Development of III-V RF Nanowire MOSFETs

Josefine Rost
tna13jro@student.lu.se
Anna Wagnström
tna13awa@student.lu.se

Department of Electrical and Information Technology
Lund University

Supervisor: Erik Lind
Co-supervisor: Fredrik Lindelöw

Examiner: Mats Gustafsson

June 14, 2018

Printed in Sweden
Tryckeriet i E-huset, Lund

Abstract

The silicon MOSFET is one of the most important components used in modern electronics. The pursuit to continue fulfilling Moore's law by scaling transistors to even smaller sizes have driven the development forward for CMOS technologies and new approaches have been necessary. These include changing the geometry of the devices for even higher performance, like creating 3D structures (finFETs or tri-gates) for better electrostatic control. With the aggressive scaling, power dissipation has become an increasing problem in CMOS applications. One way to combat this is by introducing a material in the channel with superior transport properties, like InGaAs.

As the development of MOSFET technology for CMOS applications continue, devices with excellent high frequency performance have been demonstrated. Thus, the MOSFET has also emerged as a possible alternative to the HEMT in high frequency applications. The desirable properties of an RF transistor differs somewhat from CMOS transistors, as stability and high gain are more important than size and power dissipation.

In this thesis, lateral InGaAs nanowire MOSFETs for RF applications with different geometries have been fabricated and characterised. The transistors exhibit a highest f_t and f_{max} of 150 GHz and 280 GHz respectively. The devices were fabricated in a cleanroom environment and measured in both DC and RF. Small signal modelling was performed using Matlab. A model of the devices was also made in COMSOL multiphysics and simulations were performed to investigate the effect of different geometries on the gate-source and gate-drain capacitance. The results form the simulation showed similar values to the fabricated devices. This indicates that the model was accurate.

The fabrication yielded very few working devices compared to what was fabricated. Some of the working devices showed performance compared to similar devices that have been demonstrated. Though a few devices show promising RF characteristics, the fabrication process needs to be refined.

Populärvetenskaplig sammanfattning

Om du skulle gissa på hur många komponenter det finns i din mobiltelefon, vad hade du gissat på då? 10? 100? 1 000 000? Skulle du tro mig om jag säger att svaret är runt 200 miljarder? För det är faktiskt så.

I din telefon och i all elektronisk utrustning runt omkring dig finns miljontals små elektroniska byggdelar som gör vårt tekniska samhälle möjligt. Snabba, smarta och kraftfulla datorer och mobiltelefoner är vi vana vid att ha med oss varje dag och det finns en konstant strävan till förbättring. Komponenten som har gjort denna tekniska utveckling möjlig är transistorn, som är en väldigt mångsidig byggsten. Trots dess minimala storlek på endast några nanometer i gatelängd (nu pratar vi om storlek tjockleken av ett hårstrå delat i 100 000 delar!) är transistormarknaden en miljardindustri. Gatelängd är ett mått som används för att jämföra storleken på olika transistorer, och motsvarar sträckan strömmen måste färdas i transistorn. Om det ska få plats 200 miljarder komponenter i en mobiltelefon behöver de vara rätt så små om du ska orka bära med dig den, inte sant?

Transistorn kan både användas som av/på-knapp i logiska kretsar och som förstärkare, beroende på hur den byggs. För hur konstigt det än låter att så små saker kan byggas så är det faktiskt sant! I väldigt rena miljöer, med stora och dyra maskiner, tillverkas transistorer med en rad komplicerade men väl beprövade metoder. Beroende på hur man designar geometrin och vilka material man använder i byggandet får transistorn olika egenskaper. Den vanligaste typen av transistor kallas *MOSFET* och är vanligtvis baserad på grundämnet kisel, vilket är superbra för logiska användningsområden. Men om man ändrar materialet till en sammansättning av andra grundämnen, och ändrar geometrin lite, kan man få en MOSFET som passar bra även till andra applikationer, så som förstärkare vid höga frekvenser.

I detta examensarbete har vi tillverkat transistorer för högfrekvensapplikationer med en gatelängd mellan 40 och 70 nanometer som uppvisar goda egenskaper vid höga frekvenser. För en transistor med 40 nanometer gatelängd uppmättes en högsta användningsfrekvens till 153 GHz för strömförstärkning och motsvarande 278 GHz för effektförstärkning. Transistorerna som tillverkades är MOSFETs baserade på en förening mellan indium, gallium och arsenik (InGaAs) och använder nanotrådar för att leda ström. Att använda nanotrådar och InGaAs är båda att föredra för att uppnå höga hastigheter, då elektroner rör sig mindre hindrat i denna kombinationen.

Målet var att jämföra hur olika geometrier skulle påverka prestandan, och även skapa en bild av hur transistorn kan förbättras för att nå högre frekvenser. Tre olika geometrier designades men bara två av dem gav resultat. Alla transistorer tillverkades i Lund Nano Lab och mättes sedan i både direktström (DC)- och radiofrekvens (RF)-miljö. Resultaten var mycket varierande eftersom många av transistorerna inte fungerade. Trots det har vi visat att transistorerna med de valda materialen och designen fungerar och visar lovande förutsättningar för att bli bättre!

Acknowledgements

We would like to thank our supervisors Erik Lind and Fredrik Lindelöw. Fredrik, thank you for the extreme patience, countless lab hours and unlimited positivity. Also, for the Kexchoklad. Thank you Erik for your guidance and support. We would also like to thank the remaining members of the Nanoelectronics group, as well as the LNL staff.

Dedicated to Sophie Wagnström.

Abbreviations & Symbols

Abbreviations

Al_3O_2	Aluminium Oxide
ALD	Atomic Layer Deposition
Au	Gold
BOE	Buffered Oxide Etch
CMOS	Complementary Metal Oxide Semiconductor
DC	Direct Current
DIBL	Drain Induced Barrier Lowering
EBL	Electron Beam Lithography
FinFET	Fin Field-effect Transistor
H_3PO_4	Phosphoric Acid
H_2O_2	Hydrogen Peroxide
HCl	Hydrochloric Acid
HEMT	High Electron Mobility Transistor
HfO_2	Hafnium Dioxide
HSQ	Hydrogen Silsesquioxane
InGaAs	Indium Gallium Arsenide
InP	Indium Phosphide
IPA	Isopropanol Alcohol
LRRM	Load-Reflect-Reflect-Match
MOCVD	Metalorganic Chemical Vapor Deposition
MOSFET	Metal-oxide-semiconductor field-effect transistor
$(NH_4)_2S$	Ammonium Sulfide
Pd	Palladium
PMMA	Poly(methyl methacrylate)
RF	Radio Frequency
SEM	Scanning Electron Microscopy
Ti	Titanium
TMAH	Trimethylammonium hydroxide

Symbols

C_{gd}	Gate-drain Capacitance
C_{gg}	Total Gate Capacitance
C_{gs}	Gate-source Capacitance
C_{sd}	Source-drain Capacitance
C_{ox}	Oxide Capacitance
f_T	Cutoff Frequency
f_{max}	Maximum Oscillation Frequency
g_m	Transconductance
g_d	Output conductance
h_{21}	Current Gain
L_g	Gate Length
MSG	Maximum Stable
MAG	Maximum Available Gain
R_G	Gate Resistance
R_D	Drain Resistance
R_S	Source Resistance
U	Unilateral Gain
V_{gs}	Gate Voltage
V_{ds}	Drain Voltage
V_T	Threshold Voltage
W_m	Mesa width

ω	Angular Frequency
μ	Electron Mobility
λ	Mean Free Path

Table of Contents

List of Figures

List of Tables

Chapter 1
Introduction

Since World War 2, the evolution of electronics has been sky rocketing. The war introduced the need for logical circuits that was used to break enemy codes, as well as good communication systems [1]. This led to a rapid development of electronic devices and with that a higher demand for computing power. The first electrical computer used vacuum tubes as logical components [2]. The problem was that these were both fragile and had a short life span, so something better was needed for the development to continue. This is where the transistors and the third generation of computers came in [1]. The transistors were smaller, more reliable and produced less heat. The computing power increased rapidly and has ever since [3].

The MOSFET (metal-oxide-semiconductor field-effect transistor) has been one of the most important building blocks of modern electronics since it was first invented in 1959 at Bell Labs [4]. MOSFETs today are used for a wide number of applications, ranging from digital CMOS logic to wireless communication [5]. The request for higher performance and more complex circuits challenges the boundaries of the conventional Si MOSFET [6]. To meet these requirements, new approaches need to be taken. One way of increasing the performance of the FET-devices is to change the geometry and material composition of the device, to improve performance without changing the fundamental physics [5]. By changing the material from silicon to another semiconductor, and tweaking the geometry to a non planar device, the performance and speed of the devices can be greatly increased.

As the development of the MOSFET for CMOS applications pushes on, another possible application for the MOSFET has emerged in the RF market. With its scaling potential, good high frequency performance and excellent transport properties, a RF MOSFET that can compete with the HEMT has become a reality [5].

In this diploma work, a take on the III-V nanowire MOSFET is studied and further developed in the quest of reaching even higher performing RF components.

Chapter 2

MOSFET

2.1 Background

The first patent for a field-effect-transistor was filed by Julius Edgar Lilienfeld as early as 1926 [7]. After this, researchers at Bell Labs began investigating the properties of semiconductor devices. In 1947 the first transistor was presented by John Bardeen, Walter Brattain and Walter Shockley, a three point contact device with amplification [8]. Shortly after, in 1947, Shockley published a paper describing the p-n junction and the bipolar junction transistor [9]. Bardeen, Brattain and Shockley were in 1956 awarded the Nobel prize in physics "for their researches on semiconductors and their discovery of the transistor effect" [10]. The bipolar transistor is yet today one of the most important semiconductor devices. This discovery was followed by the creation of the metal-oxide-semiconductor field-effect transistor (MOSFET) by Kahng and Atalla which was reported in 1960. This development was possible only after Atalla had discovered a technique for thermal growth of Silicon oxide [4].

The MOSFET and integrated circuits based on the MOSFET make up about 95 % of the semiconductor device market [11]. This popularity was mainly possible due to the rise of the digital computer around the same time. Integrated circuits based on MOSFETs had superior performance compared to other technologies in terms of fabrication cost and scaling. MOSFETs could be scaled down without losing performance, allowing for better performance with lower voltages, lower power dissipation and higher speed. It was this scaling ability that allowed for the rapid growth in computing power as the number of devices per area could be increased for a lower device cost [4].

This transistor downscaling is often referred to as Moore's law from the predictions made by Gordon Moore in his 1965 paper [12]. The publication predicts an exponential increase in the number of transistors per unit area and this trend was followed for many decades. However, as the transistors became smaller, physical limitations challenged the progress. Structures cannot be downsized indefinitely, as eventually they would reach the atomic levels. Simple downscaling is no longer viable, and thus other methods must be used to ensure the continued improvements in transistor performance.

2.2 MOSFET Operation

The MOSFET is a three-terminal device. There are two different types, n-type and p-type MOS [11]. Here, only the nMOS will be described as it is the one fabricated in this work. It consists of three metal contacts, gate, source and drain, a p-type semiconductor substrate and two n^+-doped regions. The pMOS has the same structure, but with an n-type substrate and two p^+-doped regions [11]. The abbreviation MOS stands for metal-oxide-semiconductor and describes the geometry of the device, shown in figure 2.1. The metal gate contact is separated from the semiconductor channel by an insulating oxide layer. The device is controlled by the field-effect, an electric field that is formed over the oxide as a voltage is applied on the gate. This allows the conductivity of the channel to be controlled by the voltage applied to the gate contact. This is the basic principle of MOSFET operation [11].

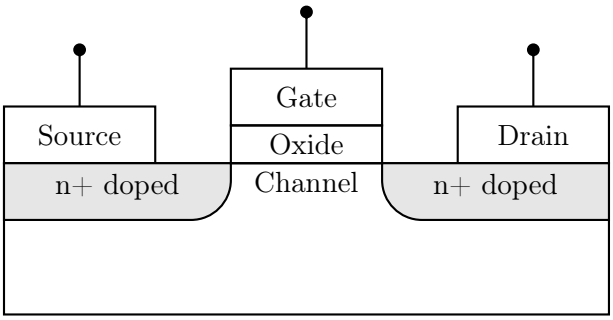


Figure 2.1: Schematic of a planar n-type MOSFET structure.

Without a voltage applied to the gate there is an energy barrier preventing charge carriers from moving from source to drain formed by the two p-n junctions between n^+-doped source and drain and p-type channel. This is illustrated in figure 2.2. When a bias is applied to the gate it creates an inversion layer between the two n^+ regions. This allows for a current to flow from drain to source, if a bias is applied between the source and drain (V_{DS}). This is shown in figure 2.3.

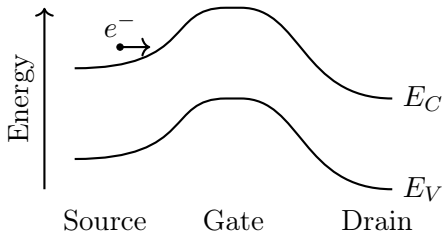


Figure 2.2: MOSFET band structure in the off-state

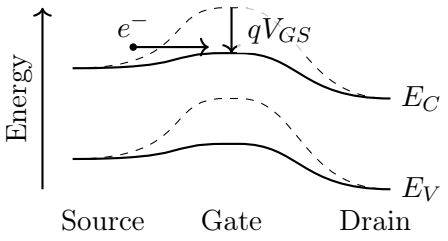


Figure 2.3: MOSFET band structure in the on-state

The bias applied between the gate and source is referred to as V_{GS}. The threshold where current starts flowing is called V_T. The transistor is said to be

in the off-state when V_{GS} is below V_T, and similarly in the on-state when V_{GS} is above V_T.

As stated in section 2.1, improving performance by only scaling down MOSFETs is no longer feasible. Not only is the scaling limited by the atomic scale, but transistors with short gate length also show so called short-channel effects [11]. This means that the device performance is degraded by e.g. drain induced barrier lowering (DIBL), when the barrier in figure 2.2 is affected by the increase in drain voltage and thus the conductance of the channel is affected [11]. With a high drain voltage, it will no longer be possible to turn off the channel.

This can be counteracted in different ways, e.g. by changing the geometry or changing the material. This is described below.

2.3 III-V Nanowire MOSFET

The electrostatic control can be improved by changing the geometry of the gate structure. Instead of only having the gate control the channel from one direction, a multi-gate device can be used, also known as FinFETs or tri-gate devices [5]. Then, the channel is no longer part of the bulk as in the planar MOSFET but a raised "fin". The gate is placed around the channel, providing much better electrostatic control as the contact area is increased without increasing the gate length. Such a device with a nanowire channel is shown in figure 2.4, where the gate stack is shown and in figure 2.5 where the source, gate and drain layout is shown.

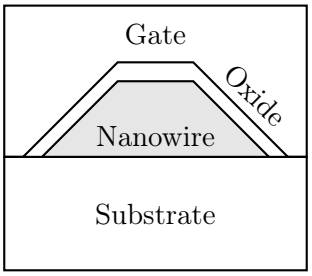


Figure 2.4: Nanowire MOSFET structure

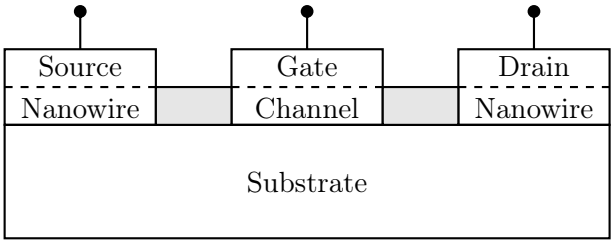


Figure 2.5: Side view of the nanowire MOSFET structure

2.4 III-V Materials

Electronics today are mainly silicon based and the combination of silicon and its native oxide remains one of the key material combinations in transistor development [11]. Silicon is abundant, and therefore cheap, and SiO_2 is a stable oxide. These characteristics has made it the foundation for all CMOS technology, and its characteristics and performance is well researched.

