LUND UNIVERSITY

Nanowire Transistors and RF Circuits for Low-Power Applications

Persson, Karl-Magnus

2014

Nanowire Transistors and RF Circuits for Low-Power Applications

Doctoral Thesis
Karl-Magnus Z. Persson

Thesis Supervisor:
Professor Lars-Erik Wernersson

Faculty Opponent:
Professor Sylvain Bollaert

LUND
UNIVERSITY

Department of Electrical and Information Technology
Faculty of Engineering, LTH, Lund University
SE-221 00 Lund, Sweden

Academic thesis for the degree of Doctor of Technology at the Faculty of Engineering, Lund University, to be defended on Friday the 14th of March, 2014. The defense will take place at 10:00 in room E:1406, found in the E-building, LTH, at Ole Römersväg 3, Lund, Sweden.

Nanowire Transistors and RF Circuits for Low-Power Applications

Doctoral Thesis

Karl-Magnus Z. Persson

Thesis Supervisor:
Professor Lars-Erik Wernersson

Faculty Opponent:
Professor Sylvain Bollaert

LUND
UNIVERSITY

Department of Electrical and Information Technology
Faculty of Engineering, LTH, Lund University
SE-221 00 Lund, Sweden

Lund, Sweden, 2014

Department of Electrical and Information Technology
Faculty of Engineering, LTH, Lund University
P. O. Box 118
SE-221 00
Lund
Sweden

ISSN 1654-790X
ISBN 978-91-7473-809-4 (printed version)
ISBN 978-91-7473-810-0 (pdf verison)
Printed in Lund, Sweden, by *Tryckeriet i E-huset*
February 2014

It's not what you look at that matters, it's what you see.
- Henry David Thoreau

Populärvetenskaplig sammanfattning

Surfplattor och mobiltelefoner med prestanda i klass med den hos stationära datorer för bara några år sedan är resultatet av den kontinuerliga hårdsatsning som i över 50 års tid stadigt lyckats göra den minsta elektronikkomponenten, transistorn, mindre, snabbare och mer strömsnål. Under senare år har det dock blivit svårare att förbättra prestandan enbart genom att minska transistorns storlek. För att längre framöver kunna få fortsatta prestandaökningar, som i sin tur kan generera nya och spännande sätt att interagera med elektronik, kan det behövas helt nya tillverkningsmetoder. Framtida transistorer konstruerade med nanotrådar är ett alternativ som kan ge möjlighet att fortsätta öka prestandan samtidigt som energiförbrukningen kraftigt kan minskas. Vi har demonstrerat nanotrådskretsar som opererar vid gigahertz-frekvenser, vilket är unikt då det visar på en högre nivå av funktion än vad som tidigare har publicerats kring nanotrådar. För att förstå hur nanotrådsteknik kan ge framtida fördelar är det nödvändigt att ge en övergripande bakgrund.

En transistor är en elektronisk ledare där flödet mellan två kontakter kan styras av och på från en tredje kontakt kallad styret. Transistorer används till att förstärka t.ex. radiosignaler, som är en analog applikation, alternativt utföra logiska beräkningar med 1:or och 0:or, som är en digital applikation. Mobiltelefonen är ett exempel på en elektronisk apparat som behöver transistorer i kretsar för både analoga och digitala applikationer.

Tidigt i utvecklingen av transistorn kom kisel att vara det halvledarmaterial som framstod som bäst lämpat. Kisel finns i överflöd, vanligen i form av sand, och det har varit relativt enkelt att omvandla denna halvledare till bra elektronledare. Det finns dock andra material som har egenskaper som i vissa avseenden är mer åtråvärda när det kommer till att bygga bra transistorer. Indium Arsenid (InAs), som är en förening av två grundämnen från grupp 3 och 5 i periodiska systemet, har förmågan att kunna transportera elektroner betydligt snabbare än vad som är möjligt i kisel. Detta ger fördelar både när det gäller hur snabbt en transistor kan slå av och på, samt hur mycket energi som förbrukas vid transporten av elektroner. Forskning tyder på att energiförbrukningen hos en transistor skulle kunna minskas med en faktor 10 i jämförelse med kiselteknik, samtidigt som arbetsfrekvensen kan ökas.

Arkitekturen hos en transistor har även den varit relativt konstant sedan den första integrerade kretsen utvecklades. Två kontakter har förbundits med en två-dimensionell (2D) yta, kanalen, som kan ändra motstånd genom att reglera en pålagd spänning på styret. En transistor med en 2D-kanal kallas för planar. Den aggressiva skalningen, för att öka prestandan, har dock lett till att kanalen har blivit väldigt kort i riktningen mellan de strömledande kontakterna, vilket i sin tur gör att fältverkan, inducerad av spänningen på styret, börjar bli otillräcklig. En lösning för att öka kontrollförmågan via styret är att ändra form på kanalen, och att istället för

en planar design konstruera den som en pinne. Styret kan då läggas runtom pinnen och påverka från alla riktningar samtidigt. Detta är idén med en nanotrådstransistor. För att förstå hur strömmen styrs av och på i en nanotrådstransistor kan det var illustrativt att göra en liknelse med ett vattenflöde i en vattenslang. Styret fungerar då som en hand som kan greppa vattenslangen och krama åt alternativt släppa upp. Vid små dimensioner har ett omslutande styre visat sig vara långt mer effektivt än ett styre som endast inducerar ett elektriskt fält från en riktning.

Vi har tillverkat nanotrådstransistorer i InAs som uppvisat mycket bra prestanda. Genom att förfina tillförlitligheten i vår process har vi även kunnat tillverka kretsar bestående av tre stycken sammankopplade nanotrådstransistorer i en mixer design. En mixer används för att konvertera en signal till en högre eller en lägre frekvens och används t.ex. i mobiltelefoner för att skicka och ta emot data. Resultatet tillhör ett av de första rapporterade för nanotrådar med en kretsprestanda som ligger i närheten av gigahertz-området och som därmed närmar sig de frekvenser som är intressanta för att skicka data.

Planära kiseltransistorer kommer troligtvis att finnas kvar en lång tid framöver på grund av den låga tillverkningskostnaden. Med framtida krav på högre dataöverföring kan det dock finnas utrymme för specialiserade utökningar. Intresset från industrin inom området för alternativa tekniker är stort och många av de betydande företagen håller på med utveckling av transistorer som nyttjar nya kanaltyper och andra material än kisel.

Abstract

The background of this thesis is related to the steadily increasing demand of higher bandwidth and lower power consumption for transmitting data. The work aims at demonstrating how new types of structures, at the nanoscale, combined with what is referred to as exotic materials, can help benefit in electronics by lowering the consumed power, possibly by an order of magnitude, compared to the industry standard, silicon (Si), used today.

Nanowires are semiconductor rods, with two dimensions at the nanoscale, which can be either grown with a bottom-up technique, or etched out with a top-down approach. The research interest concerning nanowires has gradually increasing for over two decades. Today, few have doubts that nanowires represent an attractive alternative, as scaling of planar structures has reached fundamental limits. With the enhanced electrostatics of a surrounding gate, nanowires offer the possibility of continued miniaturization, giving semiconductors a prolonged window of performance improvements.

