Width 2.5μm	Width 63.5μm
1	0.1713	0.1713
2	0.1768	0.1768
3	0.2012	0.2012
4	0.2356	0.2356
5	0.2594	0.2594
6	0.4085	0.4085
7	0.6096	0.6096
8	1.02	1.02
9	2.31	2.31
10	5.84	5.84

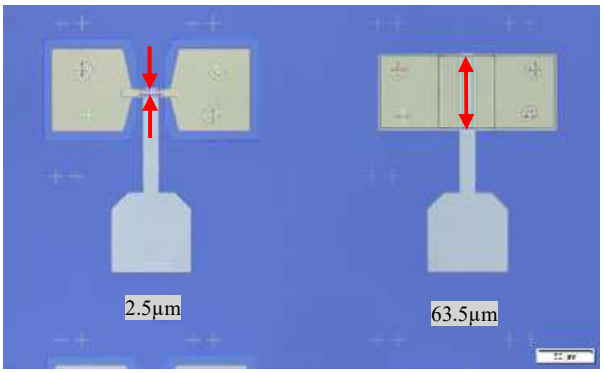
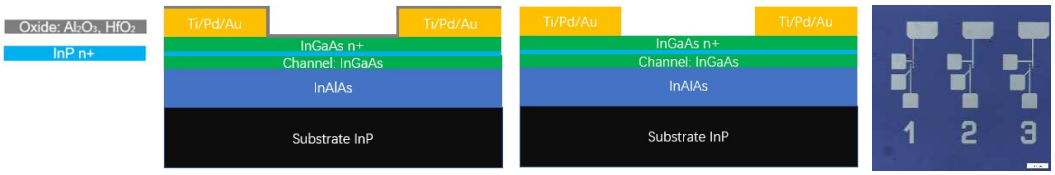


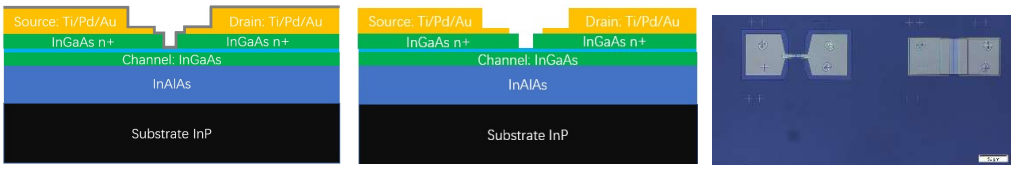
Figure 34 Transistor image of different gate width.

4.2.2 TLM with four-terminal and two-terminal

Two groups of 2×2 factorial design of experiment (DOE) investigates high-κ layer and PMA influence on R_c/R_{sh}. Group one is four-terminal TLM using n-type InGaAs as semiconductor layer, and the distance between source and drain is the similar to FET44. Group two is two-terminal TLM connecting by n-type InP as transistor pattern, so the resistance only can be measured by 2-point method. The distance between source and drain of two-terminal TLM is the same as the transistor channel length in Table 4. The different layouts of both groups were shown in Figure 35.



(a) Group one: four-terminal TLM



(b) Group two: two-terminal TLM

Figure 35 TLM schematic layout and photos under optical microscope.

4.2.3 Transistor: SS, R_c R_{sh} and μ_n

The drain bias of the transistors was also set as 50mV and 500mV. Figure 36 demonstrates the normalized transfer characteristic (like FET44) after sweeping gate bias from -2V to 0.5V. $I_{saturation}$ is in the range of 10^{-5} A under 50mV drain bias before PMA (red curves), and in the range of 10^{-4} to 10^{-3} A under 500mV drain bias after PMA (green curves). Overall, the saturation current is the same as FET44, but the transistor turn-off current is lower by four orders of magnitude compared to FET44. Given the discrepancy in transfer performance, the highest $I_{saturation}$ still nearly 100 times lower

than FET44, demonstrating leakage issue of FET53. Furthermore, the transistor turn-off current did not reach the range of 10^{-10} A after PMA (see green curves in Figure 36), or not as lower as FET44C. The inconstant processing can be detected from the spread-out curves in the same group, and the difficulty of short-gate transistor fabrication is probable an influence factor for processing.

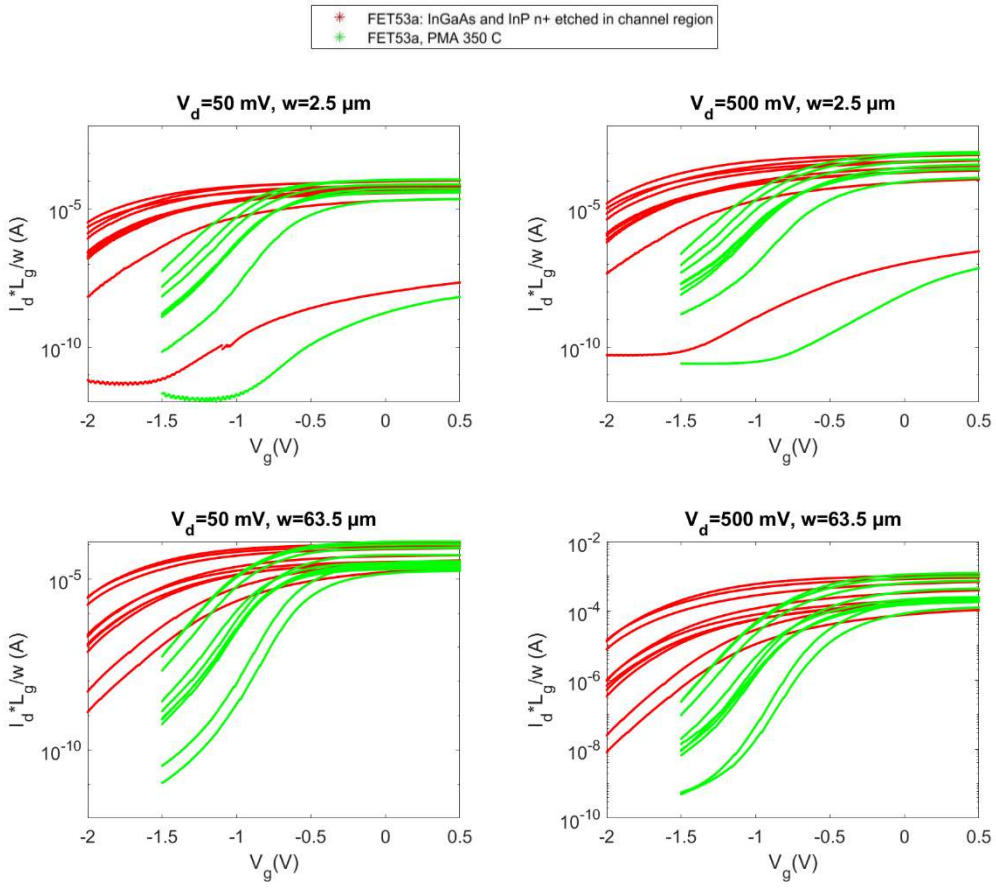


Figure 36 Transfer charcteristic of FET53.