However, silicon transistors have some drawbacks. With the aggressive scaling of transistors, power dissipation has become one of the key issues with CMOS performance. To increase performance by scaling, operating voltage needs to be

decreased resulting in a decrease in speed. One solution to this is introducing a material in the channel with a higher carrier mobility than silicon [13]. This is where the III-V compound semiconductors come in.

III-V materials are compounds of elements from group III and V of the periodic table, such as InAs, GaAs, InP and AlAs. They also include the ternary alloys, like InGaAs. These compounds show a wide variety of characteristics and some have unique applications in both optics and electronics [13]. The ternary III-V alloys have another advantage, as properties like mobility, band gap and lattice parameter can be fine-tuned by changing the composition [14]. A few III-Vs show especially high electron mobility, namely InAs and InGaAs, which have demonstrated mobilities up to 10 times that of silicon [13].

A restriction to the development of InGaAs or GaAs MOSFETs has been the poor nature of their respective native oxides. The interface formed when for instance GaAs is exposed to oxygen is highly unstable and has a high interface state density [13], thus impeding the effective modulation of the channel. A major breakthrough was made in 2003, when the first GaAs MOSFET with an ALD deposited Al_2O_3 layer was demonstrated [15].

2.5 RF Applications

III-V materials are already successfully used in high frequency applications, for instance in high electron mobility transistors (HEMTs) [11]. The HEMT is a heterojunction device that differs from the MOSFET by using a high band gap semiconductor instead of an oxide barrier [11]. HEMTs also use modulation doping, where that barrier layer is doped except for a narrow region closest to the channel and the channel is undoped. Carriers from the barrier layer will move to the channel, increasing the carrier concentration. This will result in an increased carrier mobility compared to if dopant atoms would sit in the channel, thanks to the increased carrier concentration without any impurity scattering in the channel [11]. III-V HEMTs with cutoff frequencies above 600 GHz and maximum oscillation frequencies above 1 THz have been demonstrated.

However, to further increase the high frequency performance of HEMTs, gate capacitance would need to be increased since the transconductance depends on the gate capacitance. Due to the small band gap of the barrier layer (compared to an insulator), scaling is difficult without significantly increasing the gate leakage current. MOSFETs, which use an oxide barrier that allows for much more aggressive scaling, could therefore be a viable option for high frequency applications [5].

Chapter 3
Fabrication

Manufacturing a transistor from scratch takes time and involves many processing steps. This chapter describes the geometry and fabrication of the InGaAs nanowire MOSFETs processed in this diploma work. Three different device geometries were fabricated, two of which have never been fabricated by the Nanoelectronics group before, on a total of six samples. Five of the samples were measured in DC, but only the best performing sample was studied at high frequencies and analysed in chapter 6. The different geometries are called A, B and C and are introduced in the section below.

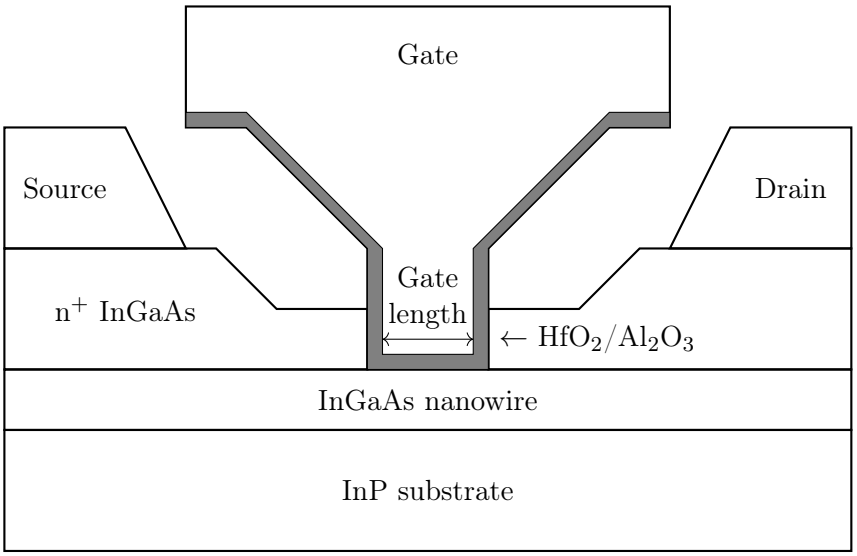


Figure 3.1: Device schematic

3.1 Device Design

This diploma work builds on previously processed devices by the Nanoelectronics group, here referred to as type A. Several papers have been published on lateral InGaAs nanowire MOSFETs by the group with very promising results [5]. Nanowires

exhibit a low crystal defect density, which allows for low surface scattering and thus higher electron mobility. By placing the channel on top of the substrate, the electrostatic control is increased significantly compared to a conventional buried channel, which also is a perk of using nanowires. This makes nanowire channels an interesting challenger to the planar MOSFET and other fin structures. A schematic side view of the devices is shown in figure 3.1. A top view of a fabricated *A*-device can be seen in figure 3.2. Here, source drain and gate is marked, as well as the structure called mesa, which is described later in this chapter.

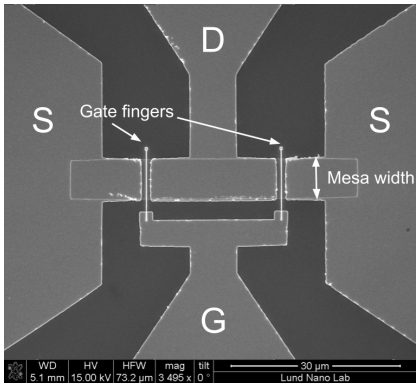


Figure 3.2: SEM overview image of a finished *A*-device with source (S), drain (D), gate (G) and mesa width marked

For these devices, the nanowires are grown parallel in an array with a density of 14.3 nanowires per μm gate width. The direction of the nanowires in relation to the InP substrate creates {110} crystal facets on the nanowire sides. The source and drain contacts are grown on each sides of the arrays, resulting in {111}-facets of the InGaAs contact layers. The gate is placed over the arrays with a 45° angle with respect to the nanowires. How the nanowires are orientated under one of the gate fingers are visualized in figure 3.3, together with source and drain contacts. The nanowires are show in grey.

To explore new grounds and hopefully achieve higher performance, the geometry of these previously fabricated devices were altered.

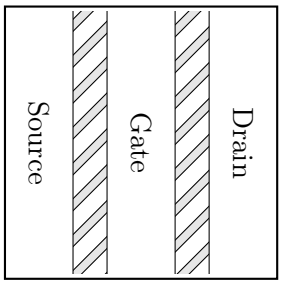


Figure 3.3: Device type A.

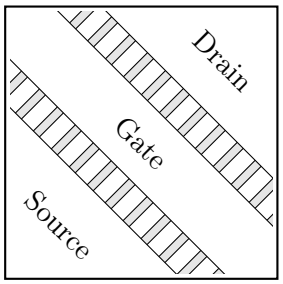


Figure 3.4: Device type B.

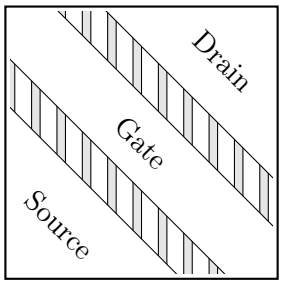


Figure 3.5: Device type C.

The first alteration was to rotate the gate 45° with respect to the *A*-gate. This makes the gate orthogonal to the nanowires. This is visualized in figure 3.4. The rotation of the gate, and thus source and drain contacts, results in new crystal facets of the contact layer. This device design is expected to create {110} crystal facets on the contact layer as well. The nanowires stay exactly the same as in type *A*. The rotation of the gate results in a higher density of nanowires in the channel per gate width, which is clearly visible in figure 3.4. The new nanowire density becomes $14.3 \cdot \sqrt{2} = 20.2$ per μm gate width. A higher nanowire density should result in higher currents and thus higher transconductance, which is a desired RF metric. This could result in lower electron scattering and thus higher transconductance. These devices are called type *B*. In figure 3.6 a top view of a finished *B*-device can be seen. Here, the contacts are marked, as well as the mesa and effective mesa widths.

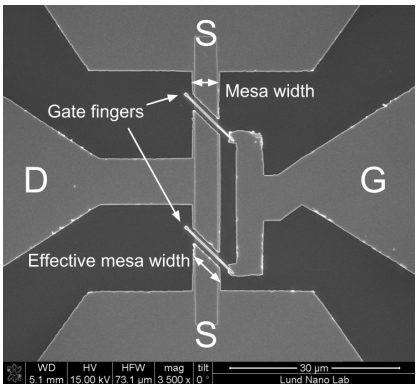


Figure 3.6: SEM overview image of a finished *B*-device with source (S), drain (D), gate (G), mesa width and effective mesa width marked. This layout also applies to the *C*-devices.

The second alteration and third geometry that was examined used the gate rotation from *B*, but with the nanowires rotated 45° as well. This layout can be seen in figure 3.5. The contacts are the same as in *B*. However, this new nanowire orientation should result in {111}-facets with a 56° angle, the same facet as for the contact layer in *A*. The nanowire density is the same as for *A*, but with new crystal facets. The current from drain to source has a straight path through the nanowires, without any bends that are very distinct in *A* and *B*. The interesting analysis for this device type would be if there is any difference in performance for the different crystal facets. One could argue that steeper walls would result in a larger cross sectional area and thus lower resistance in the channel. Another difference is that there is probably a different interface trap density, D_{it}, of the {111} facets compared to the {110}. These devices are called type *C*.

To be able to measure the DC and RF performance, pads connected to the contacts are needed. The finished pad layout can be seen in the optical microscopy images in figure 3.7 and 3.8. All three transistor types were designed with two gate fingers connected to one common gate contact, two source contacts and one drain contact. The layout of the gate pad was the result of a COMSOL simulation

presented in chapter 4.

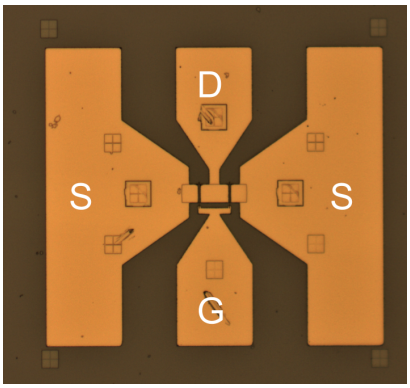


Figure 3.7: Pad layout of device type *A*

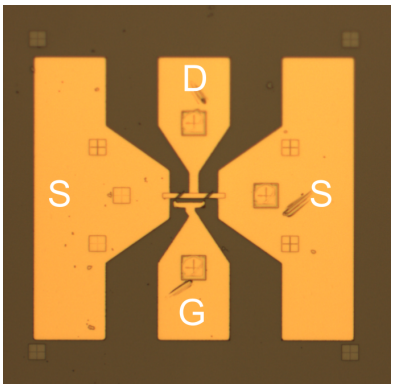


Figure 3.8: Pad layout of device type *B* and *C*

3.2 Experiment setup

To be able to compare a maximum number of parameters of the three different device types introduced above, the diversity of the components were planned before starting the processing. All three type of devices were designed with gate lengths ranging from 20 nm to 70 nm. The gate length is defined in figure 3.1. Beyond the gate length variations, a structure called mesa was also varied. The mesa is marked in figure 3.2 and 3.6. The width of the mesa sets the gate width and hence number of nanowires. By changing the mesa width, the performance can be compared when the number of nanowires are increased. The relation between mesa width and amount of nanowires in the channel for the three different types of devices is presented in table 3.1 below. For the *B* and *C* devices, it is more clear to compare the effective mesa width. Since the gate is rotated 45° related to the mesa, instead of being orthogonal, the effective gate width becomes longer with a factor of $\sqrt{2}$.

Table 3.1: Mesa widths

Device type	Mesa width (μm)	Effective mesa width (μm)	# Nanowires
Device *A*	7	7	100
	14	14	200
Device *B*	3.5	5	100
	5	7	140
	7	10	200
Device *C*	5	7	100
	7	10	140

Beyond gate and mesa geometry variations, different EBL-patterns were also compared. The first is the pattern of the nanowires, where two different exposure geometries were tested. One used rectangles (R) to define the nanowire array, and the other used six parallel single pixel lines (S). The other variation was the gate exposure. The first gate type used two single pixel lines (2L) to define the gate. The other used six single pixel lines (6L) to achieve a wider gate. This is further discussed in the corresponding sections below, and the results are presented in chapter 6.

3.3 Processing

The different process methods that were used to build the devices are briefly described in Appendix A. Some process steps are recurrent and are therefore described in less detail after the first mention. The processing described below is the general process. The real processing deviates from the general as described in section 3.3.9 in the end of this chapter.

3.3.1 Sample preparation

From a 4" InP (100) wafer, 12x10 mm samples were cut using a scriber. The substrates were cleaned in acetone for 15 seconds followed by a longer soak in acetone for 5 min and finally a 60 s soak in Isopropyl Alcohol (IPA).

3.3.2 Nanowire channel

The lateral nanowires, which makes up the channel of the MOSFET, were the first structures to be processed. The samples were pre-baked on a 200°C hotplate for 2 min to remove any water or other liquids that had been formed on the sample, to increase the adhesion of the photoresist. After cooling for one minute, the samples were placed on a spinner. Hydrogen silsesquioxane (HSQ) was used as photoresist and was diluted with Tetramethylammonium hydroxide (TMAH) 1:3 to achieve a thinner resist layer and optimize resolution. The samples were spun on a spin coater at 3000 rpm for 60 s, with a ramping of 1500 rpm/s. To increase adhesion of the HSQ to the InP substrate, the sample was baked on a 200°C hotplate for 2 min. The patterning was made using EBL. The aperture was set to 30 μm, the voltage to 50 kV and the step-size 2 nm. On two of the samples, two different exposure geometries were tested to see if there were any difference in edge roughness. The first geometry used rectangles to define where the nanowires should grow. The second exposure geometry used six single pixel lines with a higher dose than for the rectangle. This dose was also varied, increasing from left to right on the sample to see which dose gave the best result. After exposure, the resist was developed using TMAH for 90 s heated to 37°C to reduce edge roughness. When HSQ is exposed, it transforms to SiO_2, and can sustain high growth temperatures.

The $In_{0.85}Ga_{0.15}As$ nanowires were grown using metalorganic chemical vapour deposition (MOCVD) at 500°C. This step is critical for the device performance since it limits the channel mobility, which is important for RF applications. Therefore all growth was outsourced to Lasse Södergren, PhD student in the Nanoelec-

tronics group. The InGaAs only grows in the *trenches* between the HSQ strips created by the EBL pattern, not on the HSQ itself. When the HSQ is removed, it leaves lateral InGaAs nanowires. HSQ removal was done using 1:10 diluted Buffered Oxide Etch (BOE) for 2.5 min, followed by 60 s H_2O and 60 s IPA. To decrease the size and edge roughness of the nanowires, five cycles of digital etch followed. This was done by placing the samples in 1:5 Hydrochloric acid (HCl):H_2O for 12 s and ozone cleaning for 8 minutes.

3.3.3 Contact regrowth

After the nanowire channel was processed, the heavily doped drain and source contacts were added according to the images in section 3.1. To prevent the contact layer from growing over the nanowires, a 70 nm strip of HSQ was added covering the nanowire array. This strip is called a dummy-gate. Resist was deposited in the same way as in section above 3.3.2, with the only difference that undiluted HSQ was used to achieve a thicker resist layer. The patterning was made using EBL, this time with an aperture of 40 μm, a step size of 8 nm and an area dose of 1500 $\mu C/cm^2$.