As a material choice, compound semiconductors with elements from group III and V (III-Vs), such as indium arsenide (InAs), have the possibility to dramatically decrease power consumption. The reason is the inherent electron transport properties of III-Vs, where an electron can travel, in the order of, 10 times faster than in Si. In the projected future, inclusion of III-Vs, as an extension to the Si-CMOS platform, seems almost inevitable, with many of the largest electronics manufacturing companies showing great interest.

To investigate the technology potential, we have fabricated InAs nanowire metal-oxide-semiconductor field effect transistors (NW-FETs). The performance has been evaluated measuring both RF and DC characteristics. The best devices show a transconductance of 1.36 mS/µm (a device with a single nanowire, normalized to the nanowire circumference) and a maximum unilateral power gain at 57 GHz (for a device with several parallel nanowires), both values at a drive voltage of 0.5 V. The performance metrics are found to be limited by the capacitive load of the contact pads as well as the resistance in the non-gated segments of the nanowires. Using computer models, we have also been able to extract intrinsic transport properties, quantifying the velocity of charge carrier injection, which is the limiting property of semi-ballistic and ballistic devices. The value for our 45-nm-in-diameter nanowires, with 200 nm channel length, is determined to $1.7 \cdot 10^7$ cm/s, comparable to other state-of-the-art devices at the same channel length.

To demonstrate a higher level of functionality, we have connected several NW-FETs in a circuit. The fabricated circuit is a single balanced differential direct conversion mixer and is composed of three stages; transconductance, mixing, and transimpedance. The basic idea of the mixer circuit is that an information signal can either be extracted from or inserted into a carrier wave at a higher frequency than the information wave itself. It is the relative size of the first and the third stage that accounts for the circuit conversion gain. Measured circuits show a voltage

conversion gain of 6 dB and a 3-dB bandwidth of 2 GHz. A conversion mixer is a vital component when building a transceiver, like those found in a cellphone and any other type of radio signal transmitting device.

For all types of signals, noise imposes a fundamental limitation on the minimal, distinguishable amplitude. As transistors are scaled down, fewer carriers are involved in charge transport, and the impact of frequency dependent low-frequency noise gets relatively larger. Aiming towards low power applications, it is thus of importance to minimize the amount of transistor generated noise. Included in the thesis are studies of the level and origin of low-frequency 1/*f*-noise generated in NW-FETs. The measured noise spectral density is comparable to other non-planar devices, including those fabricated in Si. The data suggest that the level of generated noise can be substantially lowered by improving the high-*k* dielectric film quality and the channel interface. One significant discovery is that the part of the noise originating from the bulk nanowire, identified as mobility fluctuations, is comparably much lower than the measured noise level related to the nanowire surface. This result is promising as mobility fluctuations set the lower limit of what is achievable within a material system.

Acknowledgments

”*There are three methods to gaining wisdom. The first is reflection, which is the highest. The second is limitation, which is the easiest. The third is experience, which is the bitterest.*”
- Confucius

I would like to take this opportunity to give a big thanks to all my collaborators for their time invested in me. My first directed appreciation goes to you Lars-Erik for initially believing in me. Lars-Erik has, beyond wide-ranging knowledge of the scientific field, a deep understanding of the world at large. Besides the transferred knowledge, I am grateful for having a source of inspiration. In your role as a leader, you manage to keep the group motivated by being the infectively enthusiastic, and you always seem to have a way to address seemingly harsh problems from a different angle of vision, turning weaknesses into strengths.

Moving on I would like to acknowledge my appreciation of my two assistant-supervisors. When visiting a conference together with Erik Lind, it is made clear that he has made quite an impression in the field, with people sometimes lining up for discussions. Few would question Erik’s technical authority as he has developed a sort of unique, intuitive understanding for the non-intuitive. I am grateful for all your teachings.

My other assistant-supervisor is the talented Henrik Sjöland who is not only an authority in his field but also has an astonishing ability to pick up detailed information in other scientific branches and have sense on how it would impact the performance of a circuit. I am grateful for all your input and happy knowing that the nanowire technology has now reached a higher level of maturity where your deep knowledge of circuit-design will become an even greater resource.

Continuing, I would particularly like to thank my PhD-student colleague Martin Berg for all of our collaborations. Together we have fought many battles and although we didn’t win all of them, in my book, which this is, we stood triumphant at last. Martin, you are a perfectionist by nature and with a rock-solid patience; your set of skills and ways of approaching problems have been an appreciated complement to mine.

Further, I would like to express my gratitude towards the rest of the group for all the exchange we have had and all the fun moments we have shared. Now, instead of writing short texts on each one of you describing your many talents and skills I will leave it at thanking you for our professional interactions. First I would like to acknowledge Sofia, Anil, Mattias, Jun, Claes, Kristofer, Micke, and Johannes for our collaborations concerning the nanowires. Johannes should also be acknowledged for his kindness of giving extensive feedback on this thesis (together with my supervisors and Sofia); Johannes exceptional eye for detail has significantly improved this thesis. Secondly, I would like to thank Lars, Carl, Anders, Marcus, and Mats, for all the inputs on measurements. Thirdly I would like

to thank Cezar, Elvedin, Jiongjiong, Aein, and Guntrade for knowledge exchange in all the meetings.

I would also like to thank Gunnar Malm at KTH for our collaboration concerning low-frequency noise measurements; I am appreciative of you Gunnar, not only for sharing your recognized expertise, but also for the warm welcome I received in Kista.

Continuing I would like to thank all the people involved in making my thesis work not only possible but in most cases effortless outside the actual scope, and the list is long. From the physics department my special thanks goes to those doing the good job at keeping the Lund Nanolab up and running: Ivan, Mariusz, Anders, Peter, George and Håkan. At Chalmers I would like to thank Vladimir for the help with all the depositions. From the electronics department I would like to thank those who have aided me administratively or with knowhow: Victor, Göran, Peter, Martin, Vince, Molund, Pia, Erik, Bertil, and Josef.

To all the people involved in the study council and the PhD board I would like to give my appreciation for all the important dedicated work, improving the situation for all the peers. During the time I was involved, I enjoyed interesting discussions and learned a great deal.

At last there are many more people, friends and family, of which I am of course especially grateful towards; thank you for being part of my life.

Table of Contents

Preface

I have devoted my last 5 years to investigating the future potential of InAs nanowire field-effect transistor technology as a mean to dramatically decrease power consumption while allowing operation at higher frequencies. As a technology demonstration, we have been able to fabricate a III-V nanowire RF circuit with GHz operation. Other investigations include intrinsic transistor transport metrics extraction and low-frequency noise characterization.

The thesis is built up around two parts where the first is an introduction and overview with seven chapters. Chosen topics from the research are explained in greater detail, and minor unpublished studies related to the thesis are discussed. The second part is the reprinting of published work.

List of Papers

Circuit Performance

I. InAs Nanowire MOSFET Differential Active Mixer on Si-substrate
Karl-Magnus Persson, Martin Berg, Henrik Sjöland, Erik Lind, and Lars-Erik Wernersson.
Electronic Letters (submitted)
I co-fabricated the devices, made the DC and RF characterization in collaboration, analyzed the data, and wrote the paper.

II. InAs Nanowire MOSFETs Implemented in Active Mixers
Martin Berg, Karl-Magnus Persson, Erik Lind, Henrik Sjöland, and Lars-Erik Wernersson.
Nano Letters (manuscript)
I co-fabricated the devices, made the DC and RF characterization in collaboration, took part in discussions, and co-edited the paper.