Transistor SS is extracted from the transfer characteristic (see Figure 37). The SS value disperse between 300-400 mV/dec and 110-200 mV/dec before and after PMA, repectively. Further, transfer characteristic at 13K was not measured, because the SS value spread out seriously and minimum SS in FET53 is the same with FET44 after PMA.

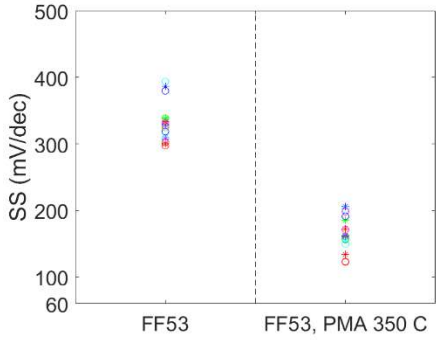


Figure 37 SS data of FET53A.

R_c, R_{sh} and μ_n were also exacted using two methods like FET44, g_m derived at V_{ds}=500 mV (see Figure 38) and R_{on} derived at V_{ds}=50 mV (see Figure 39 to Figure 41). In Figure 38, more data was achieved with short gate-length ranging from 0.1 μm to 1 μm compared with Figure 27, which indicates the fitted curve is more precise on the bending part as we expected. The results listed in Table 5 show that R_c and R_{sh} approximately shrink to half, and electron mobility μ_n approximately doubles after PMA.

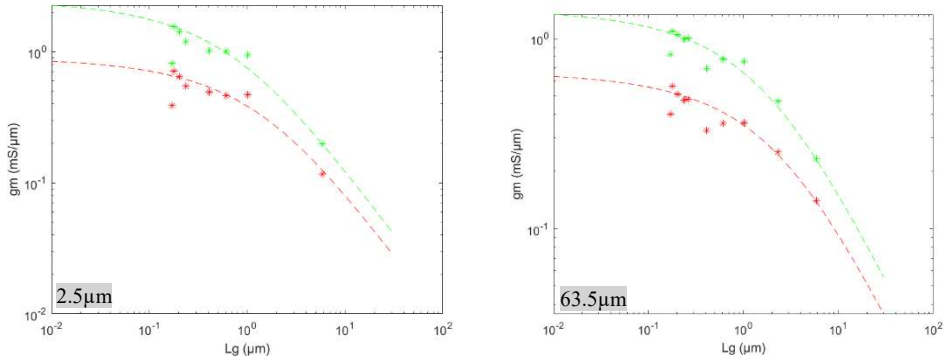


Figure 38 Fitting curve of contact resistance R_c and carrier mobility μ_n based on g_m &L_g. (Before PMA- red and after PMA- green)

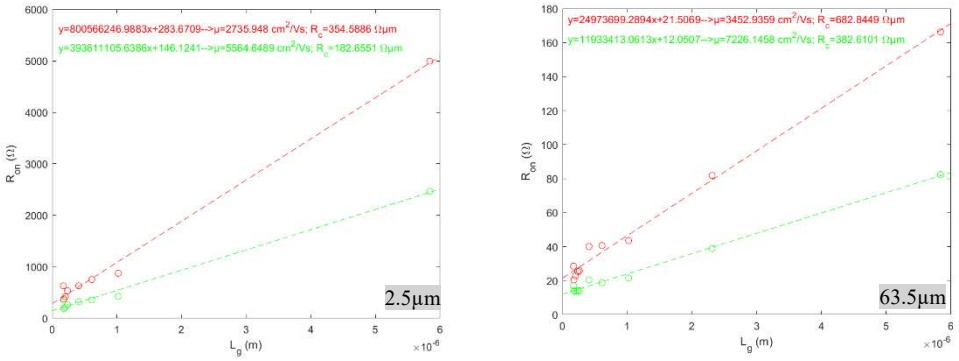


Figure 39 R_c and μ_n derive from R_{on} of sample A (V_{ov}= V_{gs}-V_{th}=0.3 V). (Before PMA- red and after PMA- green)

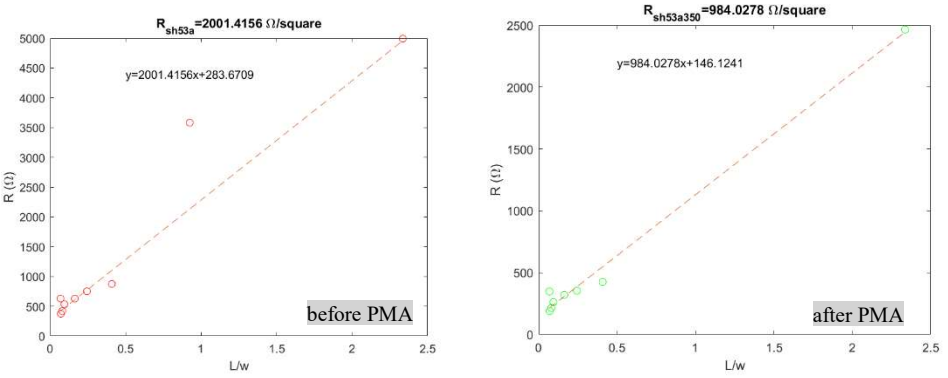


Figure 40 R_{sh} from R_{on} vs L_g on 2.5μm transistors.

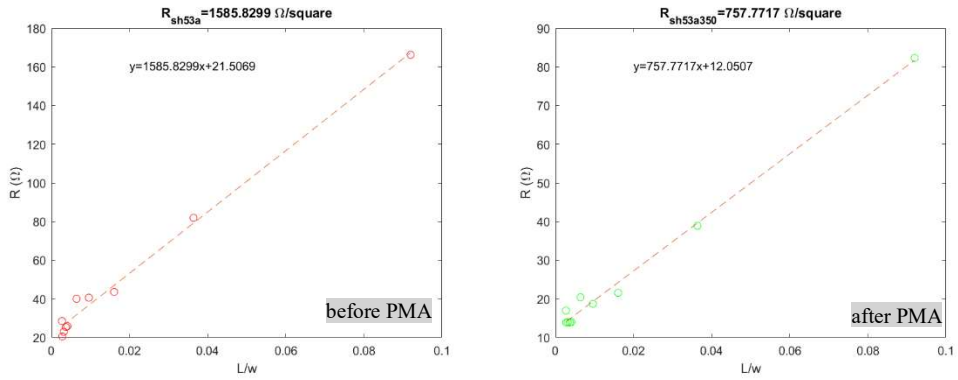


Figure 41 R_{sh} from R_{on} vs L_g on 63.5μm transistors.