After exposure, the resist was developed for 2 min in TMAH. 30 nm thick, heavily n-doped, InGaAs contacts were then grown on the samples using MOCVD.

A schematic of this step can be seen in figure 3.9. The HSQ shows where the gate will be placed in a later process step.

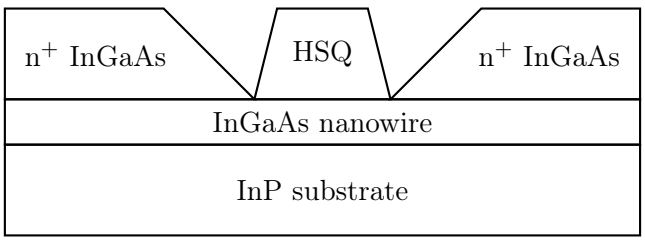
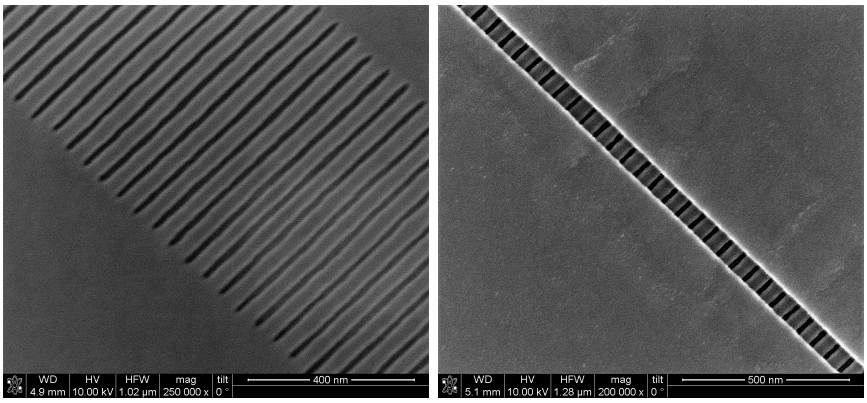


Figure 3.9: Contact growth schematic with HSQ dummy gate

The exposed HSQ was removed using BOE 10:1 for 2.5 min. In figure 3.10 the nanowire array of a *B* device before and after the contact regrowth can be seen.



a) b)

Figure 3.10: A B-type device a) before and b) after contact regrowth

3.3.4 Ledge & sacrificial InP layer

Following the contact regrowth was the ledge and InP sacrificial layer. The InGaAs ledge creates a lower contact resistance to the the channel. The sacrificial InP layer creates an air spacer between the gate and drain/source contacts, decreasing the capacitance between contacts. Air is the most desirable spacer material, since it has a low relative permittivity of 1. The resist was spun in the same way as for the contact regrowth. EBL was used once again for making the pattern, with the same settings as the previous step. Development was done for 2 min in TMAH. After development, a 8 nm heavily doped InGaAs ledge was grown, which is shown in figure 3.11.

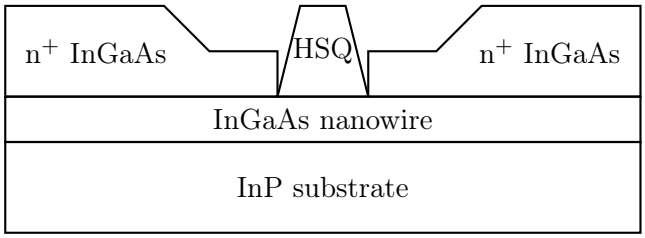


Figure 3.11: Contact ledge growth schematic with second HSQ dummy gate

After the ledge growth, the sacrificial InP layer was grown. The exposed HSQ was removed using BOE (10:1) for 2.5 min.

3.3.5 Mesa Etch

The mesa sets the gate width and thus how many nanowires that will be present in the channel. This step also makes sure that the drain and source pads are not short-circuited by the contact layer. The EBL preparations was the same as

can be seen in section 3.3.3 and 3.3.4. The exposure was done using the same settings as above. This time a wide rectangle covering both gate fingers was exposed. Development was done using TMAH soak for 2 min. The sacrificial InP outside the mesa was then etched by using 5 s 1:1 HCl:H_2O. Then the InGaAs was etched by using 1:1:25 dilution of H_3PO_4:H_2O_2:H:$_2$O for 25 s. The sacrificial InP, InGaAs ledge, InGaAs contact layer and nanowires that were not covered by the mesa structure was thus removed. Finally, the exposed HSQ was removed in BOE (10:1) for 3 min.

3.3.6 Oxide deposition

The gate oxide is one of the most important parts of a transistor. It separates the gate from the channel and prevents leakage. To get a better device performance, high-κ materials are often used as gate oxide. This oxide is deposited with Atomic Layer Deposition (ALD), which uses precursor gases at elevated temperatures. Before depositing the oxide, the native oxide needs to be removed and the samples need to be passivated. The native oxide is the natural oxide that forms on the surface of the sample when it is exposed to air. Passivation is a chemical process that makes the surface unable to react with air. Removal of the native oxide was done with 10 min of ozone cleaning. The samples were passivated by 20 min soak in $(NH_4)_2S$:H_2O 1:1. The oxide, Al_2O_3/HfO_2, was deposited using ALD. First, the sample was self cleaned with Al pulses in the ALD chamber. The Al_2O_3 was deposited using TMAl and H_2O as precursors for five cycles. This oxide is used to improve the channel-oxide interface. The *actual* oxide, HfO_2, was deposited using TDAHf and H_2O as precursors for 35 cycles. HfO_2 has a significantly higher dielectric constant (κ) of 25 compared to Al_2O_3 with κ=9 [16].

3.3.7 Gate

On top of the gate oxide is the gate contact. To lower the gate resistance, a T-gate was used. The large bulk metal of the gate increases the cross sectional area and thus reduces the resistance. To achieve the T-shape a double layer resist was used. The upper resist is exposed by the (primary) electron beam from the EBL, while the lower resist is exposed by both primary and secondary electrons. Thus, the upper resist needs to be more sensitive to electrons than the lower one since the number of exposing electrons is fewer.

PMMA 950 A2, which is a positive photo resist, was spun on the samples at 6000 rpm for 45 s. The samples were baked on a 180°C hotplate for 90 s. Another layer of positive resist, MMA EL6, was spun on top of the PMMA 950 A2 layer at 2000 rpm for 45 s. The gate contacts were patterned with EBL, using a 40 μm aperture, an area dosage of 790 $\mu C/cm^2$ and a 2 nm step size. The resists were developed in MIBK:IPA 1:3 for 90 s. The gate exposure was done in two different sizes, one using two single pixel lines (2L) and one using six single pixel lines (6L). The larger configuration was mainly to avoid alignment errors, since the gate was hard to align to where the dummy gate had been. To remove resist residues, the samples were placed in a plasma asher for 7 s at 5 mbar oxygen pressure. Gate metal was deposited using evaporation under sample rotation. The rotation

ensures an even metal layer. First 10 Å Titanium (Ti) was deposited, followed by 50 Å Palladium (Pd) and 800 Å Gold (Au). The unwanted metal was removed using an acetone soak for 20 min, with assisted lift-off using pipettes. The gate structure is shown in figure 3.12 below. The T-shape of the gate can clearly be seen, as well as the nanowires in the channel.

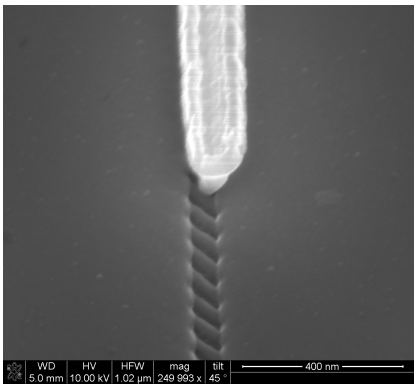


Figure 3.12: A 52° tilted SEM image of a T-gate

3.3.8 Source, gate & drain pads

To be able to measure the performance of the device in probe stations, pads for the contacts are needed. The sample was pre-baked on a 200°C hotplate for 2 min. PMMA 950 A8 was spun on the samples at 4500 rpm for 45 s with a ramping of 1500 rpm/s. The resist was baked on a 180°C hotplate for 2 min. The pad pattern was made using EBL with a 60 μm aperture, an area dose of 790 μC/cm^2 and a step size of 8 nm. The resist was developed in TMAH for 90 s and the resist residues were removed using 30 s plasma ashing at 5 mbar oxygen. The pad metal was deposited using evaporation. 100 Å Ti was followed by 100 Å Pd and 2000 Å Au. The lift-of was done with acetone using pipettes. Some final sample cleaning was done with plasma ashing for 35 s at 5 mbar oxygen pressure, and a final IPA soak for 60 s.

3.3.9 Samples

Six different samples were processed during this thesis. Below is a description of each sample deviations from the general process method.

Sample I follows the general processing with the exception that only 2L gates were processed. However, the HSQ strip that is supposed to protect the nanowires before sacrificial InP and contact regrowth was hard to align. This EBL-step, which involved resist spinning, exposure and development, was remade five times.

Sample II follows the general processing except some small deviations. This sample also lacks the 6L gate, and only has 2L. Just like sample I the

alignment of the third EBL step was remade several times to make the alignment acceptable. The evaporation of Ti in the source and drain pads was not successful for this sample. This is due to that there were two-Ti boats in the evaporator, which would have required double evaporation current during deposition.

Sample III and the following samples do not have any C-type devices since the quality of the "C-nanowires" in sample I and II was very poor after the nanowire growth. To avoid remaking the same EBL-step several times (as in the samples described above), the HSQ from the second EBL-step was etched back instead of removed. This applies to the following samples as well. To achieve different gate lengths, the HSQ-strip that was deposited before the contact regrowth (section 3.3.3) was altered in length between 20 and 70 nm.

Sample IV has several deviations from the general process. The sample lacks InGaAs ledge. The etching of sacrificial InP and InGaAs contact layer was not successful and required repeating etching steps. Because of this extra etching, the gate contact metal was not thick enough to make a contact with the gate finger. To fix this, an extra EBL and evaporation step was added to deposit a thick layer of metal at the finger-contact interface.

Sample V has the same deviations as sample III, plus that the digital etch of the nanowires was skipped to minimize the risk of etching down all the InGaAs. This step was skipped for the last device as well. To avoid any risk of separation between the gate contact and the gate finger, the thickness of the metal contact layer was increased. The Ti layer was increased from 10 to 100 Å, the Pd layer from 50 to 100 Å and the Au layer from 900 Å to 2000 Å.

Sample VI did not have the ledge or InP sacrificial layer, and is a DC reference sample for sample V. On this sample, the source and drain contacts were added before the gate-oxide and gate. To avoid an extra EBL-step the gate fingers and contacts were added together.

As was mentioned in the beginning of this chapter, the road from substrate to finished devices is complicated and takes time. The human factor as well as limitations of the processing methods will lower the yield and performance of at least some devices. Therefore one can not expect that the results always will correspond to the expected values.

Chapter 4

Simulations

The software COMSOL Multiphysics was used to model different parts of the device and simulate capacitance performance. All simulations were performed using the AC/DC module, doing a stationary analysis of electrostatics to calculate capacitance.

The AC/DC module uses FEM (finite element method) and BEM (boundary element method) to solve Maxwell's equations [17]. FEM is a numerical method of solving large mathematical problems and is often used in engineering. The idea is to divide the problem into smaller sections (finite elements) and solving the equations for each part. These sub-solutions then constitute an approximative solution for the entire problem when put together.

The relative permittivity of the materials used in the simulations is shown in table 4.1.

Table 4.1: Material parameters used in simulations

Material	Relative permittivity
InP	12.5
InGaAs	14
Oxide	20
Air	1

4.1 Pad capacitance

Simulations were made to investigate the capacitance between the gate and the drain pad. A simple model was created where both gate and drain contacts were modelled as 20 by 20 μm blocks with height 0.1 μm. The 2D model geometry is shown in figure 4.1.

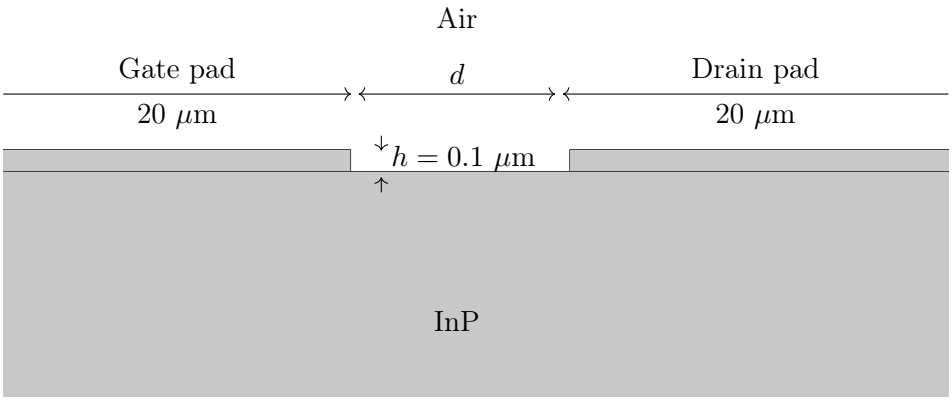


Figure 4.1: Model geometry for the pad capacitance simulation

This 2D model was simulated with an out-of-plane thickness of 20 μm so that the pads were 20x20 μm. The distance d between the pads was swept between 1 and 20 μm and the results of this simulation is shown in figure 4.2.

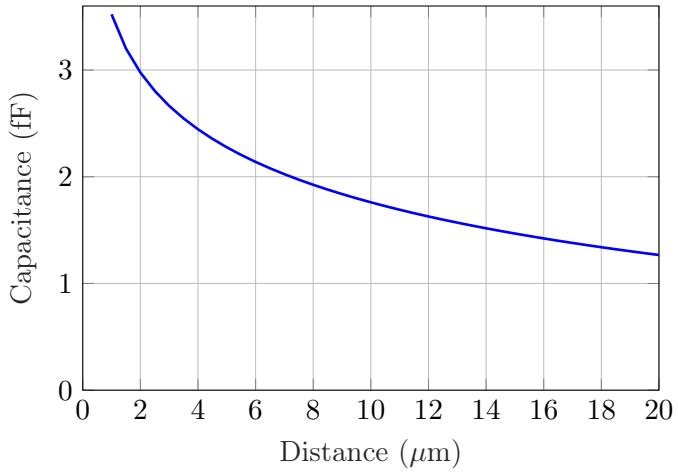


Figure 4.2: Gate to drain pad capacitance.

As can be seen by this simulation, the capacitance is quite significant. At 1 μm distance, the capacitance is more than 3.5 fF. To decrease this contribution, the gate contact was redesigned. The distance between gate and drain was increased, and smaller sections of the gate contact drawn out to ensure that gate resistance would not be affected by the change. The finished gate contact structure can be seen in figure 4.3 below where the added extensions are marked.

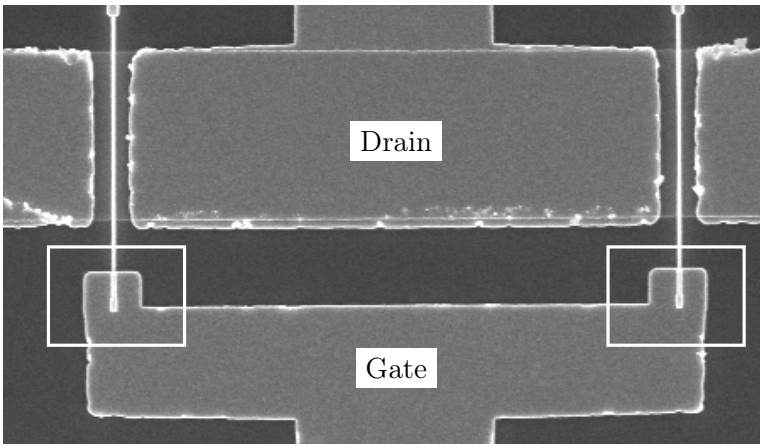


Figure 4.3: Finished gate contact layout with extended gate pad sections

4.2 Nanowire capacitance

A simple transistor model was created consisting only of a single nanowire and the gate terminal. The purpose of this model was to determine a minimum capacitance between the gate and the nanowire, without the contacts, ledges or other elements present. The model geometry is shown in figure 4.4 and 4.5 where the oxide is shown in blue.