Transistor Metrics

III. InAs Nanowire Wrap Gate Transistors With f_t > 7 GHz and f_{max} > 20 GHz
Mikael Egard, Sofia Johansson, Anne-Charlotte Johansson, Karl-Magnus Persson, Anil Dey, Mattias Borg, Claes Thelander, Lars-Erik Wernersson, and Erik Lind
Nano Letters, **10**, 809-812 (2010)
I co-fabricated the devices and co-edited the paper.

IV. Extrinsic and Intrinsic Performance of Vertical InAs Nanowire MOSFETs on Si Substrates

Karl-Magnus Persson, Martin Berg, B. Mattias Borg, Jun Wu, Sofia Johansson, Johannes Svensson, Kristofer Jansson, Erik Lind, and Lars-Erik Wernersson
Transcripts on Electron Devices, IEEE, **60**, 2761-2767 (2013)
I co-fabricated the devices and developed new fabrication procedures, did the DC and RF characterization in collaboration, made the data-analysis, and wrote the paper.

Energy Traps and Low Frequency Noise

V. Low-Frequency Noise in Vertical InAs Nanowire FETs
Karl-Magnus Persson, Erik Lind, Anil Dey, Claes Thelander, Henrik Sjöland, and Lars-Erik Wernersson
Electron Device Letters, IEEE, **31**, 428-430 (2010)
I did the measurements and data-analysis, and wrote the paper.

VI. A High-Frequency Transconductance Method for Characterization of High-κ Border Traps in III-V MOSFETs
Sofia Johansson, Martin Berg, Karl-Magnus Persson and Erik Lind
Transcripts on Electron Devices, IEEE, **60**, 776-781 (2012).
I co-fabricated the devices and co-edited the paper.

VII. Surface and core contribution to 1/*f*-noise in InAs nanowire metal-oxide-semiconductor field-effect transistors
Karl-Magnus Persson, Gunnar B. Malm, and Lars-Erik Wernersson
Applied Physics Letters, **103**, 033508 (2013)
I co-fabricated the devices, did the measurements in collaboration, made the data-analysis, and wrote the paper.

Additional Papers

The contest of these papers are either considered to overlap the included papers, or to be peripheral regarding the main topic of this thesis.

VIII. Fabrication Technology for RF Circuit Implementation of Vertical III-V MOSFETs (oral presentation)
Karl-Magnus Persson, Anil Dey, Erik Lind, Claes Thelander, Henrik Sjöland, and Lars-Erik Wernersson
Swedish System-on-Chip Conference (*SSoCC*), IEEE, May 4-5 (2009)

IX. Low-Frequency Noise in Vertical InAs Nanowire FETs with L_g=35nm (oral presentation)
Karl-Magnus Persson, Erik Lind, Anil W. Dey, Claes Thelander, Henrik Sjöland, and Lars-Erik Wernersson

GigaHertz 2010, Lund, Sweden, March 9-10 (2010)

X. 35 nm Lg Vertical InAs Nanowire FETs: DC and Noise Performance (oral presentation)
Karl-Magnus Persson, Erik Lind, and Lars-Erik Wernersson
Materials Research Society Conference (*MRS*), San Francisco, CA, USA, April 26-29 (2011)

XI. Vad är en bra lärare?
Joakim Bood, Patrik Nyman, Karl-Magnus Persson, Moa Sporre, and Can Xu
Lärande i LTH, **13**, 3 (2011)

XII. High-Frequency Performance of Self-Aligned Gate-Last Surface Channel $In_{0.53}Ga_{0.47}As$ MOSFET
Mikael Egard, Lars Ohlsson, Mats Ärlelid, Karl-Magnus Persson, Mattias B. Borg, Fredrik Lenrick, Reine Wallenberg, Erik Lind, and Lars-Erik Wernersson
Electron Device Letters, IEEE, **33**, 369-371 (2012)

XIII. InAs Nanowires for High Frequency Electronics (oral presentation)
Karl-Magnus Persson, Sofia Johansson, Martin Berg, Anil Dey, Kristofer Jansson, Mattias Borg, Johannes Svensson, Claes Thelander, Henrik Sjöland, Lars-Erik Wernersson, and Erik Lind
GigaHertz 2012, Stockholm, Sweden, March 6-7 (2012)

XIV. Vertical InAs Nanowire MOSFETs with I_{DS} = 1.34 mA/μm and g_m = 1.19 mS/μm at V_{DS} = 0.5 V (oral presentation)
Karl-Magnus Persson, Martin Berg, Mattias Borg, Jun Wu, Henrik Sjöland, Erik Lind and Lars-Erik Wernersson
Device Research Conference (*DRC*), IEEE, 195-196, University Park, PA, USA, June 18-20 (2012)

XV. Performance Advantages of III-V on Si Technology for 60 GHz RF Circuits (oral presentation)
Karl-Magnus Persson, Erik Lind, Henrik Sjöland, and Lars-Erik Wernersson
Swedish System-on-Chip Conference (*SSoCC*), IEEE, Ystad, Sweden, March 18 (2013)

XVI. 1/f-noise in Vertical InAs Nanowire Transistors (oral presentation)
Karl-Magnus Persson, Martin Berg, Erik Lind, Lars-Erik Wernersson
International Conference on Indium Phosphide and Related Materials (IPRM), IEEE, 1-2, Kobe, Japan, May 19-23 (2013)

CHAPTER 1

1 Introduction

In all things of nature there is something of the marvelous.
-Aristotle

1.1 Motivation –A Diverse Future

The internet is arguably one of the most important inventions of all time. From the coinciding development of the internet and that of digital processors and cellular communication has sprung the continuously growing market for connected mobile computers known as smartphones. A modern smartphone incorporates hundreds of millions of transistors and is as powerful as a desktop computer from the mid-2000s, yet, at much lower power dissipation. Today, a considerable part of the semiconductor industry is driven by the low power mobile sector and yesterday's stationary computer market giants are eager to take part. To lower cost and reduce power consumption, the development has gone from smartphone solutions of many dedicated chips towards few multipurpose chips with different types of electronics made with the same fabrication technology. There is circuitry for digital processing, analog radio signal amplification, transmitting, and receiving, as well as circuitry to convert between analog and digital signals. The semiconductor industry for digital applications has been so successful in cost-effective miniaturization of its Si fabrication technology that it has become the universal fabrication technology for all electronics. In the recent years, however, the Si industry has, from one generation to the next, strayed away from its former rigid path of merely decreasing structure dimensions. The introduction of enhancement boosters, coping with dimensions reaching fundamental limitations, inherently differentiates the industry. Future market demands are in one direction pursuing ultra-low-power circuits with high cost efficiency, and in another, ever increasing operation frequencies within acceptable power restraints. With requirements becoming more diversified, more extensive technology branching is perhaps the unavoidable future where different technologies will target different purposes.

Looking ahead in the market segment for ultra-low power consumption, when fabrication cost and the consumed power of radio circuits can be sufficiently reduced, among other exciting possibilities is the Internet of Things [1]. The Internet of Things aims to connect everything in our lives from the walls of our

house to the milk carton in the supermarket by putting sensors with radio circuits on them. A fully connected world with billions of sensors can for starters give collaborative minimization of waste of resources by reducing energy consumption and food chain spillage.