Table 5: R_c (Ωμm) / R_{sh} (Ω/square) and μ_n (cm^2/Vs) results

Sample	gm, 2.5 μm	gm, 63.5 μm	Ron, 2.5 μm	Ron, 63.5 μm
FET53A	550/-/3000	750/-/3800	355/2000/2700	683/1590/3500
FET53A, PMA350°C	200/-/4300	350/-/5800	183/980/5600	383/760/7200

4.2.4 TLM: R_c and R_{sh}

R_c and R_{sh} of TLM summarized in Table 6 and 7, and the extracted diagrams are shown in Appendix (Figure A2 to Figure A5). N-type InGaAs TLM has a lower R_c and R_{sh}, compared with n-type InP TLM. PMA slightly improves R_{sh} of n type InP, but not impact on R_c. The last DOE factor (high-κ) apparently minishes the R_c and R_{sh} of the two-terminal TLM.

Table 6: Contact resistance R_c (Ωμm) / sheet resistance R_{sh} (Ω/square) measured by 4-point method (Note: the negative values of resistance because of data scattering or more data needed to do accurate fitting)

sample	TLM InGaAs n+
FET53A with high-κ	-18/72
FET53A with high-κ, PMA	-9/84
FET53B34 without high-κ	-24/62
FET53B34 without high-κ, PMA	-21/62

Table 7: Contact resistance R_c (Ωμm) / sheet resistance R_{sh} (Ω/square) measured by 2-point method

Sample	TLM InP n+ (2.5 μm)	TLM InP n+ (63.5 μm)
FET53B12 with high-κ	72/165	99/219
FET53B12 with high-κ, PMA	71/139	110/197
FET53B34 without high-κ	136/225	106/336
FET53B34 without high-κ, PMA	130/157	148/247

4.3 Discussion

4.3.1 Leakage Issue of FET53

The higher turn-off current indicates an inadequate electrostatic gate control of FET53 compared with FET44. In fact, a shadow area around the gate trench was detected by SEM after wet-etching in the step of gate length definition (see the left image of Figure 41). It is deduced that the shadow area may be caused by over-etching n-type InP (see the right schematic diagram in Figure 41) during isotopic wet-etching. However, the enlarged gate-length will not lead to the current leakage issue. In addition, there is a slight chemical composition difference in the channel layer between FET53 and FET44. The Indium-rich InGaAs interlayer ($In_{0.8}Ga_{0.2}As$ see Figure 15 in chapter 3) of FET53 has a narrower band gap than $In_{0.53}Ga_{0.47}As$ of FET44 [41], which should improve gate electrostatic control but not bring about of the leakage issue. Lastly, the only possible reason for leakage we could find is that the Si doping level in the backside layer (InAlAs layer see right schematic diagram in Figure 42) of FET53 ($1\times10^{19}/cm^3$) is higher than FET44 ($0.5\times10^{18}/cm^3$).

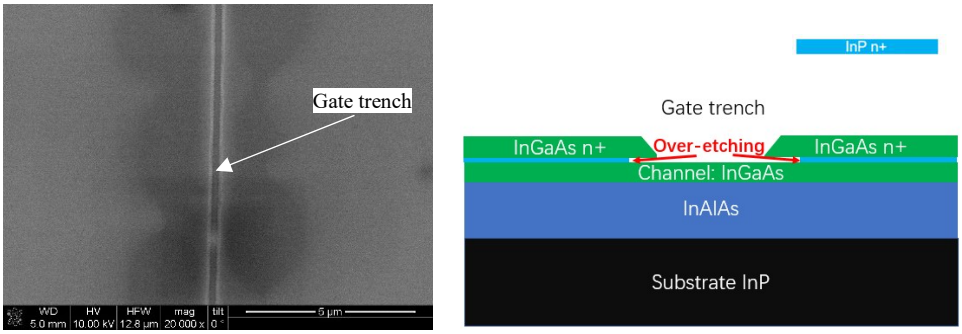


Figure 42 Gate trench SEM photo (left courtesy of Patrik Olausson) and schematic of gate trench (right).

4.3.2 PMA Impact

PMA technology significantly improves the electrostatic gate control, which is reflected by the remarkable reduction of SS after PMA at 350°C in N_2/H_2. In this study, SS of all the transistors decreased by a factor of three; for example, sample C of FET44 decreased from 300 mV/decade to 100 mV/decade. SS reduction is probably related to the density of the charge trap dropped in the MOS capacitor (Ti/Pd/Au-Al_2O_3/HfO_2-InGaAs) after annealing. By reducing Coulomb scattering, electron mobility improves, and the potential in the quantum well is better controlled by gate voltage consequently. In the PMA process, atmosphere plays an important role as well. It was reported that the forming gas of 95% N_2 and 5% H_2 helps to unpin Fermi-level in Pt-HfO_2-$In_{0.53}Ga_{0.47}As$ MOS capacitor, compared to 100% N_2 [42]. In our work, atomic hydrogen can be generated during annealing, then penetrated through the metal contact

layer and saturated the defects (like dangling bonds) at the interface of channel surface and high-κ layer (see Figure 43). The decrement of defects after PMA also explains the noticeable improvement in contact resistance, sheet resistance and carrier mobility (summarized in Table 8) in transistors. Nevertheless, an improvement on R_c and R_{sh} TLM was not found, which indicated that the defects at the source/drain side are limited and cannot decrease during the annealing process. In summary, PMA significantly improved the interface quality between the channel and the high κ layer, but not between the different layers of the source/drain.

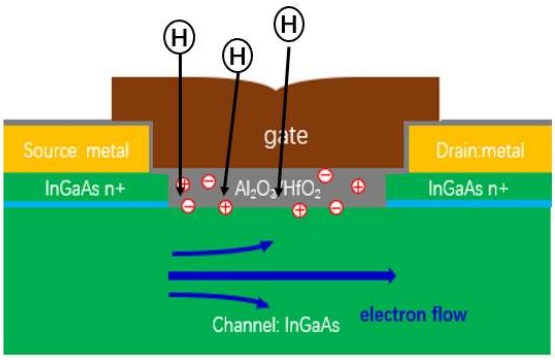


Figure 43 Schematic of PMA effect at the interface of channel and high-κ layer.

Table 8: summary of transistors characteristic exacted from R_{on}/g_m (R_{sh} cannot extract from g_m)

Sample	before PMA				after PMA			
	FET53A 2.5 μm	FET 44C 6 μm	FET 53A 63.5 μm	FET44C 70 μm	FET53A 2.5 μm	FET44C 6 μm	FET53A 63.5 μm	FET44C 70 μm
R_c (Ωμm)	355/550	315/50	683/750	606/225	183/200	73/50	383/350	165/75
Rsh (Ω/square)	2000/-	2689/-	1590/-	2225/-	980/-	1366/-	760/-	1141/-
μ_n (cm^2/Vs)	2700/3000	1948/1450	3500/3800	2354/1750	5600/4300	3833/2900	2700/5800	4587/3400

4.3.3 Measurement Temperature Effect

According to Equation (3), the SS value should be proportional to the measurement temperature, so SS tendency should be theoretically consistent with the black dot line with the temperature in Figure 44. Taking the SS value (100 mV/dec) at 300K into account, the SS value should be down to 4.3 mV/dec at 13K. However, the experimental data is 20 mV/dec, nearly five times as many as the theoretical SS value.