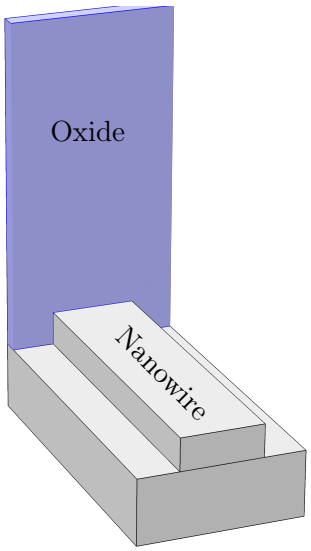


Figure 4.4: Side view of nanowire model geometry

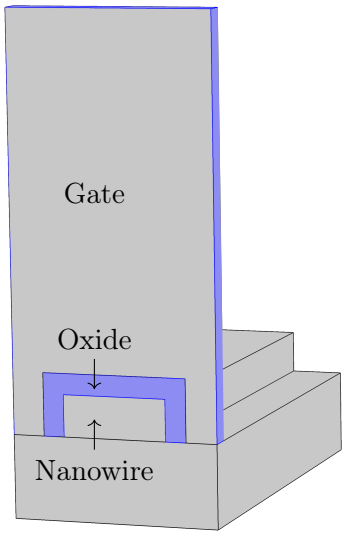


Figure 4.5: Gate view of nanowire model geometry

The nanowire shape was approximated as rectangular, though not entirely true this shape variation was not estimated to have a great impact on the result. The nanowire had the dimensions 25x10 nm and the oxide thickness was set to 5 nm.

To simulate the capacitance between the gate and the nanowire using only electrostatics, some approximations must be made. Here, the nanowire and the gate are both modelled as metallic, with one terminal on the gate surface shown in figure 4.5 and the other terminal on the surface of the entire nanowire.

The width of the model, and thus the distance between the nanowires, was swept to investigate how much it affects the capacitance per nanowire. A model width of 35 nm corresponds to a distance between the nanowires of 10 nm. A periodic boundary condition was put on the side edges of the model, since the actual devices consist of multiple nanowires. All other dimensions were kept constant. The result is shown in figure 4.6. Note that the capacitance is given in aF and is not normalized with gate width, which means the capacitance is shown per nanowire.

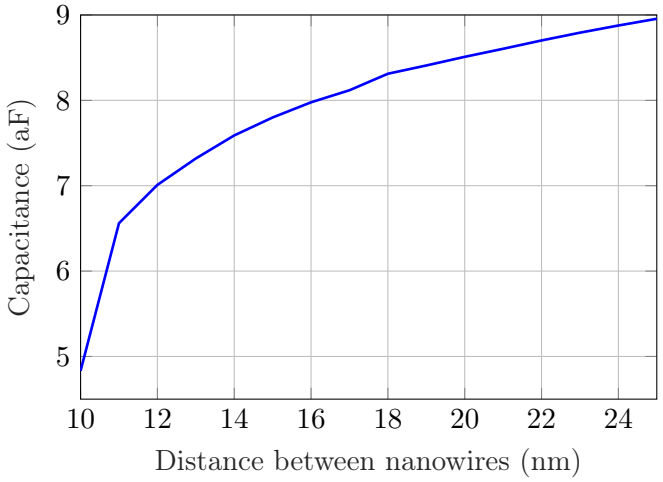


Figure 4.6: Parasitic capacitance per nanowire

As would be expected, the capacitance per nanowire is lowest when the distance between the nanowires is small. Since transconductance should scale with the number of wires, the best performance in terms of maximizing g_m and minimizing the parasitic capacitance would be achieved with wires as close together as possible.

The much lower value at 10 nm distance could be because the oxide is as thick as the distance between the wires, so there is no gate in contact with the substrate. This means that some of the parasitic capacitance to the sides of the nanowire is eliminated. This is however not a desirable design of the channel, since it would result in a channel without the tri-gate structure.

At 15 nm nanowire spacing, the capacitance is 7.8 aF corresponding to a normalized value of 0.2 fF/μm. By normalising per μm, a value is obtained that can be compared to the other simulations as well as the measured values. This

normalisation corresponds to a normalisation of a devices gate-source or gate-drain capacitance by total gate width.

4.3 Device capacitance

The model was extended to be as similar to the processed devices as possible. Since only an electrostatic simulation was performed, the whole contact area with n^+ contact and ledge are grounded while the gate domain is biased. This means that the highly doped InGaAs is modelled as metallic, which is an approximation. This results in slightly higher capacitance.

Only half the transistor device is modelled, as it is symmetric on both source and drain sides. Also, a single nanowire is modelled with periodic boundary condition to simulate a device with multiple wires. The capacitance is then normalized, corresponding to the measured capacitance normalized by the mesa width.

Dimensions were approximated from SEM images taken during processing. The nanowire has dimensions 25x10 nm and the oxide thickness is 5 nm. The width of the model is 40 nm, corresponding to a nanowire spacing of 15 nm. The gate length is 40 nm (20 nm in model) and the height of the gate is 130 nm. The thickness of the source contact is 30 nm. The geometry of the model is shown in figure 4.7 and 4.8 where the oxide is marked in blue. The nanowire is hidden by the ledge so it is only visible in the gate stack view.

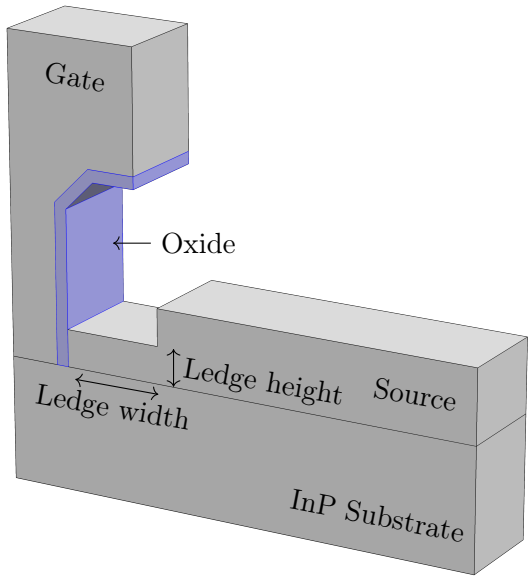


Figure 4.7: Model side view

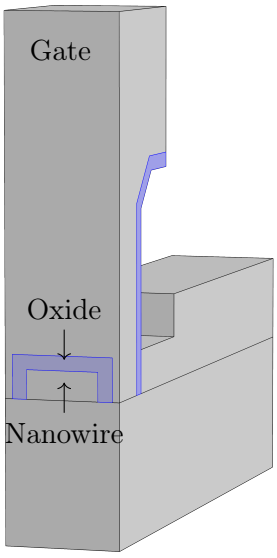


Figure 4.8: Model gate stack view

Four different ledge models were simulated. The device was first modelled without the ledge, varying the distance between the source contact region and

the gate between 0 and 40 nm. The nanowire was modelled as metallic, like in the simpler model in the previous section. Then, a ledge the same height as the nanowire (10 nm) was added. Following was two other variations, with ledge 15 and 20 nm (5 and 10 nm above the nanowire). The fact that the fabricated ledge is shaped around the nanowires is not taken into account. Instead the ledge is modelled at completely flat. In the case with 10 nm ledge height, this corresponds to filling the trenches between the wires with n^+ material. In the fabricated devices, a 8 nm ledge was grown. Thus, the model with a 15 nm straight ledge is estimated to be the closest to the fabricated devices.

The ledge distance was varied from 0 to 40 nm for the different configurations to see the significance of the parasitic capacitance between the gate contact and the n^+ source region. The result is shown in figure 4.9. The normalisation is per μm, corresponding to a normalisation per total gate width.

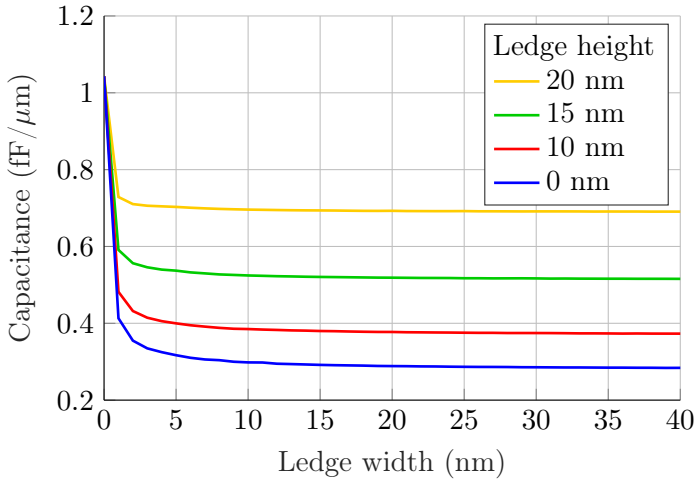


Figure 4.9: Gate to source capacitance for different ledge heights and widths

As can be seen in this figure, a small spacing between the gate and the contact region has a great impact on the capacitance. For the 15 nm ledge, increasing the ledge width from 0 to 5 nm corresponds to a decreased capacitance from 1 fF/μm to 0.54 fF/μm. The capacitance is thus almost halved.

Besides the first few nanometers, the ledge width does not have great impact on device parasitic capacitance. These results indicate that having a wide ledge does not improve performance significantly. Instead, the capacitance between the ledge and the gate contact is the dominating factor. This conclusion is also supported by the large increase in capacitance when the ledge height is increased.

At ledge width 0 nm, the source n^+ region is in direct contact with the gate oxide. This is why the different ledge heights give the same capacitance at this distance, which is dominated by the capacitance directly between the source and the gate with the oxide in between. This would also be equivalent to a ledge height of 30 nm. The value of this is found to be just above 1 fF/μm.

At 0 nm ledge height, the results can be compared to the simple nanowire capacitance model in the previous section where the normalized capacitance was 0.2 fF/μm. This value is close to the $\sim$0.3 fF/μm obtained in the device simulation for a wide ledge. The higher value can be attributed to the presence of the source contact as well as the T-gate geometry.

The electric potential for ledge height 15 nm and ledge width 0, 5 and 10 nm are shown in figure 4.10.

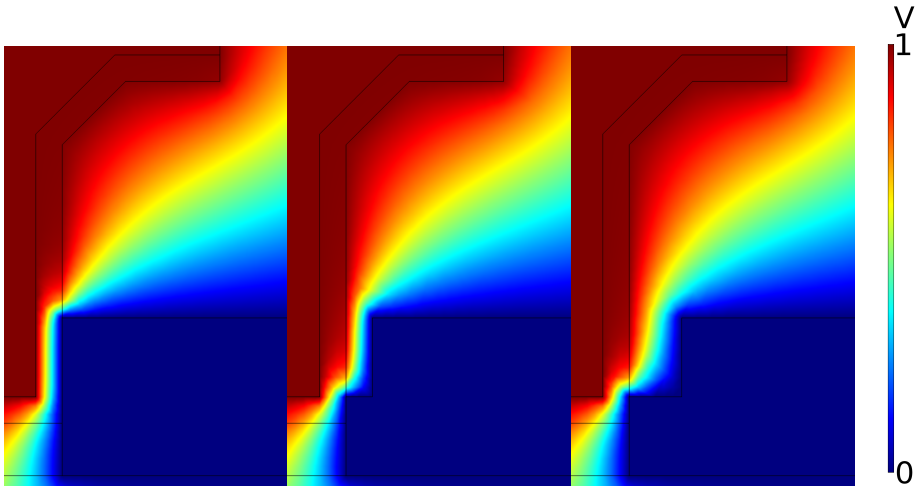


Figure 4.10: Electric potential cross-section plot for different ledge widths a) 0 nm, b) 5 nm and c) 10 nm

Without the spacer, the electric field is strongest inside the oxide, as can be seen in figure 4.10 a). When the ledge width is increased to 5 nm, that electric field is instead moved to the air spacer between the oxide and the contact. Because the permittivity of air is much lower than that for the oxide, this decreases the capacitance. However, the capacitance between the gate and the ledge through the oxide is still large and does not change by increasing the ledge width. This can be seen in figure 4.10 b) and c).

Chapter 5
Measurements & Modelling

To be able to form a perception of the device performance, benchmarking is necessary. This is done by evaluating DC- and RF-performance. This chapter introduces the different benchmarking metrics and how they are obtained.

5.1 DC measurements

5.1.1 DC performance

DC-evaluation is done by measuring the output and transfer characteristics at DC voltage. The *transfer characteristics* is measured by sweeping the gate voltage at a set drain bias. The drain current is then plotted against gate voltage. From the results, threshold voltage and transconductance can be found.

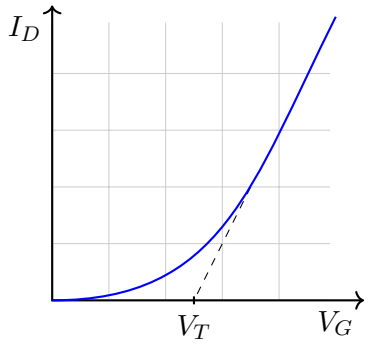


Figure 5.1: Transfer characteristics.

The *output characteristics* is measured by sweeping drain voltage for different set gate biases. This results in several plots with a similar shape. From these plots, output conductance and on-resistance can be extracted.

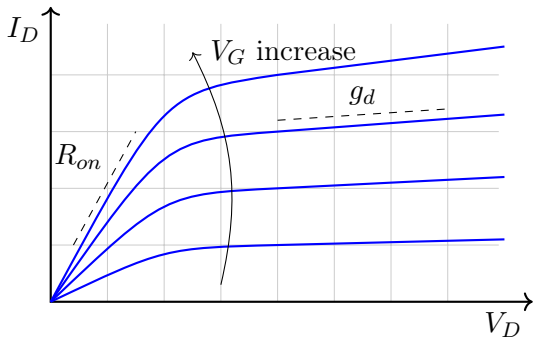


Figure 5.2: Output characteristics.

5.1.2 Transconductance

Transconductance is a very important metric for benchmarking the MOSFET and can be extracted from the transfer characteristics (figure 5.1) as the slope of the drain current. It is defined as the derivative of the drain current when the gate voltage is changed.

$$g_m = \left.\frac{\partial I_D}{\partial V_{GS}}\right|_{V_{DS} \text{ constant}} \tag{5.1}$$

The transconductance is affected by the conduction properties of the channel and the gate control. It can be increased by improving the transport properties of the channel, e.g. by using a material with high carrier mobility μ and long mean free path λ. Transconductance can also be improved by decreasing gate length and thus the channel resistance, and by increasing the oxide capacitance C_{ox} and thus the gate control.

5.1.3 Output conductance

The output conductance is defined as the derivative of the drain current but instead when the drain voltage is changed.

$$g_d = \left.\frac{\partial I_D}{\partial V_{DS}}\right|_{V_{GS} \text{ constant}} \tag{5.2}$$

The output conductance can be extracted from the output characteristics as the slope of the current in the saturated region. Ideally, the output conductance should be zero as the drain voltage should not affect the current in the saturated region. However, the channel length modulation of the drain current has an effect on the resistance of the channel and thus the current. Output conductance is also affected by short-channel effects.

5.1.4 On-resistance

The on-resistance of the MOSFET can be obtained from the slope of the output characteristics in the linear region, seen in figure 5.2. The on-resistance describes

how large the total resistance is between the source and drain contact. To achieve a high current and low current loss, a low R_{on} is desirable.

$$R_{on} = \left(\frac{dI}{dV_{DS}}\right)^{-1}\Bigg|_{V_{DS}\to 0} \tag{5.3}$$

The on-resistance can be reduced by having a short and wide channel with few crystal defects, to minimize scattering.