Non-Si technologies are often referred to as exotic. Commercially available transistors made with compound materials from elements of group III and V (III-Vs), such as those implementing InAs technology, are expensive and exclusively found in applications with extreme requirements of speed. Present technologies for tremendously high-frequency circuits (defined as 300 GHz and above) is based on fast electron transport at the lack of an isolating barrier between the electron channel and the controlling gate, resulting in electrostatics not suitable for further channel reduction below 30 nm. Another aspect is that the integration area is tightly coupled with fabrication cost, and the area size difference between a contemporary tremendously high-frequency technology device and a contemporary digital Si technology device, can be up to a factor of a thousand.

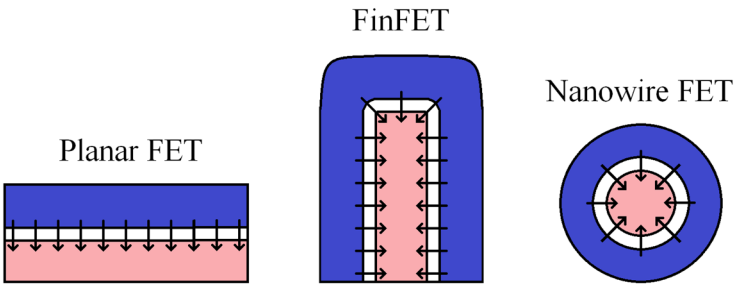

Fig 1.1. *Schematic comparison of three different MOSFET electrostatic architectures, where the arrows indicate the directions of field control. The blue material is a gate metal, the white material is an isolating dielectric, and the light red material is a semiconductor. The direction of the channel current flow is inwards into the figure. For the three structures, the electrostatic efficiency is increasing going from left to right* [2].

Nanowire metal-oxide-semiconductor field-effect transistors (MOSFETs) enable advantageous electrostatic control by asserting a potential on all sides of the channel [2], as shown in Fig. 1.1. With the incorporation of a gate-insulator, effectively functioning as a tall separating energy barrier, the technology offers possibility for high scalability along with low power applications. Vertical integration also decouples area-foot print and channel length thus possibly increasing integration density. Further, bottom-up nanowires, with its large surface to volume ratio, have several possibilities for crystal relaxation, making it viable for integration directly on a substrate with different atom spacing [3], or with the use a thin planar buffer [4]. If InAs nanowire MOSFET (NW-FET) technology could be fabricated on Si wafers alongside its Si counterpart, as a platform extension, it could become attractive not only for ultra-high frequencies, but also for certain low power applications. A first point of large scale application

convergence of different technologies is the broad future market of radio circuits in mobile devices in the non-licensed frequency band at around 60 GHz, a worldwide standard for high bitrate and short distance communication.

1.2 The Transistor – An Overview

The transistor is the fundamental building-block in electronic circuits that offer the possibility of logic computations and radio communication. The transistor is a device with three essential terminals where one is used to switch the conduction between the other two, on and off. For a computational circuit, the ability to have distinct conduction levels, enabling the digital representation of 1 and 0, is fundamental. For an analog radio circuit, the ability to selectively amplify a specific input signal is a key ability. A circuit can be assembled with discrete components, where transistors and passive elements (resistors, capacitors, and inductors) are linked together by metal stripes. For circuits with physically large components, such as protective power relay capacitors, this is common. However, when it comes to logic or radio circuits or more than a handful of not necessarily large components, integrated circuits are the preferred choice for reasons of practicality and cost. The fabrication of integrated circuit chips is very complicated with processing steps up to thousands and with fabrication facilities, by any measure, very expensive to setup. The main reasons integrated chips can still be bought at a relatively low cost are high integration density and mass production.

1.2.1 Evolvement

Although there are transistors incorporating mechanical switching, most transistors rely on switching merely with the movement of charge. In fact, any material that can be made to have semiconductor-like behavior can in principal be made into a transistor; for example, organic transistors are a growing research field [5]. The vast majority of transistors found in computers and cellphones, however, are based on just one element; Si. The fabrication of integrated Si transistors has been refined since its introduction under the workings of a process called scaling. Scaling is where the dimensions of a structure are miniaturized while keeping function intact. For transistors this means that the electrical field strength should be kept constant when shrinking the thickness of a constituent slab. The implication is that the applied voltage across the slab can be decreased proportionally. The gains of scaling include increased number of transistors in the same area, lowering of power consumption, and faster switching speed. In a paper from 1965, written by Gordon Moore [6], co-founder of Intel Corporation, it was stated that from an observation of the 7 years since the invention of the integrated circuit, it seemed as if the number of transistors per unit area doubled every two years. This statement is referred to as Moore’s law and became a self-fulfilling prophecy for the industry. Moore’s law has now been valid for over 50 years and as a result, device

dimensions of modern transistors can be measured in sub-hundred atoms across, with some layer thicknesses down to just a few monolayers. In fact, if scaling could continue, without the laws of physics setting a limit, the entire transistor would be made up of just a single Si atom in the year 2033. Understandably, other approaches than scaling have to be applied to continue enhancing performance; the trick is that whatever technology the future holds, it can only reach market appeal if it is cost effective in its niche of possible applications.

1.2.2 The Present and the Near Future

The latest transistor node is currently down to 22 nm meaning that within 100 mm^2 of chip area, it would be possible to fit up to 200 billion transistors. The actual transistor density on a chip is, however, nowhere near as dense as the theoretical maximum and is instead closer to 1% or less, mostly due to all sorts of interconnects. The largest commercial chips to date, measured in transistor count, have roughly 7 billion transistors at an area of around 500 mm^2 [7]. That number can be compared to the number of neurons in the human brain, around 100 billion [8]. Transistor density is also limited by the generated heat; there is an upper limit to how much heat that can be dissipated per unit area. The result of hitting the roof of heat dissipation about a decade ago is that instead of going for faster switching speeds for each generation, the gain of scaling has been taken out in pure power efficiency. Consumers might have noticed that computer processors have been operating at about the same frequency for several years and instead, increased performance comes from an increased number of computational units, often referred to as cores. Depending on the level of parallelization of a computer task, more or less cores are simultaneously active. With steadily increasing number of transistors and with many computer tasks needing only one or a few cores under a limited period of time, each transistor is spending more time in the off-state. A turned-off transistor is, however, not fully off due to thermally assisted conduction and thus it has a constant leakage. The leakage in the off-state causes sub-threshold power dissipation, which is increasing in proportion to the active power dissipation, and could eventually surpass the average energy consumed in the active mode. Concern over off-currents has fueled the research of so called steep slope devices, aiming to reduce the current in the off-state. Since about 7 years back, the industry has not only officially welcomed incorporation of structural and material novelties, but it has been required to make the transition from one generation to the next. Some enhancements, such as future steep slope devices may, however, not be desirable for high-end RF circuits if the on-performance cannot be substantially improved [9], and could eventually lead to platform extensions of specialized RF fabrication technologies.