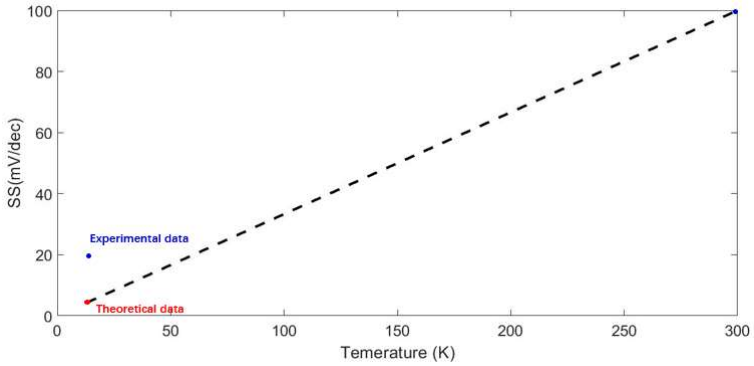


Figure 44 SS vs T (black dot line-theoretical tendency).

The discrepancy between theoretical and experimental SS probably is related to InGaAs band gap change due to band tail effect at low temperature. Band tail existence near the bottom of the conduction band provides more states for electrons passing through at lower temperature (see the schematic from Figure 45). As a result, the potential in the quantum well change leads to the degradation of gate electrostatic control in comparison to the theoretical condition.

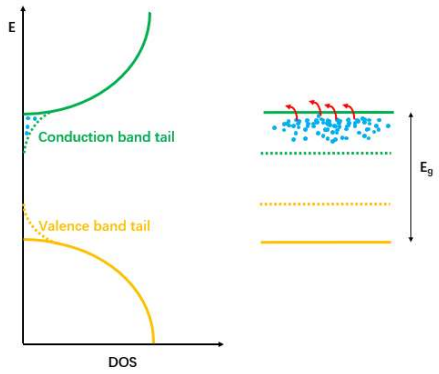


Figure 45 Schematic of band tail effect (blue dot-electrons in band tails, adapted from [43]).

The temperature influence on electron mobility has already been mentioned in chapter two. In this project, the electron mobility μ_n was separately derived from g_m and R_{on} at room temperature and 13K. The value of μ_n extracted from R_{on} method is much higher than g_m on the same samples, especially under low temperature (see the Figure 46). The g_m method Equation (18) may need to take short-gate effect into account. If we assume that only lattice scattering mechanism dominates carrier mobility, based on Equation (12), μ_n should increase to 3.2×10^5 cm^2/Vs at 13 K (use the minimum μ_n=2900 cm^2/Vs as an example). However, the practical value is 5300 cm^2/Vs, so the impurity scattering mechanism is an essential factor in the InGaAs QW-MOSFET as well. It is concluded that the defects at the interface of channel surface and high-κ layer probably limit the carrier mobility at low temperature 13K.

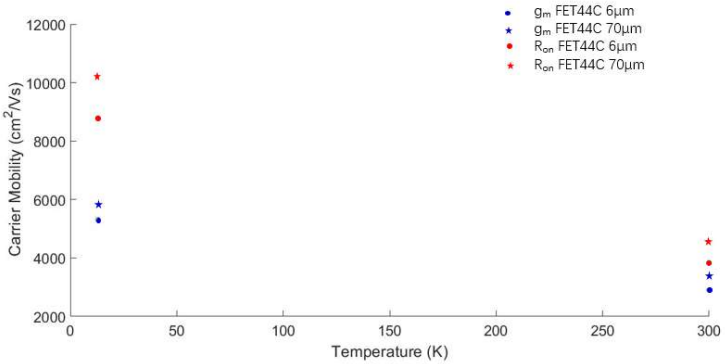


Figure 46 Carrier mobility vs temperature

Chapter 5 Conclusion and Outlook

By fabricating InGaAs quantum-well MOSFET using different techniques and measuring performance of transistors and TLMs using probe station, we get the following conclusions:

- Multiple layers N-type InGaAs/n-type InP over channel is a better structure compared with monolayer n-type InP, it was shown two orders of magnitude improvement on contact resistance in this work. The optimized source/drain contact structure will contribute to attaining a high-speed, low-power consumption transistor.
- The transistors without n-type InP layer in MOS stack show excellent gate electrostatic control, and SS narrows down to 100mV/dec after PMA.
- The transfer characteristic of short-gate QW-FET presented higher electron mobility. The evidence is found from high saturation current, and electron mobility derived using two methods. Electron mobility of more than 5000 cm^2/Vs can be extracted from both Equations related to g_m and R_{on}.
- Short-gate QW-FETs were successfully fabricated based on our design. The narrowest gate length is 171.3nm, and seven out of tenth gate-length are lower than 1μm. As a result, more accurate contact resistance and sheet resistance can be derived. Contact resistance is lower than 400Ωμm, and sheet resistance is less than 1000Ω/square, which indicates InGaAs QW-FET is a potential candidate for HEMTs.
- The performance of transistors significantly improved after PMA at 350°C in N_2/H_2 for 5 minutes. The improvement is probably related to the density of charge traps, which significantly drop at the interface of channel and high-κ layer during annealing. N_2/H_2 contribution should also be considered, which helps solve Fermi-level pinning issue for III-V compound. However, TLM contact resistance and sheet resistance only have a tiny improvement after PMA, implying that the contact resistance between Ti/Pd/Au and high-doped semiconductor is already very low. In summary, PMA technology is mainly used to improve the interface's quality between channel and high-κ layer, and the PMA effect between Ti/Pd/Au and n type InGaAs is minimal in this case. It is noticed that PMA also works between n type InGaAs and n type InP.
- Measurement temperature impact on electron mobility was also studied in the work. The experimental results further prove that two scatter mechanisms occur during electron drift, lattice scattering and impurities scattering. The experimental SS is much higher than theoretical SS at 13K, band tail effect probably is one reason for the degradation of the gate electrostatic control in practice compared with theory.

The findings indicate that the InGaAs QW MOSFET is a promising low power-consumption and high-efficiency transistor. It is worthy of doing further research on InGaAs QW MOSFET.

Chapter 6 Further Work

- FET53 Leakage issue may be related to n-type InP layer (above channel) over-etching, so some etching trails to avoid it is expected to try, such as shorten the etching time in solution of HCl:H_2O_2 1:1, or digital etching with lower etching rate.

- Gate trench profile should be optimized for short-gate transistors. Etching hard mask combining with dry etching probably could be one options for narrower gate-length (shorter than 100nm) and smooth wall of the gate trench.

- In the investigation of testing temperature effect, the transistor performance is better to test under different temperature, not only 13K.