5.1.5 Threshold voltage

The threshold voltage, V_T is the gate voltage at which the device switches from off-state to on-state. It can be extracted from the transfer characteristics (figure 5.1) by extrapolating the slope of the I_D curve at the maximum transconductance down to zero drain current. It is generally defined as the voltage required to create an inversion layer in the MOSFET channel but the extraction of this parameter is quite arbitrary and can be done in different ways.

5.1.6 Measuring DC

In this diploma work, a Cascade 11000B Probe station was used to measure the transfer and output characteristics of the devices. Three probes were connected to gate, drain and one of the source contacts respectively. This means that one of the source-pads was not biased and one of the gate fingers was unmeasured. For the transfer characteristics, the gate voltage was swept from -0.5 V to 1 V for a constant drain bias of 0.5 V. For the devices with promising transfer characteristics ($g_m > 0$), the output was measured as well. For the output characteristics, the drain voltage was swept from 0 V to 1 V for gate voltages between -0.5 V to 1 V with a step size of 0.1 V. From the DC data, an evaluation of which devices to measure RF performance on can be made.

5.2 RF measurements & modelling

5.2.1 RF performance

For good RF performance, the important metrics are somewhat different from that of DC performance. One big difference is the importance of capacitances. In a DC environment the capacitances can be seen as open circuit element, but when the frequency increases the impedance of the capacitances decreases as $\frac{1}{\omega}$. At gigahertz frequencies, small parasitic capacitances can be limiting the to RF performance. The capacitances along with the transconductance sets the upper limit of frequency where there is gain. Gain and maximum frequency are very important metrics for high-frequency applications.

5.2.2 Capacitance

There are both wanted and unwanted capacitances in a transistor. The MOSFET design relies on a capacitance that opens the channel from the gate over the oxide,

using the field effect. Here, it is beneficial with a large oxide capacitance (C_{ox}) to increase the electrostatic control of the channel.

However, there are also some parasitic coupling between the transistor contacts that limits the performance. These can be divided into *intrinsic* and *parasitic* capacitances.

The intrinsic capacitances are located *inside* the MOS-region of the transistor, between the channel and the source and drain contacts. This is visualized in figure 5.3, where the intrinsic capacitance between gate and source are denoted C_{gs} and the capacitance between gate and drain are denoted C_{gd}. The intrinsic capacitances are determined by the internal properties of the transistor, eg. geometry and material properties. The intrinsic gate-source capacitance is approximated (in saturation) with [18]

$$C_{gs,i} \approx \frac{2}{3} W L_g C_{ox} || C_q \tag{5.4}$$

C_q is the quantum capacitance and stems from the is a limited density of states. Large intrinsic capacitances result in a slow response when the drain and gate voltages are changed. This causes a low switching speed and high power losses [19].

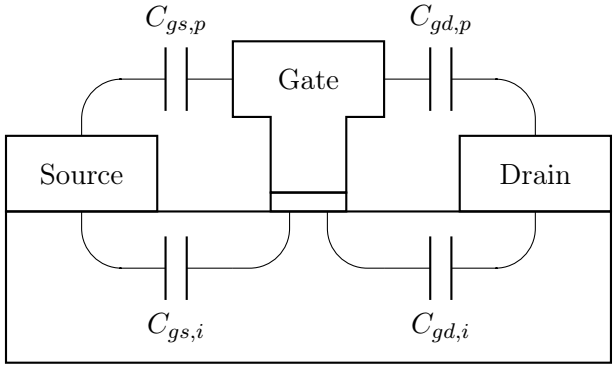


Figure 5.3: Schematic describing the intrinsic and parasitic capacitances in MOSFET structure.

Between source and gate pads there is a parasitic capacitance denoted $C_{gs,p}$ and between gate and drain there is a parasitic capacitance denoted $C_{gd,p}$. Both of these capacitances can be seen in figure 5.3. The parasitic capacitances are a limiting factor of the transistor speed, which is critical for RF applications. This is discussed further in the sections about maximum frequencies later in this chapter. To achieve low parasitic capacitances there are some parameters that can be tweaked. A larger distance between the contacts and a low permittivity material between them results in a low parasitic capacitance [20].

The parasitic and intrinsic capacitances constitute the total gate-source ($C_{gs,t}$) and gate-drain ($C_{gs,t}$) capacitances, as can be seen in the equations listed below. $C_{gg,t}$, the total gate capacitance, is the sum of the gate-source and gate-grain

capacitances.

$$C_{gs,t} = C_{gs,i} + C_{gs,p} \tag{5.5}$$

$$C_{gd,t} = C_{gd,i} + C_{gd,p} \tag{5.6}$$

$$C_{gg,t} = C_{gs,t} + C_{gd,t} \tag{5.7}$$

To achieve good RF properties, the device layout should be optimized for a large oxide capacitance and for minimum parasitic capacitances between the contacts.

5.2.3 Gain

The gain of a device describes its ability to enhance a signal from the input to the output. The gain is defined as the ratio between the amplitude of the output signal and the input signal and is often displayed in dB. There are different gain definitions used for different applications. Current gain (h_{21}), are used to define the cutoff frequency, f_T. At f_T, h_{21} is equal to unity. Maximum available gain (MAG), maximum stable gain (MSG) and unilateral gain (U) are examples of power gain. These are used to define the maximum oscillation frequency, f_{max}. At f_{max}, U and MAG/MSG are both equal to unity. The MAG/MSG also shows where the circuit goes from unconditional to conditional stability.

5.2.4 Cutoff frequency

The cutoff frequency, denoted f_T is defined as the frequency where there is unity current gain, that is when h_{21} is equal to one (0 dB) [5].

$$\frac{1}{2\pi f_T} = \frac{C_{gg,t}}{g_m} + \frac{C_{gg,t}}{g_m}(R_S + R_D)g_d + (R_S + R_D)C_{gd,t} \tag{5.8}$$

For small R_S and R_D, f_t can be approximated to be only dependent on transconductance g_m and the total gate capacitance $C_{gg,t}$.

$$f_T \approx \frac{g_m}{2\pi C_{gg,t}} \tag{5.9}$$

This approximation shows which parameters are most important for maximizing cutoff frequency, g_m and $C_{gg,t}$.

5.2.5 Maximum oscillation frequency

The maximum oscillation frequency, denoted f_{max}, is defined as the frequency where the power gain U is equal to unity (0 dB) [5].

$$f_{max} = \sqrt{\frac{f_T}{8\pi R_G C_{gd,t}\left(1 + \frac{2\pi f_T}{C_{gd,t}}\Psi\right)}} \tag{5.10}$$

where

$$\Psi = (R_D + R_S)\frac{C_{gg,t}^2 g_d^2}{g_m^2} + (R_D + R_S)\frac{C_{gg,t}C_{gd,t}g_d}{g_m} + \frac{C_{gg,t}^2 g_d}{g_m^2}$$

This shows that f_{max} is affected by the gate resistance R_G and the output conductance g_d.

5.2.6 RF measurements

In this diploma work, a Karl Suss PSM6 probe station and a network analyzer was used to extract the frequency dependence of the devices. First calibration of the setup was performed with the LRRM (Line-Reflect-Reflect-Match) method. The calibration was repeated until an acceptable noise-level was achieved. Open and short de-embedding structures, fabricated on the same sample as the transistors, were measured to extract and remove the pad impedance later. Here both gate and drain voltage was set to 0 V. To be able to de-embed the contact resistances, the off-state was measured. To measure off-state, the drain voltage was set to 0 V and the gate voltage to -0.8 V. The S-parameters were measured in the operating frequency span 10 MHz to 67 GHz at a port power of -27 dBm. The drain voltage was swept from 0 V to 1 V, and the gate voltage from -0.6 V to 1 V.

5.2.7 RF modelling

To further analyse the behaviour of a fabricated device, a hybrid-π model can be extracted. The collected S-parameters from the RF-measurements are converted into y-parameters. From these, small signal parameters can be extracted. The expressions used was similar to the ones presented in an article by Sofia Johansson [21].

$$g_m = \mathrm{Re}(y_{11})\big|_{\omega^2=0} \tag{5.11}$$

$$g_d = \mathrm{Re}(y_{22})\big|_{\omega^2=0} \tag{5.12}$$

$$R_g = \frac{\mathrm{Re}(y_{21})}{(\mathrm{Im}(y_{11}))^2} \tag{5.13}$$

$$C_{gd} = \frac{-\mathrm{Im}(y_{12})}{\omega} \tag{5.14}$$

$$C_{dg} = \frac{-\mathrm{Im}(y_{21})}{\omega} - g_m R_g (C_{gs} + C_{gd}) \tag{5.15}$$

$$C_{gs} = \frac{\mathrm{Im}(y_{11}) + \mathrm{Im}(y_{12})}{\omega} \tag{5.16}$$

$$C_{sd} = \frac{\mathrm{Im}(y_{22})}{\omega} - C_{gd} - g_m R_g C_{gd} + \omega^2 C_{gd} C_{dg} (C_{gd} + C_{gs}) R_g^2 \tag{5.17}$$

The modelling was done using a Matlab-script that makes an automatic fitting of the real and imaginary part of the y-parameters. It then extracts the values of the parameters listed above. This script was written by associate professor Erik Lind, and thus not a part of this diploma work.

A schematic of a hybrid-π model can be seen in figure 5.4 below.

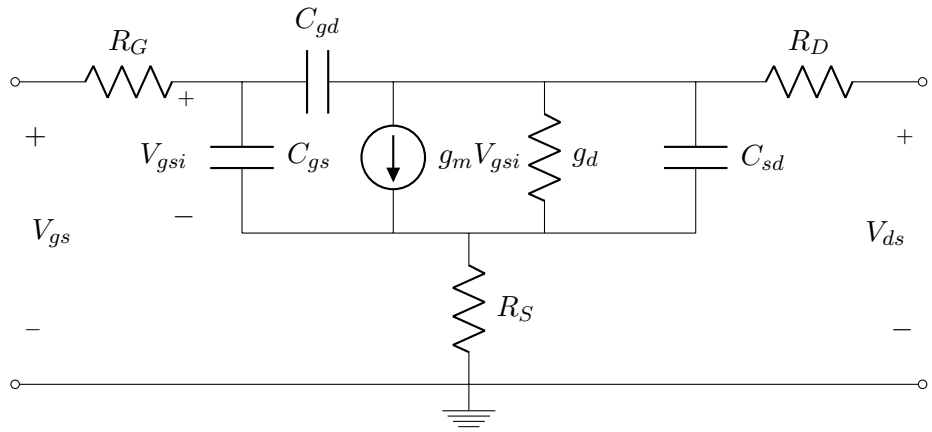


Figure 5.4: A hybrid-π model of a MOSFET

5.3 Limitations

There are some limitations of this small signal modelling. First of all, the hybrid-π model is only an approximation of a real device. To achieve a model that is closer to a real device, more circuit elements would be needed for the small signal model. The automatic fitting of the model to the real y-parameters is written for an A type device with a mesa width of 7 μm. This could be a problem for the B devices and for the devices with mesa 14, resulting in worse fitting and less accurate parameter extraction.

Chapter 6
Results & Analysis

During this diploma work, six samples were fabricated with a total number of 464 devices. However, sample I was not finished and lack pad metal, and thus was not analysed in a DC or RF environment. This means that 380 devices, on four samples, were analysed in a DC environment. Out of the 380 devices, only 124 showed transistor behaviour (transconductance>0 mS). This corresponds to a yield of 33%. The DC measurements showed that sample V had a superior performance compared to the other samples, and it was therefore the only sample measured in RF. The best performing devices on sample V were analysed further in Matlab.

In figure 6.1 and 6.2, SEM images of one finished *A*-device and one finished *B*-device are shown.

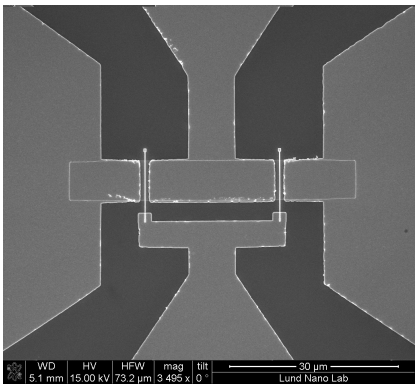


Figure 6.1: A overview of a finished *A* device with mesa 7 μm

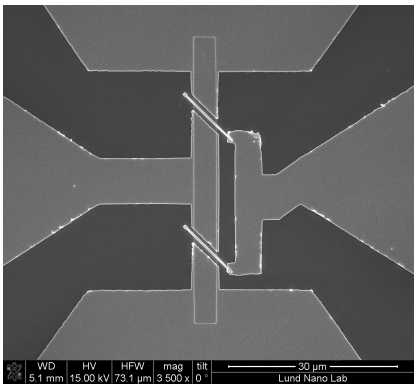


Figure 6.2: A overview of a finished *B* device with mesa 3.5 μm

6.1 Processing

Before any electrical measurements are performed, the samples can be analysed in terms of how well the resulting devices correspond to the intended fabrication. Some of the samples encountered more problems during the processing than others.

This results in large variations of the quality between the samples, and also between devices on the same sample. In this section, all six processed samples are presented with some problems that were encountered during processing.

Sample I & II: Tilted SEM images of the nanowire channel of one A- and one B-device is shown in figure 6.3 and 6.4 below.

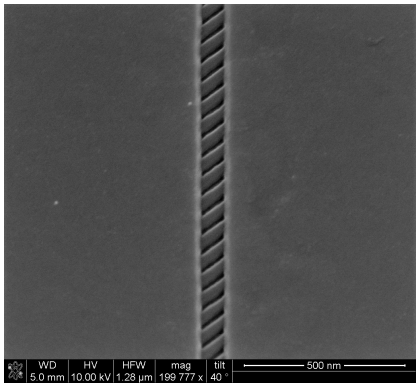


Figure 6.3: A SEM image of the nanowire channel of a *A*-device

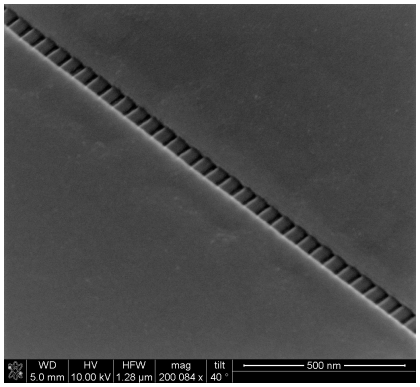


Figure 6.4: A SEM image of the nanowire channel of a *B*-device

The nanowires of the C-type devices were discovered to be very rough in the edges after growth compared to A and B-devices. This may result in poor performance of device, such as higher resistances due to crystal defects and lower carrier mobility, resulting in a lower g_m. Since there is no interest of fabricating devices with a probable poor performance, the C-type devices were not fabricated on the subsequent samples. There was also a problem with aligning the dummy gate since the EBL exposure deviated 30 nm from the designed placement. This was corrected by adding a 30 nm offset in the EBL design file. This resulted in well-aligned A-devices, but the offset of B- and C-devices was still 30 nm. To correct this, these device design files were moved an additional 30 nm. The third exposure and development resulted in wavy HSQ dummy gates, and the resist was removed and the exposure was made again. This resulted in even worse alignment. Using the first corrected files, a new exposure was made, and the alignment was somewhat better. One reason the B- and C-type devices seemed to be more sensitive to poor alignment is that they need to be aligned both in the vertical and horizontal direction of the layout. Since they are rotated 45° to the A-gate, a misalignment in the vertical direction can look like a horizontal misalignment. A vertical misalignment of a A-device would only result in a shift of the gate, making it a bit shorter on one side but not affecting performance.