1.3 Fundamental Electron Device Transport Mechanics

In electronic devices, electrons and holes are responsible for carrying a flow of charge, a current, and they are called carriers. Holes are in reality the absence of electrons but can still be treated as particles very similar to electrons, although with opposite charge. One profound scientific discovery was that all particles are dualistic in nature; their behavior shows characteristics of both a classical Newtonian object and a wave. For particles of matter, the relation derived by de Broglie links together particle momentum and corresponding wave length,

$$\bar{\lambda} = \frac{h}{p}, \tag{1.1}$$

where $\bar{\lambda}$ is the de Broglie wavelength, h is the Planck's constant, and p is the particle momentum [10]. Besides being one of the foundations of quantum mechanics, $\bar{\lambda}$ indicates the distance at which wave properties have a significant influence on the particle behavior. For a traveling electron or hole in a condensed matter, atomic distances are within the electron's de Broglie wavelength and hence, wave properties will determine how a carrier "perceives" its surroundings. The wavenumber, k, reciprocal to $\bar{\lambda}$, determines the direction of the spatial sinusoidal wave of an electron, Ψ. In one dimension, Ψ can be written as $\exp(ikx)$. Using Bloch's theorem, the solution of the Schrödinger's equation (see Eq. A1) in a lattice of atoms can be written as a periodic function U:

$$\Psi(x) = \exp(ikx)U(k, x). \tag{1.2}$$

In solid state physics, Eq. 1.2 is essential for energy band theory and constitutes the basics for electronic transport in semiconductors. The following sections will describe the movement of electrons in a more qualitative fashion.

1.3.1 Electronic Materials

Quantum mechanics can describe electron transitions between different atomic orbitals. One significant property is that allowed orbitals are found at certain electron-nucleus binding energies, specific for each chemical element. If a material is structured in periodic lattices, in crystalline form, the description of the possible energy states of an electron can be expanded to the form of bands, stretching through-out the crystal. Knowing the energy of the electron, it is thus possible to calculate inter atom displacement. An atomic unit cell, with which the periodical structure of a crystal can be described, is illustrated in Fig. 1.2a. It is the interactions of the atomic orbitals that generate continuous bands of energy states, distributed at different energy depths. Between bands of allowed energy states there

are gaps of non-allowed energy states. For a crystal lattice with 2 or 3 dimensions being significantly larger than $\bar{\lambda}$, the bands of allowed states typically overlap and all energies, except a single material specific gap between the bound and the free states, are accessible.

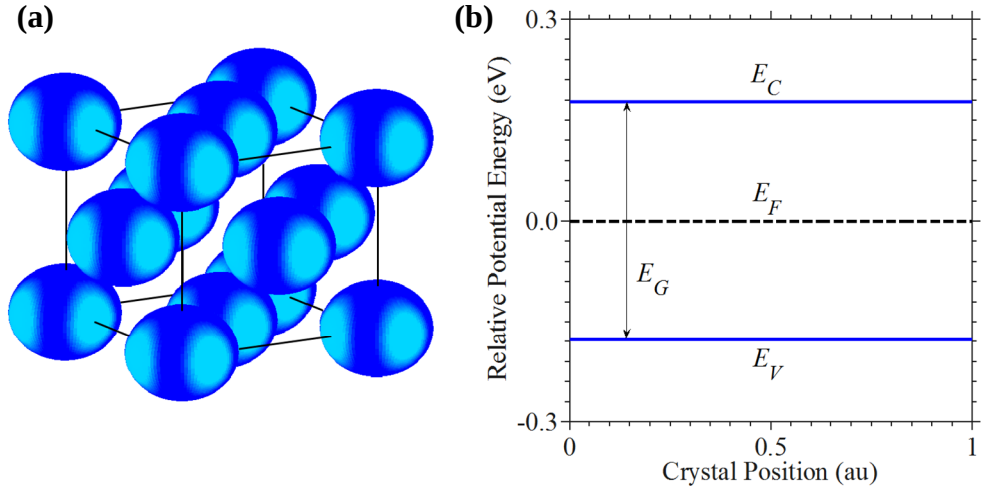

Fig 1.2. (a) *Atoms placed in a periodic lattice. The movement of electrons can be modelled using bands of accessible states, originating from periodic atomic orbitals.* **(b)** *Semiconductor band energy diagram displaying the valance band edge, E_V, and the conduction band edge, E_C, separated by the energy gap, E_G. The Fermi level of the material, E_F, is here situated in the middle of the gap.*

Atoms in covalent crystals, in contrast to ionic crystals known as salts, are held together by sharing the outer electrons, where neighboring atoms have shared orbitals. Any atom with an excess outer electron not used for binding with a nearby atom, contributes with a free electron, instead bound to the entire crystal. The opposite is likewise true for an atom lacking an outer electron to share, where this atom instead binds a borrowed electron and thereby contributes with a free hole carrier. In metals, all atoms have loosely bound valence electrons, one or two, and these gather in a shared mobile cloud, hence the concentration of free carriers is in the same order as the concentration of atoms. In isolators, on the other hand, valence electrons are so tightly bound to the atoms that the work needed to free an electron corresponds to several volts, and practically no atoms contribute towards a common carrier population. In between metals and isolators, there are the semiconductors. Semiconductors are characterized by that they can be manipulated to be either conductive or isolating; when looking at the band diagram of a semiconductor, there are bands of allowed energy states divided by a gap of non-allowed energy states. The two sides of the gap correspond to the valance band and the conduction band, respectively. A carrier can either occupy an energy state where it is bound to an atom, in the valance band, or it can be in a conducting energy state where it can travel freely, in the conduction band. When referring to the associated energy state of a carrier, it is in terms of probability as exact energy

and position cannot be simultaneously determined (due to Heisenberg's uncertainty principle). It is, in fact, desirable to leave out the particles altogether and focus on the probabilities of the possible energy states, whether or not they are occupied. To do this, statistical mechanics with a concept of a chemical potential referred to as the Fermi level is introduced. The Fermi level is the corresponding energy depth where there is an equal probability to find an energy state to be occupied and non-occupied, respectively. In a pure semiconductor crystal, called an intrinsic material, the Fermi level is situated roughly in the middle of the energy band gap, E_G, and this is illustrated in Fig. 1.2b. There are no energy states in the energy gap but the likelihood of finding a state to be occupied above the Fermi level is decreasing exponentially with increasing energy and if the distance between the Fermi level and the conduction band edge is large, there are hardly any free carriers. Similarly, at a decreasing energy depth, the probability of finding a state to be occupied goes exponentially towards one at the valance band edge. The likelihood $n(E)$ of finding an occupied state at the energy level of E is described by the Fermi-Dirac distribution:

$$n(E) = \frac{1}{1+\exp(\frac{E-E_F}{q\varphi_T})}. \qquad (1.3)$$

In Eq. 1.3, E_F is the Fermi level energy, q is the fundamental charge corresponding to that held by one single electron, and φ_T is the thermal voltage, $\varphi_T = k_BT/q$. T is the temperature and k_B is the Boltzmann's constant. One important aspect of semiconductors is that position of the Fermi level is not fixed but can be moved up or down with the introduction of doping. Doping is the insertion of atoms with more or fewer outer electrons than the inherent crystal atoms. An increase in doping concentration raises the Fermi energy and will move the Fermi level closer to either the conduction band edge, or the valance band edge, depending on doping type. When the Fermi level crosses either of the two band edges, the conductivity becomes linear in respect to the applied acceleration voltage, similar to the conduction in a metal. It is often desirable to have highly doped semiconductor regions when forming a contact to a metal, not to create a barrier in between. A semiconductor-metal junction without a barrier is referred to as an Ohmic contact while a junction with a barrier is called a Schottky contact.