- Equation (18), the relation of g_m with mobility and contact resistance, could be optimized. Short-gate effect probably need to be considered.

Appendix

A1: FET44C TLM R_{sh} Data

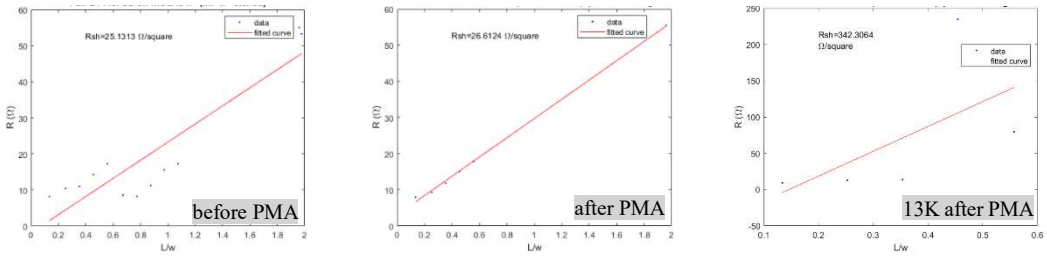


Figure A1 R_{sh} of TLM of FET44C

A2: FET53 TLM R_c and R_{sh} Data

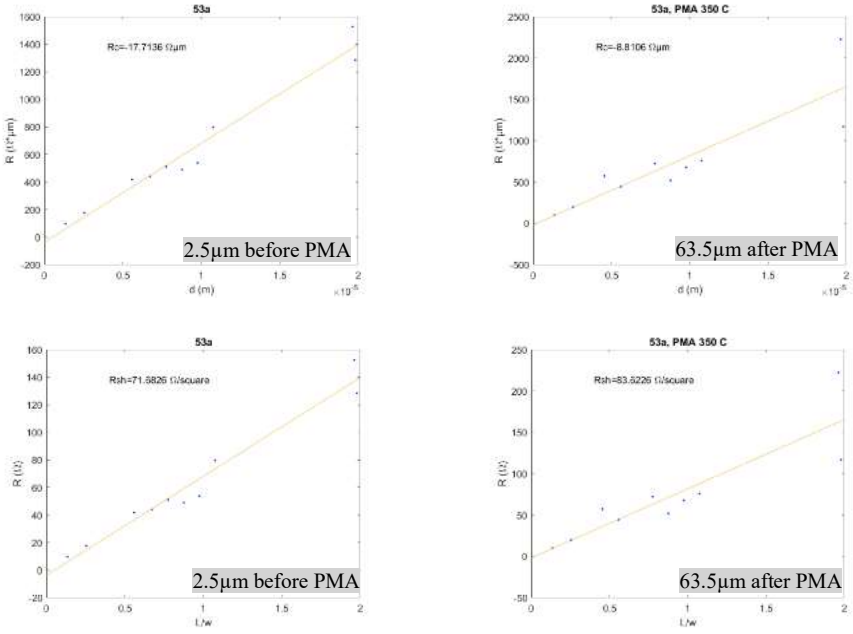
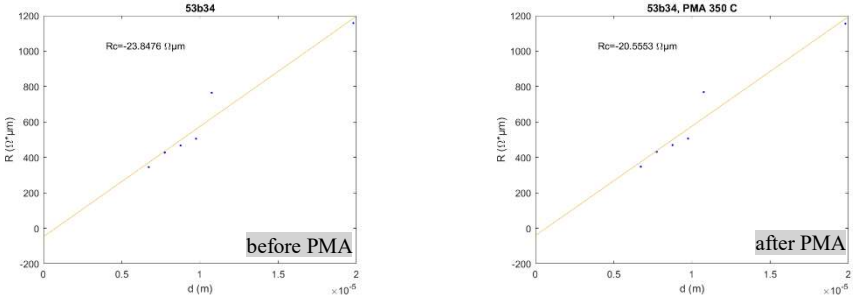