Other than redesigning the EBL-files, the devices with gate length 50 nm was replaced with 20, 30 and 40 nm devices. The 70 nm gate length devices

were also reduced in number. For longer gate lengths, there is a smaller margin of error, making them harder to align. This could result in a very low yield of working long gate devices. In total, the sample was exposed four more times than intended in the dummy gate EBL step, and thus the resist was removed four additional times. Every time the HSQ is removed using BOE, it also works as a digital etch on the nanowires. This means that part of the InGaAs wires were etched away and the BOE etched down in the InP substrate. This results in nanowires that are InGaAs in the top, and InP at the bottom.

Sample III & IV: The biggest problem encountered in the fabrication of these devices was the InP and InGaAs mesa etch. Etch pits were visible after InP etching and did not seem to vanish when the etching time was increased. When the concentration of HCl was increased the etching was somewhat successful. Some remaining etch pits and surface roughness was present. The gate lift off was difficult and resulted in low yield, as several devices were shorted due to residual metal.

Sample V & VI: The nanowire growth was not completely successful and resulted in poor edges and lower mobility than expected. From Hall measurements on an InGaAs layer grown in the same way as the nanowire layer, the mobility was found to be approximately 700 cm^2/Vs. This had been previously measured to around 3000 cm^2/Vs for similar structures [22].

The 20 and 30 nm wide gates had a yield close to zero due to collapsing of the thin dummy gate structure. Therefore, no 20 and 30 nm gate lengths are included in the results. The InP etch is anisotropic and has a higher etch rate in the (111)B plane than in the (111)A [23], which resulted in a large under etch of the mesa in the (111)B-plane. The under etch of a *B*-device mesa is shown in figure 6.5.

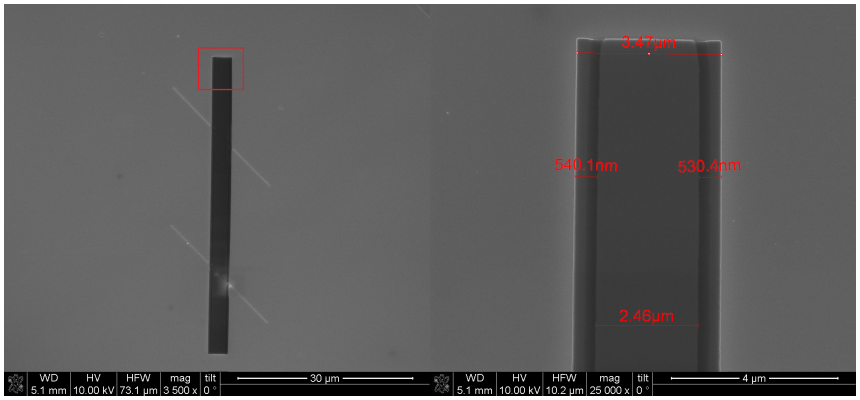


Figure 6.5: SEM image of under etch on a *B*-type device

The under etch mainly affects the mesa of the *B*-devices, since they are placed so the width of the mesa is etched. The under etch was measured to around 530-540 nm on each side, resulting in a 1 μm shorter mesa width.

Therefore the effective mesa widths are recalculated and presented as the actual effective mesa width in the following sections.

The alignment of the gate on sample VI was very poor on the *B*-devices, resulting in short-circuited gates. This is shown in figure 6.6 below. This large misalignment of 1.2 μm is most likely caused by poor contrast in the alignment markers that are scanned before every write field exposure in the EBL patterning. The *B*-type devices is, as was discussed above, much more sensitive to misalignment than the *A*-type, since they need to be aligned both in the vertical and horizontal direction.

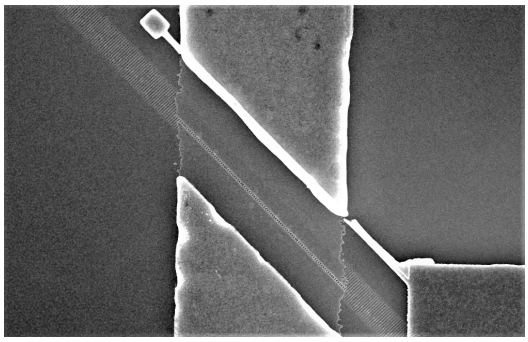


Figure 6.6: SEM image of a misaligned *B*-device

For the *B*-devices on all six samples, the highly doped InGaAs contacts that were supposed to get {110}-facets was observed to be steeper than the expected 45° angle. This could result in poor deposition of the gate metal, by hindering the metal from covering the sides of the sacrificial InP all the way down. This effect is called bridging. Bridging leads to a low electrostatic control since the gate metal is not in contact with the gate oxide, creating a low electric field.

6.2 DC measurement & analysis

All devices that were not visibly shorted were measured in a DC environment as described in chapter 5. From the DC measurements, a rough estimation of the device performance was made. The devices that didn't show transistor behaviour were not analysed further. In this section, the DC performance of sample II, III, IV, V and VI are shortly discussed as well as why some samples were not analysed further.

Sample II: The DC measurements showed that all except one device had large leakage currents from gate to source. Such large leakage current can be explained by lack of gate oxide. This may result from an unsuccessful ALD process or an overly aggressive oxide etch. Either way, no devices on sample II work and this sample was not analysed further.

Sample III & IV: The samples have some working devices but with a very poor performance. Very few devices have a normalized gm over 0.2 mS/μm (nor-

malized with mesa width). Little or no gate modulation was visible. These two samples were not analysed further. The last step of added pad metal on sample IV was unsuccessful for the *B*-devices and thus they could not be measured.

Sample V: This sample showed the most promising DC results of all six processed samples. It had a relatively high yield compared to the other samples, and some of the devices showed high transconductance and gate modulation. The DC results of some devices from sample V are presented further down in this section.

Sample VI: The DC measurements showed low yield of working devices and a low performance of those who worked. No *B*-devices were functional. Since this is a DC reference, it was not as relevant to do RF measurements on.

The only sample that was analysed further is V. However, analysis of the other samples could give a higher understanding of how the processing affects the final results.

6.2.1 DC analysis of *Sample V*

Here, an analysis of the best performing devices on sample V is made. In table 6.1, the devices and how they differ in geometry and exposure methods are shown.

Table 6.1: Summary of gate length, mesa width, nanowire design and gate design for measured devices.

Device	L_g (nm)	Mesa (μm)	NW design	Gate design
Device type A				
A1	40	7	S	2L
A2	40	7	S	2L
A3	40	7	S	2L
A4	50	7	R	6L
A5	50	7	S	2L
A6	60	7	R	6L
A7	60	7	S	2L
A8	70	7	R	6L
A9	70	7	S	2L
A10	40	14	R	6L
A11	70	14	R	6L
A12	70	14	R	6L
Device type B				
B1	50	5.7	R	6L
B2	60	5.7	R	6L
B3	70	5.7	R	6L
B4	70	8.5	R	6L

The mesa listed is the effective mesa, where 1 μm has been subtracted on the B-devices due to the under etch described above. The different mesa widths are later referred to M followed by the mesa size. For example, a mesa width of 7 μm is referred to as M7.

The devices are listed in table 6.2 together with the DC results. The measured transconductance and the transconductance normalized with the total gated nanowire circumference is presented. The data is extracted at drain bias 0.5 V but with different gate bias points depending on where the device has its maximum.

The normalized g_m shows how high the transconductance is, unrelated to the amount of nanowires and their dimensions. This means that the normalized g_m only scales with the gate length and makes it easy to compare the transconductance between devices with both different mesa widths and different gate lengths.

Table 6.2: DC results of some devices from sample V.

Device	L_g (nm)	W_m (μm)	max g_m (mS)	norm g_m (mS/μm)
Device type A				
A1	40	7	4.0	1.08
A2	40	7	1.1	0.30
A3	40	7	4.5	1.21
A4	50	7	4.2	1.13
A5	50	7	3.6	0.97
A6	60	7	3.4	0.91
A7	60	7	2.8	0.75
A8	70	7	3.7	1.00
A9	70	7	2.5	0.67
A10	40	14	6.5	0.87
A11	70	14	7.7	1.04
A12	70	14	7	0.94
Device type B				
B1	50	5.7	1.7	0.40
B2	60	5.7	1.9	0.45
B3	70	5.7	1.2	0.29
B4	70	8.5	3.0	0.48

The devices with the highest measured transconductances are A10, A11 and A12 in table 6.2. This is not surprising since they all have a 14 μm mesa, which gives a higher transconductance due to the larger number of nanowires. A10 should however have a better performance than A11 and A12 since the gate length is shorter, but instead it has a lower g_m. The difference is not very significant, and it may be explained by that there are processing variations that make the results somewhat uncertain.

Among the devices with 7 μm mesa, the devices with 40 nm gate length (A1, A2 and A3) should have the highest g_m. However, there are some variations between

the result for the different devices, even though they are identical by design. It can be seen that the normalized g_m for A2 is much lower than for the other two, even though they were processed on the same sample and same row. Again, this is most likely due to process variations. The device with best normalized transconductance is the third from the top, A3, with a value of 1.21 mS/μm.

Doing the nanowire EBL exposure with single pixel lines (S) instead of the rectangle (R) seems to have achieved a better result. Furthermore, the highest dose resulted in the best performing device. This could mean this method gives a higher quality of the HSQ edges, and thus the nanowire surface.

The I_D - V_{DS} curves for four devices (A1, A5, A7 and A9) are compared in figure 6.7. The devices all have mesa width 7 μm and same nanowire (S) and gate exposure (2L), but different gate lengths. It is clear that the current increases with shorter gate. This is expected, since the drain current is proportional to $\frac{1}{L_g}$. It can also be seen from this result that the shorter gate length devices have a much higher off-current, indicating that there are some short channel effects making it difficult to turn off the transistor.

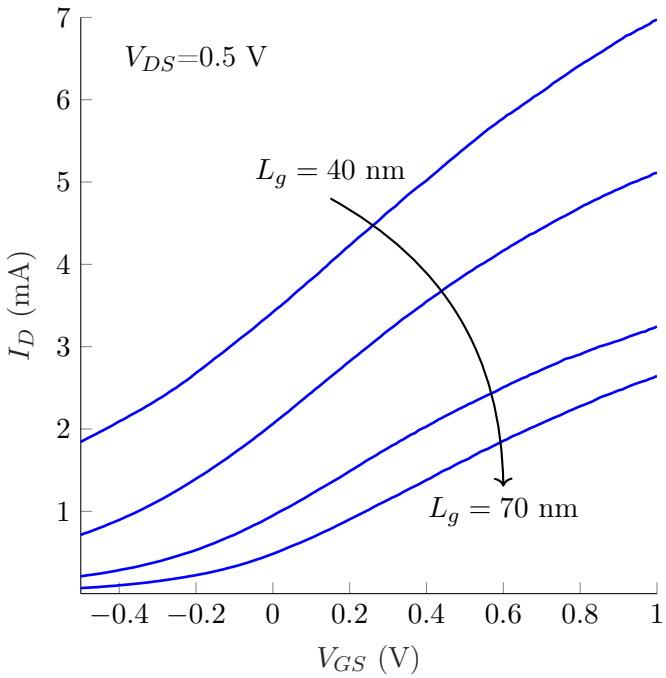


Figure 6.7: I_D-V_{GS} curve for devices with gate lengths of 40, 50, 60 and 70 nm

Overall, the *B*-type devices seem to have a much lower normalized transconductance than the *A*-type. This may be explained by poor alignment of the gate fingers which was a recurring problem during the processing. However, it is hard to say what differences in performance are due to geometry and what are due to process variations. The low yield and large process variations make it difficult to

make valid comparisons between the different devices.

Some further analysis was done on device A3, since it had the best DC performance of all the devices. The output and transfer characteristics are shown in figure 6.8 and 6.9, respectively. It should be noted that DC measurements are only performed on one gate finger, so roughly twice the magnitude of current and transconductance can be expected from the entire device.

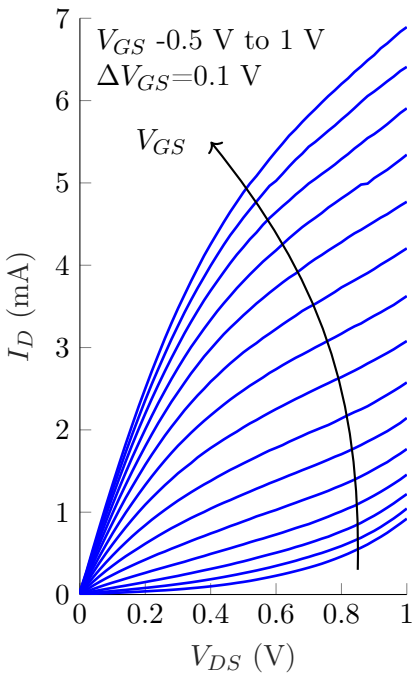


Figure 6.8: Output characteristics.

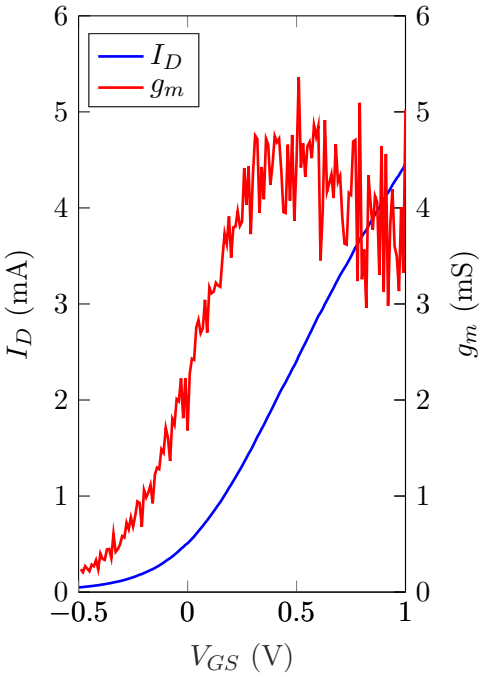


Figure 6.9: Transfer characteristics.

A3 showed a lower off-current than the other devices, and quite high transconductance. The off-current is measured and normalized to 13.0 nA/μm. However, as can be seen from the lack of saturation in the output characteristics, it has a high output conductance.

6.3 RF analysis of sample V

The RF measurements are used to get a thorough evaluation of the device performance. The transconductance and maximum frequencies are used to evaluate the overall performance of the transistors, while the capacitances are used to understand how the geometry of the device can be improved. These are extracted according to the model described in chapter 5. Since the only sample with promising DC-results was sample V, the other samples were chosen not to be measured in a RF environment.

6.3.1 Conductance & capacitance

The extracted transconductance g_m, output conductance g_d, drain-source capacitance C_{gs} and gate-drain capacitance C_{gd} for device A3 are shown in figure 6.10 and 6.11 below. The parameters are shown for every bias point, eg. V_{gs} from -0.6 V to 1 V and V_{ds} from 0 V to 1 V.