If the dopant atoms contribute with electrons, the material is referred to as n-doped. For contribution of holes, the material is instead called p-doped. The doping level can be written as *N*, *N*-, *N*+, *N*++, etc., where the + and – signs indicate a two decade range, orders of magnitude higher or lower than *N*, which ranges between 10^{14}-10^{15} cm^{-3}. Modern electronics fabrication implements high precision doping control able to make not only sharp and distinct regions of high carrier concentrations but also graded doping profiles where the doping concentration is gradually increased along a certain axis.

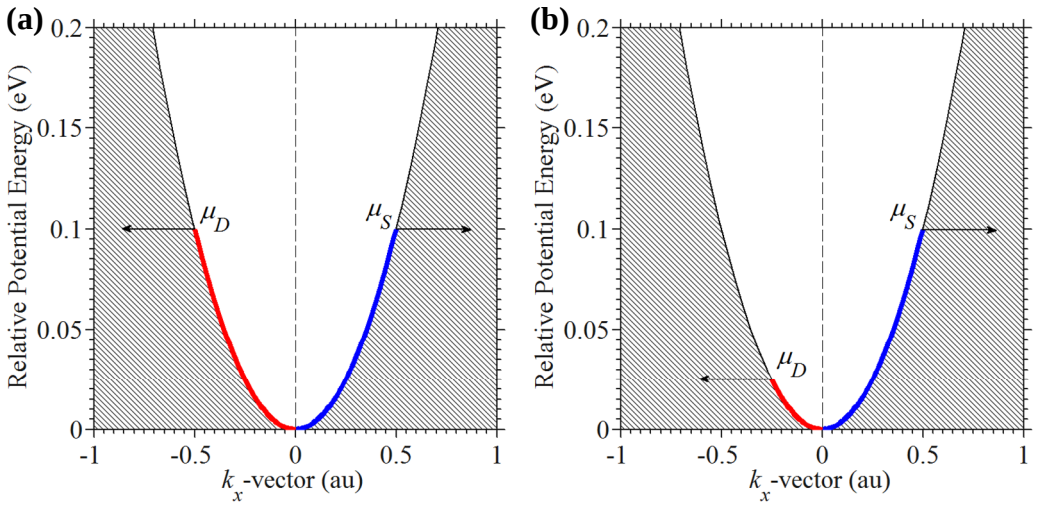

Fig 1.3. *The energy-distribution of electrons as function of k_x-vector in an arbitrary point along the x-direction in a crystal lattice. The available states are occupied up to the corresponding level of the potentials μ_S (blue) and μ_D (red) found at left and right of the spatial point, respectively.* **(a)** *μ_S and μ_D are equal and the average directional transport along the x-direction is zero. This would correspond to a flat conduction band as shown in Fig. 1.2b.* **(b)** *μ_S is larger than μ_D and the average directional transport is larger than zero in the positive x-direction. This would correspond to a tilted conduction band.*

1.3.2 **Electronic Transport**

Going back to the electron wave behavior, carriers do not only interact with surrounding atoms but may also interact with each other. For non-metals, however, the concentration of carriers being at least several magnitudes lower than the atomic concentration, it is low enough that the average distance between carriers is much longer than the de Broglie wavelength, and wave interactions can thus be disregarded. The entirety of carriers in a semiconductor is often referred to as a classical gas, as it can be treated with the same classical physics as a gas found elsewhere, such as in in the atmosphere, although with the difference that the gas of carriers is “trapped” in a crystal with allowed and non-allowed energies. A gas is called degenerate if the accessible energy states are filled according to the lowest energy configuration, where the possibility of spin up and down allows for two electrons at each energy state. A gas is fully degenerate at zero temperature, where all accessible states up to the Fermi level are occupied and none above. At a higher temperature, as described by Eq. 1.3, some states above the Fermi level are going to be filled as well, relative to the temperature. Carriers found in a state above the Fermi level are called non-degenerate. Non-degenerate carriers have an energy corresponding to the available thermal energy, while carriers at or below the Fermi level have an associated maximum energy corresponding to the Fermi energy (more on one of the implications in section 2.1).

Modern electronic devices are transitioning to confinements in two dimensions, meaning that the possibility for a carrier to move freely is only available one direction. Confinements affect the bending of the energy bands and will impact carrier transport. A nanowire, if thin enough, can be seen as a 1-dimensional (1-D) object and this will have substantial effects on its conductivity. Taking into consideration the effects of structural confinements, such as non-overlapping energy sub-bands, the transport in an approximately non-confined direction can still be calculated with the use of classical physics. Under the assumption of a travel direction greatly exceeding $\bar{\lambda}$, the movement of electrons and holes is attributed to processes called drift and diffusion. Surrounding thermal energy constantly provide carriers with a finite energy, elevating their associated energy state. Diffusion is a process where a carrier is relocating at random, which at a many-particle scale effectively counteracts any carrier concentration gradients. Drift on the other hand, is a transport process where a carrier is accelerated by a potential gradient, seen as a tilt in the energy band structure. The tilt can be induced by a difference in doping, or a heterojunction, where materials with different bandstructure are combined. A tilt can also be achieved by an externally applied field. The energy distribution of states as function of *k*-vector is shown in Fig. 1.3a and Fig. 1.3b for a crystal with and without an applied external field, respectively. In the case of an isolated semiconductor slab with a flat band structure, the net flow of current, measured over long enough time, will be zero as there will be equal amount of carriers traveling in both directions, and is referred to as equilibrium state. If, however, there is an applied potential difference, the band structure is going to be tilted and carriers will be less likely to diffuse in the direction of higher potential as it requires a larger thermal elevation. Unsurprisingly, it turns out that the requirement for a steady-state current flow is an external injection of carriers.

1.3.3 **Charge Control**

For semiconductor electrical devices, current control is achieved by direct or indirect applied voltages. In a charge based transistor, a controllable potential barrier is used to turn current flow on and off. Diffusive conduction (thermally fueled) will limit the steepness in turning a device off with an exponential dependence to the barrier height. The bipolar transistor, a device that is similar to two merged diodes, is constructed in a way that the diffusion process is the governing mechanism in all of its operation modes. The other, more common type of transistor device, the field effect transistor (FET), instead relies on drift conduction and carrier acceleration with the use of an applied field. The band diagrams of a FET for two different biases are shown in Fig. 1.4a and Fig. 1.4b. For the description of drift conduction, a parameter named mobility is introduced. Mobility relates to the average velocity of a carrier in relation the electric field strength and is dependent on the effective carrier mass. The effective mass is a complex theoretical construction related to the material energy band structure. To make a simplified explanation, the effective mass is an equivalent mass of the