Figure A2 R_c and R_{sh} of 4-terminal TLM with high-κ



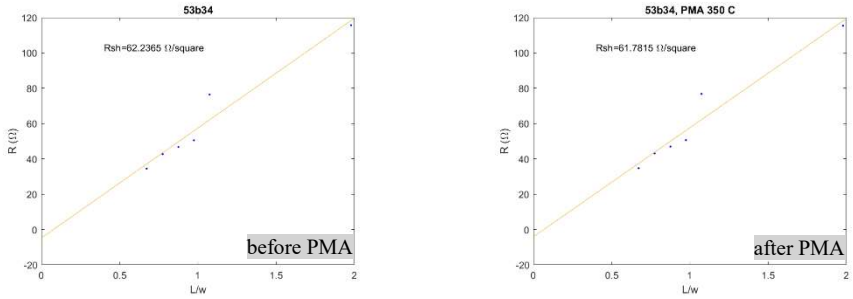


Figure A3 R_c and R_{sh} of 4-terminal TLM without high-κ

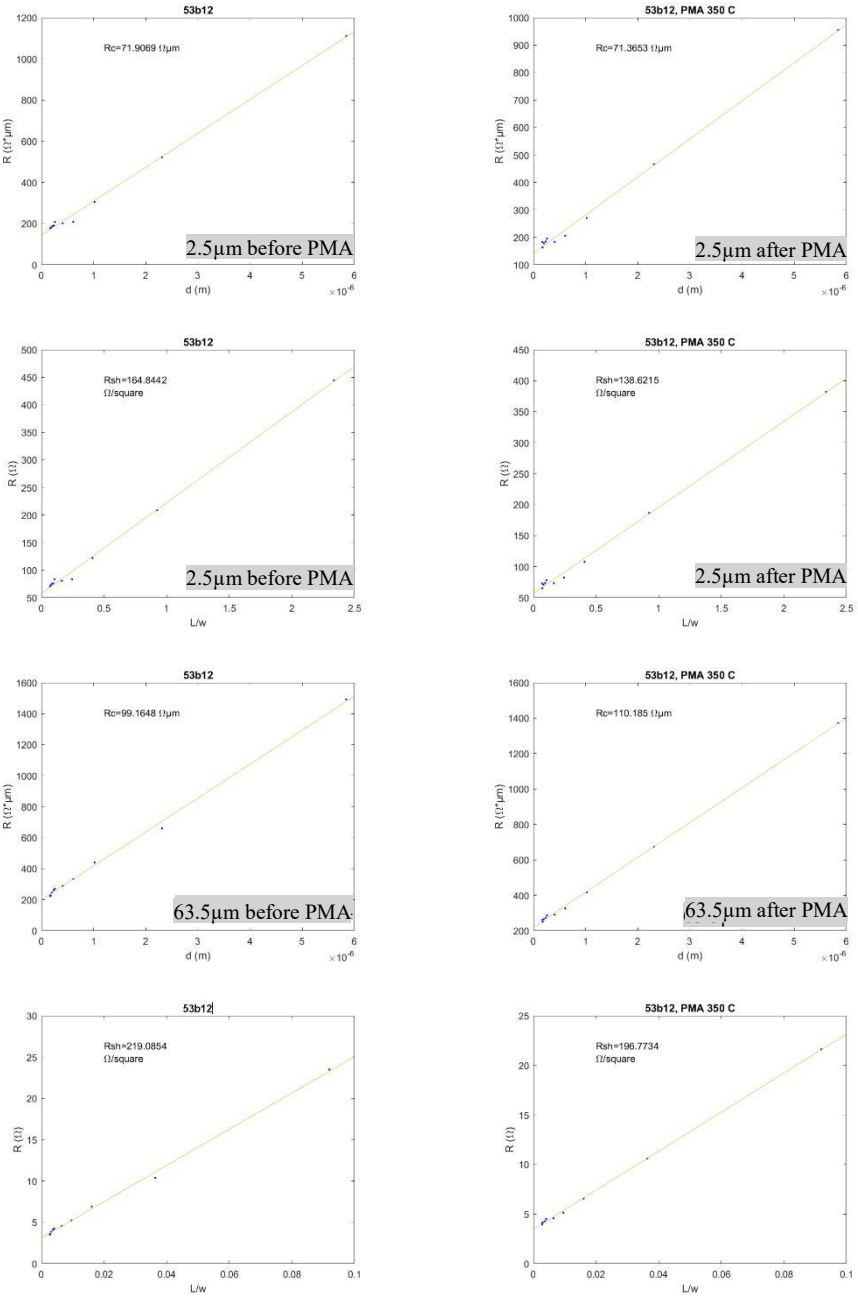


Figure A4 R_c and R_{sh} of 2-terminal TLM with high-κ

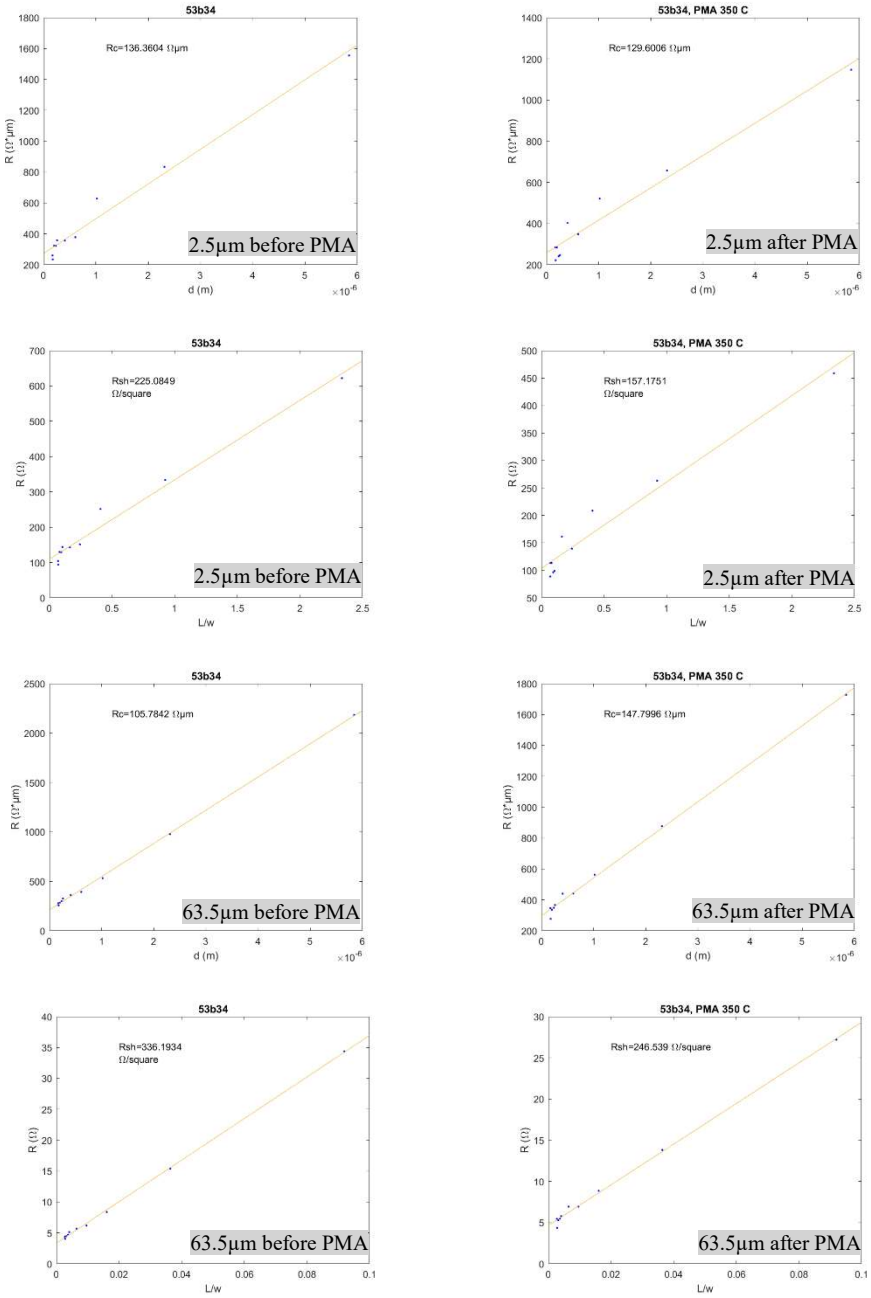


Figure A5 R_c and R_{sh} of two-terminal without high-κ

A3: Hard Mask Investigation

To fabricate short-gate QW-MOSFET, we tried different hard masks from FET47 to FET52, and the processing may work on RTD3 wafer but not on IET3. The results are summarized in Table A1.

Table A1

Sample name	Hard mask	Wafer	Acid	Result
FET47	silicon nitride	IET3	citric	failed *
FET48	silicon nitride	RTD3	citric	Super good
FET49	silicon nitride	IET3	citric	failed-hard mask crack (see Figure A6)
			H3PO4	failed-hard mask crack (see Figure A7)
FET50	Silicon	RTD3	citric	failed-hard mask crack (see Figure A8)
			H3PO4	failed-hard mask crack (see Figure A9)
FET51	Silicon dioxide	RTD3	citric	good
FET52	Silicon dioxide	IET3	citric	failed-not relate to hard mask**

* The failed reason may be related with that we did mesa etching before gate definition