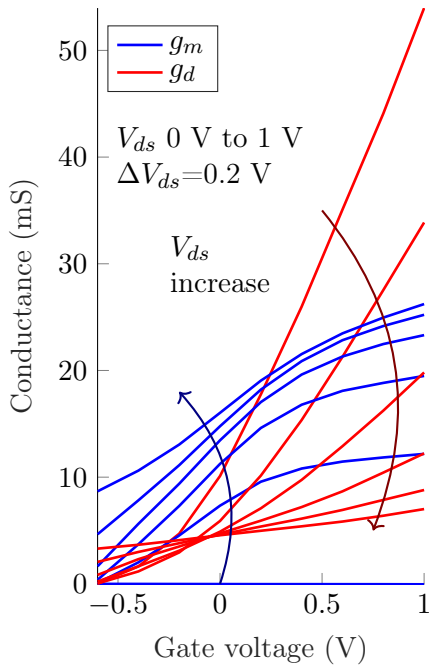


Figure 6.10: Extracted transconductance and output conductance for each bias

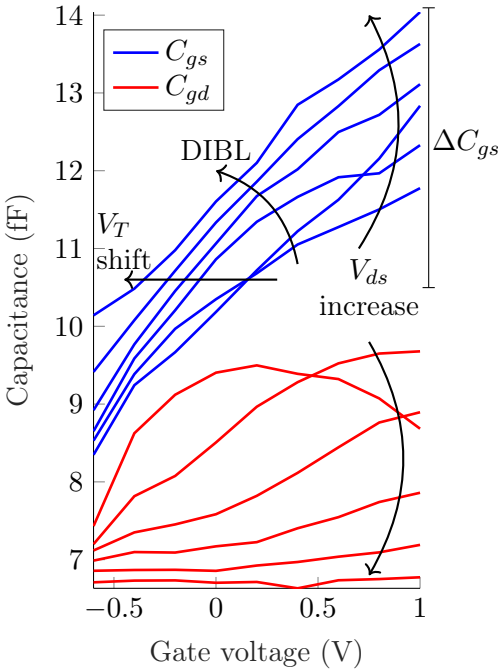


Figure 6.11: Extracted gate-source and gate-drain capacitance for each bias

From the source-drain capacitance shown in figure 6.11 it is clear that this capacitance increases with gate voltage. The drain gate capacitance however, seem to be somewhat constant. The increase of C_{gs} with V_{gs} is explained by studying the intrinsic capacitance, that is introduced in chapter 5 equation 5.4 and showed again below. $C_{gs,i}$ increases with C_{ox} which in turn increases with gate voltage.

$$C_{gs,i} \approx \frac{2}{3} W L_g C_{ox} || C_q$$

At higher drain biases, short channel effects like DIBL is present. This lowers the threshold voltage, creating a shift of the C_{gs} to higher values.

From the RF measurement data for every device measured, the RF parameters from a single bias point was obtained. The chosen bias point was $V_{gs} = 1$ V and $V_{ds} = 1$ V. The transconductances and capacitances for the same devices introduced before are summarized in table 6.3 below. Also included is the normalized

transconductance. Unfortunately, no g_m for the devices with 14 μm mesa width was extracted. This may be due to low accuracy of the Matlab model, which resulted in unreasonable high or low values.

Table 6.3: RF results of same devices from sample V as presented in previous section

Device	L_g (nm)	W_m (μm)	g_m (mS)	norm g_m (mS/μm)	C_{gs} (fF)	C_{gd} (fF)
Device type A						
A1	40	7	25.1	3.37	14.6	5.64
A2	40	7	8.7	1.17	10.0	6.87
A3	40	7	26.2	3.53	14.0	6.77
A4	50	7	23.8	3.20	20.2	9.48
A5	50	7	24.0	3.22	13.6	6.86
A6	60	7	20.8	2.8	19.1	9.43
A7	60	7	16.1	2.16	14.6	6.14
A8	70	7	20.2	2.71	11.4	7.60
A9	70	7	13.7	1.84	11.1	6.17
A10	40	14	-	-	55.8	12.3
A11	70	14	-	-	66.9	13.7
A12	70	14	-	-	51.9	14.3
Device type B						
B1	50	5.7	8.5	1.01	8.05	5.10
B2	60	5.7	8.7	1.03	12.6	4.80
B3	70	5.7	7.1	0.84	7.17	5.63
B4	70	8.5	16.2	0.64	10.3	5.90

To better visualize the results presented in the table above, all devices with corresponding normalized capacitances and normalized transconductance are plotted in figure 6.12.

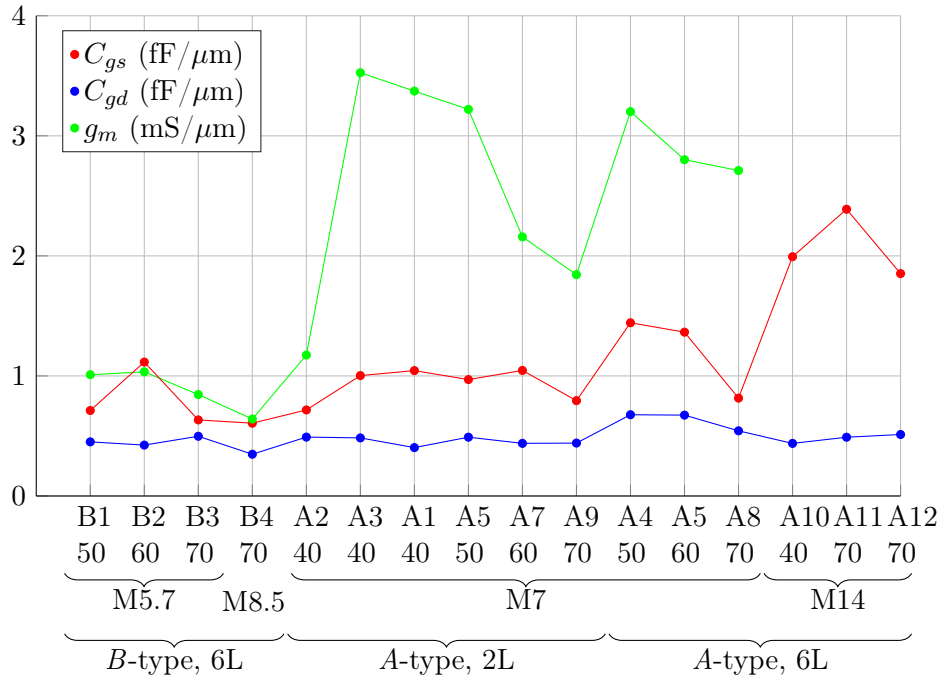


Figure 6.12: Transconductance and capacitances for the devices measured on sample V. Device name and gate and mesa dimensions are showed underneath the plot.

The denotation M5.7, M7, etc. was introduced above, and specifies the mesa size. The first four devices in figure 6.12 are all B type. For these, the gate-source capacitance is lower than for the A devices. This is a promising quality, since a low C_{gs} is a requirement for a high f_T. Alas, the transconductance is also lower. This may be caused by the steep InGaAs walls which could lead to bridging, as was discussed in section 6.1. Comparing the A-devices only, the 6L devices seem to have slightly higher capacitances than the 2L devices. An approximated value of this is C_{gs} increases with $\sim$0.6 fF/μm going from 2L to 6L, and C_{gd} increases with $\sim$0.25 fF/μm. The relatively low capacitances are thanks to the air spacer that was created by the sacrificial InP layer. However, as stated previously, it is hard to say which differences in performance depend on geometry variations and which depend on processing variations.

In figure 6.12, the normalized gate-drain capacitance seem to be rather constant for all devices, with a value around 0.5 fF/μm. This value is highly comparable with the one obtained from the COMSOL simulations, which was 0.54 fF/μm for the ledge height and width 15 nm and 5 nm respectively. This shows that the simulations correspond well to measured values.

The difference between C_{gs} at V_g=1 V and at V_g=-0.6 V is denoted ΔC_{gs}, and is visualized in figure 6.13. If ΔC_{gs} is plotted against L_g an increase is expected with increasing L_g. This is done in figure 6.13. Here, the A-devices with mesa width 7 μm are plotted, as well as the 5.7 μm B-devices. The A-devices are divided

into two groups, depending on if they have the 2L or 6L gate exposure. However, these devices do not seem to follow the trend of decreasing ΔC_{gs} with increasing L_g. This could be explained by process variations and few data points.

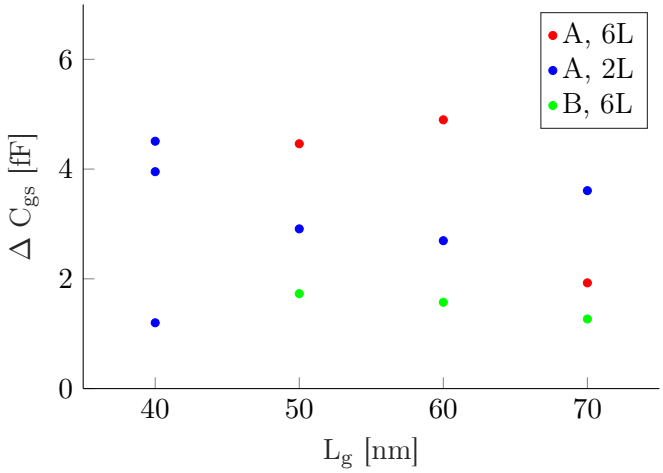


Figure 6.13: ΔC_{gs} against gate length

The RF transconductance in the table 6.3 above is extracted at gate and drain voltage 1 V. To be able to compare the DC and RF values to each other, the RF g_m was extracted at V_{ds}=0.6 V V_{gs}=1 V, which is the closest bias point to the one used in the DC measurements. The transconductances are both normalized with the total nanowire circumference. The RF g_m is also divided by two to obtain the transconductance for one gate finger, since only one finger was measured during DC. The two values are plotted in figure 6.14 below for each device where RF g_m was obtained, eg. for all except for the M14 devices.

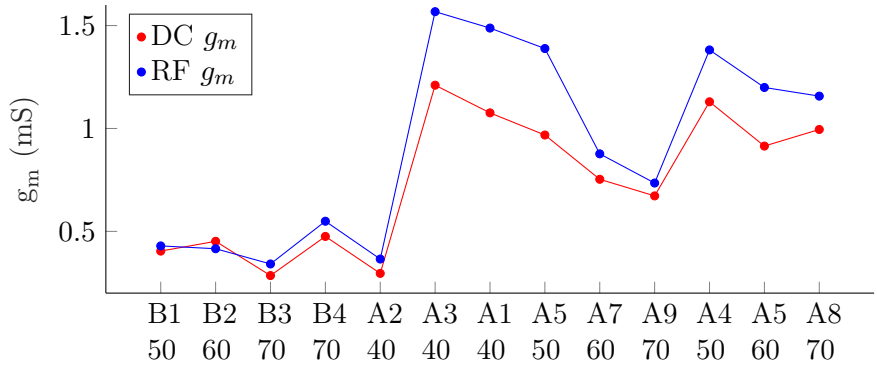


Figure 6.14: Transconductance from both DC and RF measurements

The transconductance obtained from the RF measurements shows a slightly

higher value than for the DC measurements. This may be explained by traps in the oxide that lowers the current in DC measurements, but are unnoticed in RF measurements. The traps can be caused by impurities in the ALD process, and can be reduced by cleaning the ALD chamber thoroughly.

6.3.2 Gain

The gain plot for device A3 on sample V is shown in figure 6.15. The dots show the measured data from the RF measurements, while the dashed lines show extrapolated values.

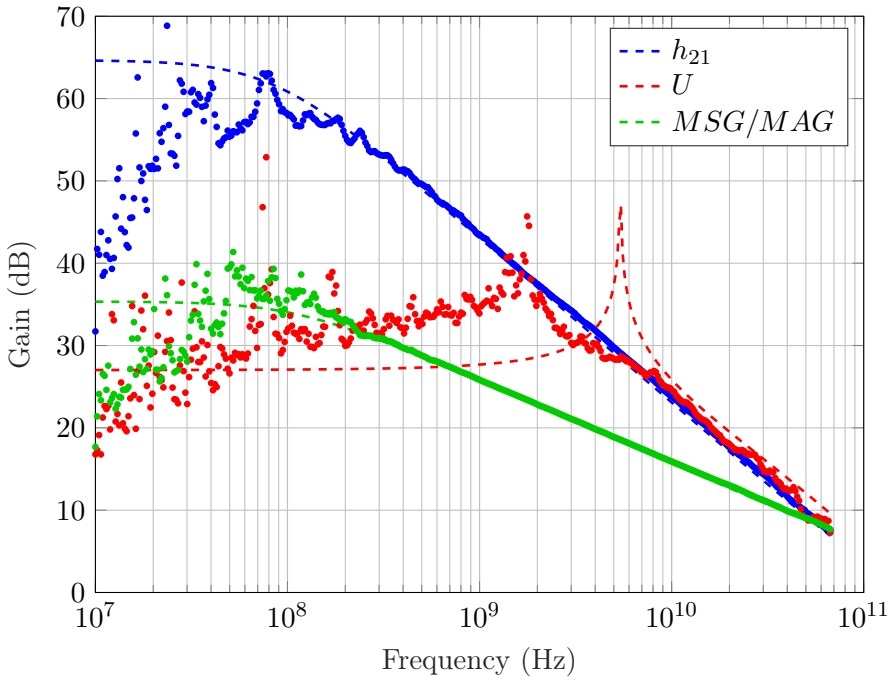


Figure 6.15: Current and power gain.

It is clear that the model overall follows the data well, except for a small deviation for the gains in high frequencies where h_{21} seems to be a little bit too low compared to the measured values, and U seems to be too high. This would result in a slightly lower cutoff frequency and a slightly higher maximum oscillation frequency. This is good to have in mind when reading the following section.

6.3.3 Cutoff frequency & maximum oscillation frequency

Maybe the most interesting merit for RF devices is the frequency operation. As described in chapter 5, there are two different frequency merits that are of interest: f_T and f_{max}. Equation 5.9 and 5.10, that is repeated below, are introduced in chapter 5 as well.

$$f_T \approx \frac{g_m}{2\pi C_{gg,t}}$$

and

$$f_{max} = \sqrt{\frac{f_T}{8\pi R_G C_{gd,t}\left(1 + \frac{2\pi f_T}{C_{gd,t}}\Psi\right)}}$$

From the Matlab script, f_T and f_{max} were extracted for V_{gs} going from -0.6 V to 1 V at V_{ds}=1. The obtained values for device A3 is shown in figure 6.16. The cutoff frequency shows a clear increasing trend with both gate and drain bias. The maximum oscillation frequency is also increasing at higher bias, but is clearly more fluctuating at higher voltages.

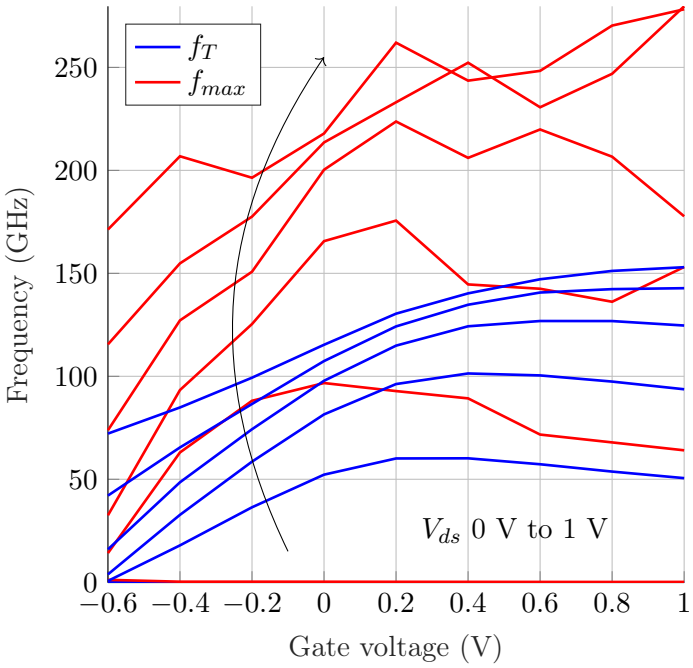


Figure 6.16: Extracted cutoff and maximum oscillation frequency for each bias

The extrapolated f_T and f_{max} for all previously introduced devices on sample V can be seen in table 6.4. All values are extracted at $V_{ds} = 1$ V and $V_{gs} = 1$ V. Some devices lack a value for f_{max} because of the fluctuation being too high which makes it difficult to obtain a reasonable value. Also, because of these fluctuations there is some uncertainty in the f_{max} values. Another metric that is used to compare the gain properties of a transistor is how large the Maximum stable gain is at a lower frequency. In the table below, MSG is shown for each device at a oscillation frequency of 20 GHz.