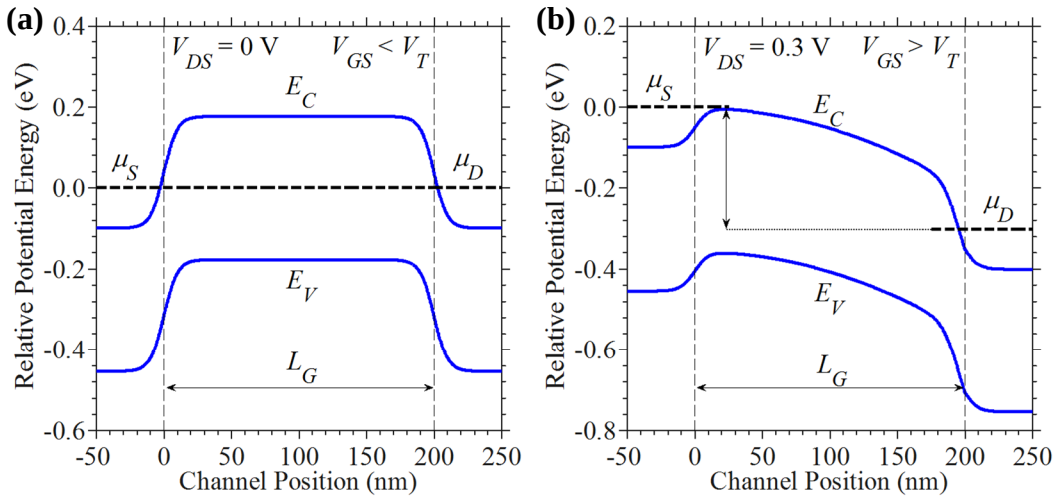

Fig 1.4. (a) *Band energy diagram of a FET with the channel contact potentials, μ_S and μ_D, levelled, and with a barrier in between with a height above the contact potentials. L_G is the channel length, here corresponding to 200 nm.* **(b)** *Band energy diagram of a FET with the contact potentials shifted by an applied drain-source voltage, V_{DS}.*

spread-out of a wave of a carrier, making it possible to calculate the transport of a wave as if it was a particle. The lower the effective mass, the less acceleration field strength is needed to reach a certain average velocity and the higher the mobility. Different semiconductors have different associated mobilities; bulk InAs has an electron mobility above 20,000 cm^2/Vs while, in contrast, the value for bulk Si is around 1,400 cm^2/Vs, meaning that it is, in theory, possible to transport an electron in a slab of InAs, at the same velocity, with less than 1/10 of the field-strength and 1/100 of the power ($P \sim V^2$) used for an electron in a slab of Si. This has long been the underlying motivation for low effective mass III-V transistors.

1.3.4 **Scattering**

For modern transistors, drift and diffusion is no longer sufficient to describe the transport characteristics since the channel is so short that the approximation of an infinitely long channel is no longer accurate. In a deeper analysis of drift conduction, transport consists of a series of accelerations and collisions, as shown in Fig. 1.5. The average distance between collisions is called mean free path, λ_{mp}. In the acceleration phase of the transport, the carrier undergoes a ballistic trajectile. Ballistic carrier transport well below a relativistic speed (the case in an electronic device) will follow the same principles as a much larger Newtonian projectile, like a cannon ball. One difference is that while a cannonball, after being shot out of the cannon, faces deceleration due to drag resistance, a carrier may instead be accelerated if it is traveling in an electrical field. Secondly, at the atomic scale, gravity is relatively a very weak force and so it is disregarded. Thirdly, a carrier

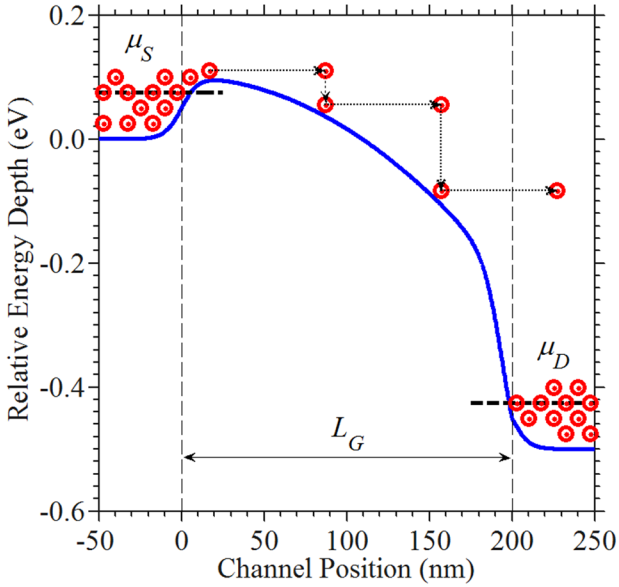

Fig 1.5. *The diagram shows the relative energy depth of the conduction band of a FET. The different contact potentials, μ_S and μ_D, are biased at different levels and separated by a barrier that constitutes the lower doped channel in between the contacts. Electrons, represented by red circles, are injected at μ_S and then drift towards μ_D. During the transport, the carriers endure sets of collisions in relation to the material mobility. As the channel is tilted proportionally to μ_S-μ_D, carriers are continuously and directionally accelerated towards the low potential.*

traveling ballistically in a crystal is better likened with a cannon ball being shot into a forest; sooner or later, the carrier encounters a collision. A carrier collision is referred to as a scattering event and there are multiple mechanisms in which a carrier can scatter. When scattered inelastically, the carrier may lose momentum and subsequently lower its energy state. If the carrier is instead scattered elastically, the carrier will only change its direction, where the direction of travel can be described with the wavenumbers k_x, k_y and k_z. If a carrier encounters an inelastic scattering event, it will subsequently lower its energy state. At the new energy state, if the carrier is in an electrical field, it will accelerate and start another trajectile.

A transistor channel length within a few λ_{mp} is called semi-ballistic as the entire travel distance is covered by a few ballistic transport events. If the travel distance is equal or less than the λ_{mp}, the transport is called fully ballistic. For a ballistic transistor, the mobility becomes somewhat peripheral as the time for traversing the channel is small compared to the mean time between scattering events. Instead, velocity at which the carrier is injected becomes dominating. Injection velocity is, however, like mobility also related to the effective mass and materials with high mobility typically also have high injection velocities. For a semi-ballistic device, both mobility and injection velocity are important.

1.4 The Metal-Oxide-Semiconductor Field Effect Transistor (MOSFET)

The commercial market workhorse with its favorable digital device performance is the metal-oxide-semiconductor field-effect transistor (MOSFET). By its broad application range and wide market adoption, the MOSFET effectively marginalizes all other technologies today. MOSFETs are fabricated with various architectures for various purposes and are often specialized for either low-power or high performance. Another classification is the type of carriers they conduct; a MOSFET with electron carriers is called an n-type while a MOSFET with hole carriers is called a p-type. Having both an n-type and a p-type MOSFET is vital for logic circuits, as explained in section 1.5.

1.4.1 The Planar MOSFET

Large-scale commercial integrated circuits are commonly fabricated on Si substrates due to its high cost efficiency. Wafers that measure up to 300 mm in diameter go through up to thousands of different processing steps, before being sliced into chips that are mounted on circuit boards. One of the initial steps in fabricating planar MOSFETs is to define the regions for each device by implanting low concentrations of dopant atoms, contributing with either free electrons or free holes, by bombarding the substrate with ions. To create the two basic contacts between which there will be a flow of current, namely the source and the drain contacts, ions are implanted once again, locally forming regions with a much higher carrier concentration. The contacts are doped with opposite charge compared to those from the ions previously implanted in the substrate. The reason for having the substrate contain carriers of opposite charge is that, for an n-type device, it effectively raises the conduction band around the source and drain regions, making the contact carriers face an uphill potential on all sides. At this stage, the source and the drain have no conduction path between them.

To keep the substrate at a fixed potential during operation, a body contact may also be introduced. The body contact is non-essential and it is not present in all types of FETs; to make a FET, a source-, a drain- and a gate electrode is needed. By applying a potential, V_G, on the gate electrode, a temporary path may be formed, a channel, between the source and drain. When conduction is switched on, V_G will be of the same sign as the charge of the substrate carriers, thus opposite sign to the charge in the contact regions. For an n-type device, a positive V_G will repel positive charges and eventually attract negative charges, something called inversion, as the charge sign of the carriers under the gate is inverted.