** The etching rate is completely different with reference sample

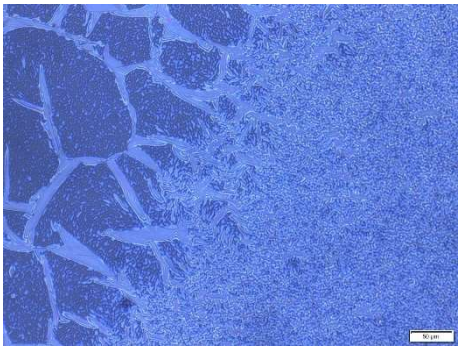


Figure A6 FET49 after Citric etching

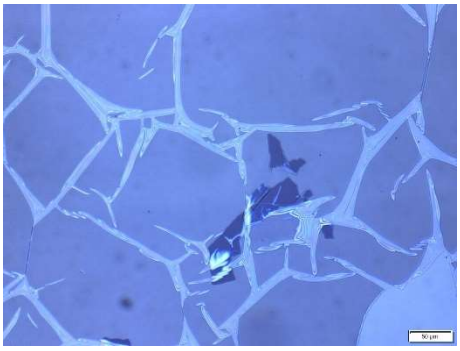


Figure A7 FET49 after H3PO4 etching

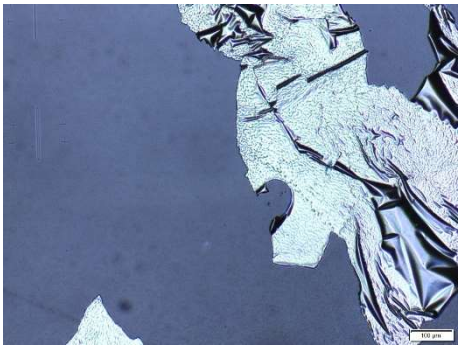


Figure A8 FET50 after Citric etching

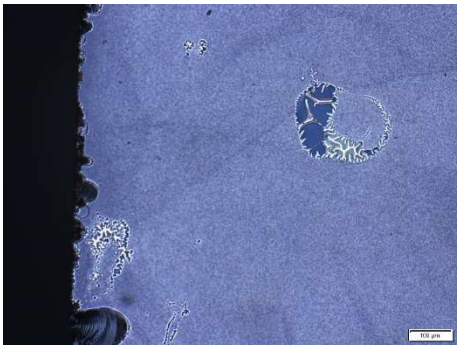


Figure A9 FET50 after H3PO4 etching

Reference

[1] Nair, B. Somanathan, “Digital electronics and logic design”, New Delhi, India, *Prentice-Hall of India*, 2006, p.240.

[2] Shunri Oda, David Ferry. Silicon Nanoelectornics, Taylor&Francis, 2005, p.178

[3] R. Chau, B. Doyle, S. Datta, J. Kavalieros and K. Zhang, “Integrated Nanoelectronics for The Future”, *Nature Mater, vol.* 6, pp. 810–812, Nov. 2007, doi: 10.1038/nmat2014.

[4] C. B. Zota, L. -E. Wernersson and E. Lind, "High-Performance Lateral Nanowire InGaAs MOSFETs with Improved On-Current," in *IEEE Electron Device Letters*, vol. 37, no. 10, pp. 1264-1267, Oct. 2016, doi: 10.1109/LED.2016.2602841.

[5] M. Egard *et al*., "High-Frequency Performance of Self-Aligned Gate-Last Surface Channel In0.53Ga0.47As MOSFET," in *IEEE Electron Device Letters*, vol. 33, no. 3, pp. 369-371, March 2012, doi: 10.1109/LED.2011.2181323.

[6] J. Nah, H. Fang, C. Wang, K. Takei, M. Y. Lee, E. Plis, S. Krishna and A. Javey*,* “III–V Complementary Metal-Oxide-Semiconductor Electronics on Silicon Substrates”, *Nano Lett.*, vol 7, pp. 3592–3595, June 2012, doi: 10.1021/nl301254z.

[7] A. Jönsson, J. Svensson and L. Wernersson, "A Self-Aligned Gate-Last Process Applied to All-III–V CMOS on Si," in *IEEE Electron Device Letters*, vol. 39, no. 7, pp. 935-938, July 2018, doi: 10.1109/LED.2018.2837676.

[8] More Moore report, international roadmap for devices and systems, 2017

[9] D. -H. Kim, J. A. del Alamo, P. Chen, Wonill Ha, M. Urteaga and B. Brar, "50-nm E-mode In0.7Ga0.3As PHEMTs on 100-mm InP substrate with fmax > 1 THz", *2010 International Electron Devices Meeting*, 2010, pp. 30.6.1-30.6.4, doi: 10.1109/IEDM.2010.5703453.

[10] D.-H. Kim, D. Alamo and A. Jesus, “30-nm InAs PHEMTs With f_T = 644 GHz and f_{max} = 681 GHz”, *IEEE Electron Device Lett,* vol. 31, no. 8, pp. 806-808, Aug. 2010, doi: 10.1109/LED.2010.2051133.

[11] A. Leuther, S. Koch; A. Tessmann; I. Kallfass; T. Merkle; H. Massler; R. Loesch; M. Schlechtweg, S. Saito and O. Ambacher, “ 20 nm Metamorphic HEMT with 660 GHz *f*T”, *Int. Conf. Indium Phosphide Relat. Mater., 23rd International Conference on Indium Phosphide and Related Materials*, pp. 1-4, May 2011

[12] M. Berg, O.-P. Kilpi, K.-M. Persson, J. Svensson, M. Hellenbran, E. Lind, and L.-E. Wernersson, “Electrical Characterization and Modeling of Gate-Last Vertical InAs Nanowire MOSFETs on Si,” in *IEEE Electron Device Letters*, vol. 37, no. 8, pp. 966-969, Aug. 2016, doi: 10.1109/LED.2016.2581918.

[13] D. Alamo, “Nanometre-scale electronics with III–V compound semiconductors” *Nature,* vol. 479, pp. 317–323, Nov. 2011, doi: 10.1038/nature10677.

[14] S. M. SZE, M.L. LEE, “Semiconductor Devices Physics and Technology”, John Viley & Sons, Inc., 2012, pp.171

[15] Markus Hellenbrand, “Electrical Characterisation of III-V Nanowire MOSFETs” doctor thesis of Department of Electrical and Information Technology in Lund University, May 2020, pp.23