Table 6.4: Frequency results of the same devices from sample V as presented in previous section. Gain is given as MSG at 20 GHz.

Device	L_G (nm)	W_m (μm)	f_T (GHz)	f_{max} (GHz)	Gain (dB)
Device type A					
A1	40	7	138.4	223	13.4
A2	40	7	72.0	145.8	9.7
A3	40	7	153.0	278.2	13.5
A4	50	7	91.7	213	11.2
A5	50	7	137.0	166.9	12.9
A6	60	7	87.8	155.7	11.0
A7	60	7	115.7	177.4	12.3
A8	70	7	130.0	-	11.1
A9	70	7	106.0	162.7	11.9
A10	40	14	79.7	-	10.7
A11	70	14	99.6	-	11.8
A12	70	14	99.0	-	11.6
Device type B					
B1	50	5.7	95.6	121.3	10.8
B2	60	5.7	71.8	208.9	10.9
B3	70	5.7	75.4	123.2	9.7
B4	70	8.5	135	169.0	12.1

The highest extrapolated cutoff frequency is 153 GHz with an associated f_{max} of 280 GHz. These values belong to A3, that also showed the highest normalized transconductance with relatively low capacitances, which is beneficial for a high f_T. As was discussed in section 6.1, the electron mobility of the InGaAs nanowires were overall much lower than previously grown by the Nanoelectronics group. Some devices achieved frequencies close to the best devices fabricated earlier in the group.

The last listed B-type device, B4, has an impressive f_T of 135 GHz. This is the fourth highest cutoff frequency obtained from this sample. Considering the low transconductance of the device (0.64 mS/μm), this is a surprisingly high result. The devices that has a f_T higher than device B4 all has 2L gates, which would result in lower parasitic capacitances. The B-devices which all has 6L gates, and the A-devices with 6L gates, are highly comparable in f_T performance. Still, the A-devices with 2L gates are superior compared to the other devices. This suggests that a even higher performance can be achieved if the 2L gate can be combined with the B-type device.

To find the optimum bias point for maximum f_T and f_{max}, a contour plot can be made. In figure 6.17 and 6.18, f_T and f_{max} for device A3 are plotted as a function of both gate and drain voltage.

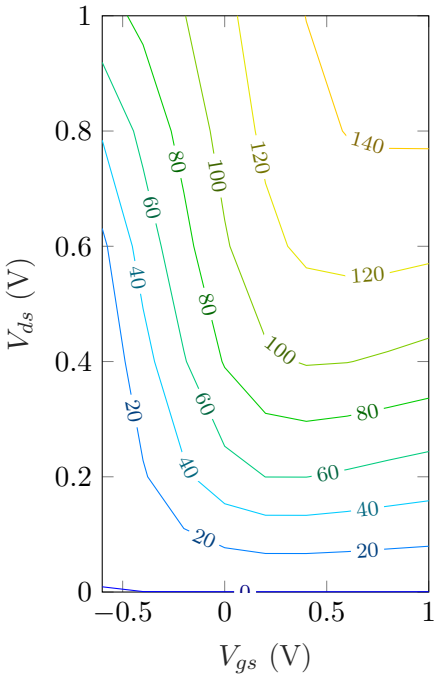


Figure 6.17: Contour plot of f_T

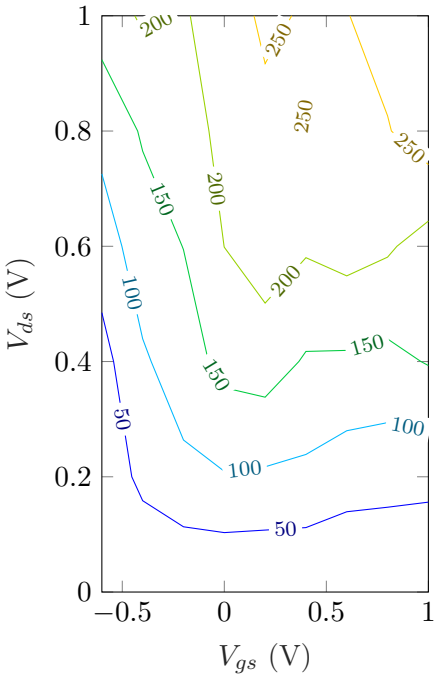


Figure 6.18: Contour plot of f_{max}

The cutoff frequency seem to increase with both gate and drain voltage, and a global maximum value has not been found. The maximum oscillation frequency however has several *islands* that gives a maximum value for different bias points. It is unclear what happens at higher voltages. To see the actual maximum frequencies, the voltage ranges need to be increased.

Chapter 7
Conclusions

The main goal of this diploma work was to discover how different geometries affect the performance of lateral InGaAs nanowire MOSFETs. The low yield and large process variations make it hard to make the comparisons that was intended in the beginning of this work. To be able to draw more accurate conclusions, the processing needs to be refined.

With that said, some conclusions can be drawn from the results presented and analysed in chapter 6.

First of all, both A- and B-type transistors were successfully fabricated. The yield of the A-type devices were much higher than that of B. It is not completely clear why, but one possibility is alignment difficulties. Because the gate in B-devices needs to be aligned with high precision in both the vertical and horizontal directions, bad alignment could be more noticeable for the B-devices. No working devices of type C was fabricated, and thus no conclusions about this geometry can be made.

The characterization of the working devices showed that the A-type was superior to B when it comes to transconductance, cutoff frequency and maximum oscillation frequency. However, B showed promising results on the parasitic capacitances which also is an important metric for the maximum frequencies. If the transconductance can be increased for these devices they may be better than the A-type. However, since so few devices were compared, no definitive conclusion can be drawn as to the performance difference between the two geometries.

The COMSOL simulations on gate drain capacitance had a value close to the one obtained from the RF modelling. This is an indication that the COMSOL model was a good approximation of the fabricated devices.

Some clear trends were observed for the different gate lengths. A higher current and transconductance was observed for shorter gates, as expected. The 20 and 30 nm gate lengths proved to be too short for this processing method, as they seemed to have collapsed. Either, these gate lengths needs to be excluded on future samples, or the method needs to be altered to support the dummy-gate and preventing it from collapsing. It was also noted that when making the nanowire EBL patterning, the devices with nanowires defined using six single pixel lines and a high dose achieved higher performance than the devices using a rectangle. Further experiments on this should be performed, but this suggests that this exposure method with a high line dose should be used in future nanowire processing. For the gate exposure, the large gate structures that was exposed with six single pixel

lines gave rise to a higher parasitic capacitance than the two single pixel lines, as expected. To achieve good RF performance the latter should be used, which increases the need for better alignment.

COMSOL simulations show the significant decrease in capacitance when a small air spacer between the contact region and the gate is introduced. For a 5 nm wide, 15 nm high ledge, the capacitance is almost halved compared to without a spacer. Increasing the ledge width even more, and thus the spacer, does not have great impact as the capacitance between the ledge itself and the gate becomes dominant. This capacitance is also shown to be greatly increased by increasing the ledge height. Even the 15 nm ledge introduces a significant parasitic capacitance of 0.5 fF/μm, compared to 0.3 fF/μm for only a wire. The purpose of the ledge is to reduce contact resistance, but it is a trade-off with the parasitic capacitance introduced. An even thinner ledge could be implemented, or other methods of decreasing the contact resistance could be investigated.

Even with the process variations, a few devices show promising characteristics, e.g. with f_T around 150 GHz and f_{max} above 250 GHz. These numbers are especially promising due to the fact that the mobility in the nanowires is believed to be quite low in the measured samples compared to what can be achieved, and the transconductance is directly proportional to the mobility. However, these results were only obtained for the geometry previously developed, so further testing with the new geometry is needed to determine the reason for the low yield and the lower transconductance.

References

[1] Nationalencyklopedin. (2018) Dator. [Online]. Available: http://www.ne.se/uppslagsverk/encyklopedi/lång/dator

[2] ——. (2018) ENIAC. [Online]. Available: http://www.ne.se/uppslagsverk/encyklopedi/lång/eniac

[3] N. Routley, "How the computing power in a smartphone compares to supercomputers past and present," Nov 2017. [Online]. Available: http://www.businessinsider.com/infographic-how-computing-power-has-changed-over-time-2017-11?r=US&IR=T&IR=T

[4] R. G. Arns, "The other transistor: early history of the metal-oxide semiconductor field-effect transistor," *Engineering Science and Education Journal*, vol. 7, no. 5, pp. 233–240, Oct 1998. [Online]. Available: https://ieeexplore.ieee.org/document/730824/

[5] C. Zota, "III-V MOSFETs for high-frequency and digital applications," Ph.D. dissertation, Department of Electrical and Information Technology, Lund University, 2017.

[6] R. W. KEYES, "Fundamental limits of silicon technology," *PROCEEDINGS OF THE IEEE*, vol. 89, no. 3, March 2001. [Online]. Available: https://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=915372

[7] J. E. Lilienfeld, "Method and apparatus for controlling electric currents," 1930.

[8] J. Bardeen and W. H. Brattain, "The transistor, a semiconductor triode," *Proceedings of the IEEE*, vol. 86, no. 1, pp. 29–30, Jan 1998. [Online]. Available: https://ieeexplore.ieee.org/document/658753/

[9] W. Shockley, "The theory of p-n junctions in semiconductors and p-n junction transistors," *The Bell System Technical Journal*, vol. 28, no. 3, pp. 435–489, July 1949. [Online]. Available: https://ieeexplore.ieee.org/document/6773080/

[10] Nobelprize.org, "The nobel prize in physics 1956," 2014. [Online]. Available: http://www.nobelprize.org/nobel_prizes/physics/laureates/1956/

[11] S. Sze and M. Lee, *Semiconductor Devices: Physics and Technology*, 3rd ed. John Wiley & Sons, 2013.

[12] G. E. Moore, "Cramming more components onto integrated circuits," *Proceedings of the IEEE*, vol. 86, no. 1, pp. 82–85, Jan 1998. [Online]. Available: https://ieeexplore.ieee.org/document/658762/

[13] J. A. del Alamo, "Nanometre-scale electronics with III-V compound semiconductors." *Nature*, vol. 479, no. 7373, pp. 317 – 323, 2011. [Online]. Available: http://ludwig.lub.lu.se/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=a9h&AN=67327843&site=eds-live&scope=site

[14] U. W. Pohl, *Epitaxy of Semiconductors: Introduction to Physical Principles*. Springer, 2013.

[15] P. D. Ye, G. D. Wilk, J. Kwo, B. Yang, H. J. L. Gossmann, M. Frei, S. N. G. Chu, J. P. Mannaerts, M. Sergent, M. Hong, K. K. Ng, and J. Bude, "GaAs MOSFET with oxide gate dielectric grown by atomic layer deposition," *IEEE Electron Device Letters*, vol. 24, no. 4, pp. 209–211, April 2003. [Online]. Available: https://ieeexplore.ieee.org/document/1206840/

[16] J. Robertson, "High dielectric constant oxides," *THE EUROPEAN PHYSICAL JOURNAL APPLIED PHYSICS*, vol. 28, p. 265–291, 2004. [Online]. Available: https://web.stanford.edu/class/ee311/NOTES/Robertson%20JAP04.pdf

[17] COMSOL Multiphysics. (2018) Simulate static and low-frequency electromagnetics with the AC/DC module. [Online]. Available: https://www.comsol.com/acdc-module

[18] Gray, Hurst, Lewis, and Meyer, *Analysis and Design of Analog Integrated Circuit*, 5th ed. John Wiley & Sons, 2010.

[19] G. Consentino and G. Ardita, "A simplified and approximate power MOSFET intrinsic capacitance simulation: Theoretical studies, measures and comparisons," in *2008 IEEE International Symposium on Industrial Electronics*, June 2008, pp. 38–43.

[20] M. Wang, D.-M. Ke, C.-X. Xu, and B.-T. Wang, "A 2-D semi-analytical model of parasitic capacitances for MOSFETs with high k gate dielectric in short channel," *Solid-State Electronics*, vol. 92, pp. 35 – 39, 2014. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0038110113003328

[21] S. Johansson, M. Egard, S. G. Ghalamestani, B. M. Borg, M. Berg, L. E. Wernersson, and E. Lind, "RF characterization of vertical InAs nanowire wrap-gate transistors integrated on Si substrates," *IEEE Transactions on Microwave Theory and Techniques*, vol. 59, no. 10, pp. 2733–2738, Oct 2011. [Online]. Available: https://ieeexplore.ieee.org/document/6015503/

[22] C. B. Zota, D. Lindgren, L.-E. Wernersson, and E. Lind, "Quantized conduction and high mobility in selectively grown $In_xGa_{1-x}As$ nanowires," *ACS Nano*, vol. 9, no. 10, pp. 9892–9897, 2015, pMID: 26387961. [Online]. Available: https://doi.org/10.1021/acsnano.5b03318

[23] B. Tuck and A. J. Baker, "Chemical etching of {1 1 1} and {1 0 0} surfaces of InP," *Journal of Materials Science*, vol. 8, no. 11, pp. 1559–1566, Nov 1973. [Online]. Available: https://doi.org/10.1007/BF00754890

Appendix A
Fabrication methods

This appendix describes the processing methods mentioned in chapter 3. More thorough descriptions are found in Semiconductor Devices: Physics and Technology by Sze, S.M. and Lee, M.K. [11].

A.1 Spin coating

To get an even distributed resist on a sample, a spin-coater (or spinner) can be used. A chuck holds the sample in place by using vacuum. By rotating the sample at a high speed, the deposited resist is forced to spread out over the sample. The excessive resist is forced off the sample, leaving only a thin film of resist. Ramping, spinning speed and time is set to match the properties of the resist and what thickness is required.

A.2 Electron Beam Lithography

Electron beam lithography (EBL) is a lithography method with high resolution and flexibility. EBL uses a focused, high energy electron beam to expose an electron sensitive resist in a desired pattern. The probe is moved, pixel by pixel, over a sample which produces a high resolution pattern in the electron sensitive resist. However, there is a trade-off between resolution and throughput. A low current gives a high resolution pattern but a low throughput, and a high current gives a high throughput but a lower resolution. Depending on the resist type, the exposed areas becomes less or more soluble. For a negative photoresist, exposure with an electron beam leads to cross-linking of the resist, making the exposed areas less soluble than the unexposed. The unexposed areas can be removed with a process called development leaving a pattern of cross-linked photoresist. An example of a negative photoresist is HSQ. For a positive photoresist the exposed areas becomes more soluble and creates an inverted pattern compared to a negative photoresist. An example of a positive photoresist is PMMA.

A.3 Plasma ashing

Plasma ashing is often used as a compliment to development of resist to make sure there is no resist residues left. The sample is placed in a chamber where a low pressure of oxygen is ignited with high power radio waves. The ionized oxygen is highly reactive and will react to the remaining resist.

A.4 Atomic Layer Deposition

Atomic layer deposition (ALD) is a method to deposit thin films, like gate oxides, in a highly controlled way. The sample is placed in a vacuum chamber with an elevated temperature. The chamber is filled with reactive gases called precursors, one at a time. The gas reacts with the surface of the sample, leaving a single atom thick layer. The first gas is completely removed before the second gas enters the chamber. This creates a layer with a thickness decided by the number of cycles. For example, a five cycle ALD process will create a ten atom thick layer if two precursors are used.

A.5 Evaporation

Evaporation is a method to deposit a metal film with highly controlled thickness. The sample is placed together with the metals in a high vacuum chamber. The metals are placed in *boats* and the sample is placed up side down facing the metals. When the vacuum level is sufficient a high current is generated through the boats, increasing the temperature. Together with the high vacuum this causes the metal to evaporate. The thickness of the metal layer is measured with a piezoelectric crystal. To achieve an even metal layer, the sample holder is rotating during the evaporation.