[16] P. Olausson, L. Södergren, M. Borg and E. Lind, " Optimization of Near-Surface Quantum Well Processing", Phys. Status. Solidi A, vol 218, pp.2000720, Jan 2021, doi: 10.1002/pssa.202000720.
[17] John H. Davies, "*the physics of low-dimensional structures and quantum devices",* Cambridge University Press, 1997, pp.132
[18] A. M. Ionescu and H. Rjel, "Tunnel filed-effect transistors as energy-efficient electronic switches", *Nature*, vol. 479, pp. 329–337, Nov 2011, doi: 10.1038/nature10679.
[19] N.G.Tarr, D.J.Walkey, S.B.Hewitt and T.W.MacElwee, "Short-channel effects on MOSFET subthreshold swing" , *solid-state electronics*, vol.38, pp. 697-671, Mar. 1995, doi: 10.1016/0038-1101(94)00147-8.
[20] J. Q. Lin, T.-W. Kim, D. A. Antoniadis and J. A. del Alamo "A Self-Aligned InGaAs Quantum-Well Metal–Oxide–Semiconductor Field-Effect Transistor Fabricated through a Lift-Off-Free Front-End Process", *Appl. Phys. Express*, vol. 5, pp. 064002, May 2012, doi: 10.1143/APEX.5.064002
[21] T. -W. Kim, H.-M. Kwon, S. H. Shin, C.-S. Shin, W.-K. Park, E. Chiu, M. Rivera, J. I. Lew, D. Veksler, T. Orzali and D. H. Kim, "Impact of H_2 High-Pressure Annealing onto InGaAs Quantum-Well Metal–Oxide–Semiconductor Field-Effect Transistors With Al_2O_3/HfO_2 Gate-Stack," in *IEEE Electron Device Letters*, vol. 36, no. 7, pp. 672-674, July 2015, doi: 10.1109/LED.2015.2438433.
[22] J. Lin, X. Zhao, I. M. Clavero, D. A. Antoniadis and J. A. del Alamo, "A Scaling Study of Excess OFF-State Current in InGaAs Quantum-Well MOSFETs," in *IEEE Transactions on Electron Devices*, vol. 66, no. 3, pp. 1208-1212, March 2019, doi: 10.1109/TED.2019.2891751.
[23] J.Ajayan, D.Nirmal, P.Prajoon, J.Charles Pravin, "Analysis of nanometer-scale InGaAs/InAs/InGaAs composite channel MOSFETs using high-K dielectrics for high speed applications," *Int. Electron. Commun. (AEU)*, vol.79. pp. 151-157, Sep 2017, doi: 10.1016/j.aeue.2017.06.004
[24] Sheet resistance, from Wikipedia, https://en.wikipedia.org/wiki/Sheet_resistance
[25] handout in the lecture of semiconductor physics (on Convas)
[26] Transconductance, from the internet https://www.techtarget.com/whatis/definition/transconductance
[27] Erik Lind, personal communication, Feb 25th, 2022
[28] D. Gupta, M. Katiyar and D. Gupta, "Mobility estimation incorporating the effects of contact resistance and gate voltage dependent mobility in top contact organic thin film transistors," *Proc. of ASID*, 06, pp. 425-428, Oct, 2006
[29] MLA150 specification from the Internect, https://snfexfab.stanford.edu/guide/equipment/heidelberg-mla-150-heidelberg
[30] M.M. Frank, G. D. Wilk, D. Starodub, T. Gustafsson, E. Garfunkel, Y. J. Chabal, J, Grazul and D. A. Muller, "HfO_2 and Al_2O_3 gate dielectrics on GaAs grown by atomic layer deposition", *Appl. Phys. Lett.*, vol. 86, pp. 152904, Apr. 2005, doi: 10.1063/1.1899745.
[31] M. L. Huang, Y.C. Chang, C.H. Chang, Y. J. Lee, P. Chang, J. Kwo, T.B. Wu and M. Hong, "Surface passivation of III–V compound semiconductors using atomic-layer-

deposition-grown Al_2O_3," *Appl. Phys. Lett.*, vol. 87, pp. 252104, Dec 2005, doi: 10.1063/1.2146060
[32] R. Suzuki1, N. Taoka1, M. Yokoyama1, S. Lee1, S. H. Kim1, T. Hoshii1, T. Yasuda, W. Jevasuwan, T. Maeda, O. Ichikawa, N. Fukuhara, M. Hata, M. Takenaka1, and S. Takagi, "1-nm-capacitance-equivalent-thickness HfO_2/Al_2O_3/InGaAs metal-oxide-semiconductor structure with low interface trap density and low gate leakage current density," *Appl. Phys. Lett.*, vol. 100, Pp. 132906, Mar 2012, doi: 10.1063/1.3698095
[33] B.Brennana, M.Milojevicb, C.L.Hinkleb, F.S.Aguirre-Tostadob, G.Hughesa, and R.M.Wallaceb, "Optimisation of the ammonium sulphide (NH4)2S passivation process on In0.53Ga0.47As," *Applied Surface Science*, vol. 257, pp. 4082–4090, Feb. 2011, doi: 10.1016/j.apsusc.2010.11.179
[34] A. Alian， G. Brammertz, R. Degraeve, M. Cho, C. Merckling, D. Lin, W. E. Wang, M. Gaymax, M. Meuris and K. Heyns, " Oxide Trapping in the InGaAs–Al_2O_3 System and the Role of Sulfur in Reducing the Al_2O_3 Trap Density," in *IEEE Electron Device Letters*, vol. 33, no. 11, pp. 1544-1546, Nov. 2012, doi: 10.1109/LED.2012.2212692.
[35] W. K. Chong, E. F. Chor, C. H. Heng and S. J. Chua, "(Pd, Ti, Au)-based ohmic contacts to p- and n-doped $In_{0.53}Ga_{0.47}As$," *Compound Semiconductors 1997.* Proceedings of the IEEE Twenty-Fourth International Symposium on Compound Semiconductors, 1997, pp. 171-174, doi: 10.1109/ISCS.1998.711607.
[36] C. B. Zota, C. Convertino, D. Caimi, M. Sousa and L. Czornomaz, "Effects of Post Metallization Annealing on InGaAs-on-Insulator MOSFETs on Si," *2019 Joint International EUROSOI Workshop and International Conference on Ultimate Integration on Silicon (EUROSOI-ULIS)*, 2019, pp. 1-4, doi: 10.1109/EUROSOI-ULIS45800.2019.9041855.
[37] C.-Y.Chang, M. Yokoyama and S.-H. Kim, "Impact of metal gate electrodes on electrical properties of InGaAs MOS gate stacks, " *Microelectronic Engineering, vol.* 109, pp. 28–30, Sep. 2013, doi: 10.1016/j.mee.2013.03.086
[38] Raith EBL manual, https://raith.com/product/voyager/
[39] Four-terminal sensing, from Wikipedia, https://en.wikipedia.org/wiki/Four-terminal_sensing
[40] Chandra, H., S.W. Spencer, S. W. Oberloier, N. Bihari, J. Gwanuri and J. M. Pearce, "Open-Source Automated Mapping Four-Point Probe," *Materials* , vol. 10(2), pp. 110, Jan. 2017, doi: 10.3390/ma10020110
[41] $In_xGa_{1-x}As$ band gap, from the internet, https://www.batop.de/information/Eg_InGaAs.html
[42] Yoontae Hwang, Roman Engel-Herbert, Nicholas G. Rudawski, *and* Susanne Stemmer, "Effect of postdeposition anneals on the Fermi level response of $HfO_2/In_{0.53}Ga_{0.47}AsHfO_2$ gate stacks," *Journal of Applied Physics*, vol.108, pp. 034111, may 2010), doi: 10.1063/1.3465524
[43] band tail effect, from the Internet, https://www.researchgate.net/figure/Effects-of-band-tail-states-on-photoluminescence-a-The-normalized-spectral_fig8_346698015