To avoid a current flow in and out of the gate, an energy barrier is introduced between the gate contact and the substrate. If the barrier consists of an insulator and the gate is of metal, the material stack under the gate is referred to as metal-oxide-semiconductor, MOS. The low leakage properties of the MOSFET makes it the

preferred choice for digital applications, enabling very-large-scale-integration (VLSI). For an n-type MOSFET, a strong enough positive voltage on the gate will at a critical level attract a sheet of negative charges, an inversion layer, in the channel close to the oxide surface. The critical field occurs at the threshold voltage, V_T. An applied gate voltage, V_G, above V_T will increase the depth of the channel and thus lower the resistance between source and drain. To get a current flow between source and drain, I_{DS}, there must be a difference between the source potential, V_S, and the drain potential, V_D, providing an accelerating voltage, $V_{DS} = V_D - V_S$. For an n-type MOSFET, if V_{DS} is larger than the net voltage above threshold at the gate, $V_G - V_T$, the excessive voltage, $V_{DS} - V_G - V_T$, will instead drop in the end of the channel region, at the drain side, and thereby not contribute to further carrier acceleration. This is referred to as pinch-off and one outcome is that the drain contact region will start expanding in proportion to V_{DS} and thus shorten the effective channel length, something that can be problematic for short channel devices.

A device biased above pinch-off is said to be in saturation, in active mode, as an increase of the V_{DS} will, ideally, not increase the I_{DS}. With few important exceptions (digital logic), a device should always be in saturation so that current is only controlled by V_G. If a FET is biased below saturation but above threshold, it is said to be in the triode region. In the triode region, the FET acts as a gated resistor

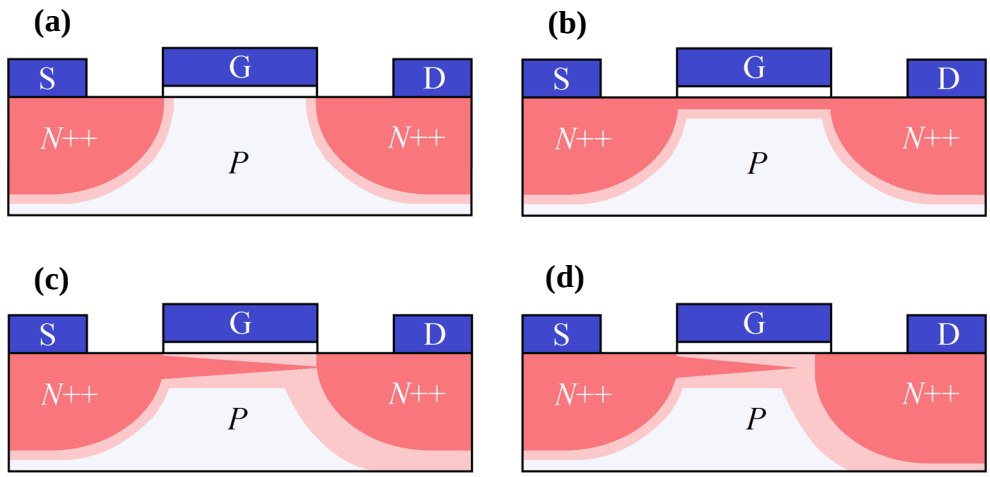

Fig 1.6. *Schematic illustrations of different biasing conditions of a planar MOSFET. The device has metal deposited in direct contact with highly doped (N++) regions at the source, S, and the drain, D. The lower doped channel (P) is controlled by the potential in gate, G, where the metal contact is separated by a dielectric film. Between the positive and negative charges there is a depletion layer.* **(a)** *No gate bias is applied and there is no conducting path between S and D.* **(b)** *A gate bias beyond threshold is applied, meaning that an inversion layer is formed at the channel surface.* **(c)** *The gate is biased at a higher potential than in (b) and the potentials of S and D differ to an extent corresponding to channel pinch-off.* **(d)** *The potential difference between S and D is increased beyond pinch-off and thus the effective channel length is reduced.*

where current control is achieved with all potentials. In Fig. 1.6, a MOSFET under 4 different bias conditions (a, b, c, and d) is shown, illustrating the spatial charge concentration.

1.4.2 A Classical MOSFET Model

The description of the behavior of a FET can be accomplished with different types of models, depending on the purpose. To develop a model, a first step is to identify the nodes and potentials of importance. In Fig. 1.7a, a model of the MOS interface is shown, indicating the different potentials of the band structure; if the portrayed spatial axis is in the transverse direction, then I_{DS} flows in the perpendicular, longitudinal direction, or the *x*-direction. For simplicity, $x = 0$ is the channel at the source contact and $x = L_G$ is the channel at the drain contact. Assuming that the transverse field is much stronger than longitudinal field in every point along the channel, the channel in the longitudinal direction can be seen as a series connection of infinitely small resistive elements, where the bulk Fermi potential, $E_{F,S}(x)$, at each node is set by the superposition of V_S and V_D [10]. The total channel charge area density, $Q_{Ch}(x)$, can be thus described as the superposition of the charge area density attributed to the gate-source voltage, V_{GS} and the gate-drain voltage, V_{GD}, Q_S(x) and Q_D(x), respectively:

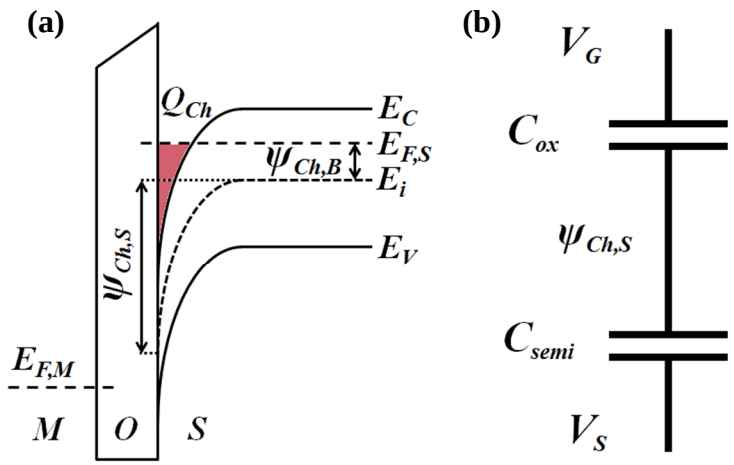

Fig 1.7. (a) *A schematic illustration of the MOS interface for an n-doped MOSFET channel, biased at a positive gate-metal potential, $E_{F,M}$. The channel charge, Q_{Ch}, has accumulated at the semiconductor surface. The surface potential, $\psi_{Ch,S}$, is the difference between the intrinsic potential, E_i, in the bulk semiconductor and at the channel surface. The bulk potential, $\psi_{Ch,B}$, is the difference between the intrinsic level and the semiconductor Fermi level, $E_{F,S}$.* **(b)** *A schematic illustration of the most important nodes and the different capacitors between gate and source. By replacing V_S with V_D, the illustration would instead show the most important potentials and series capacitances between gate and*

$$Q_S(x) = C_G(V_{GS} - \psi_{Ch,S}(x)), \quad (1.5)$$

$$Q_D(x) = C_G(V_{GD} - \psi_{Ch,S}(x)). \quad (1.6)$$

In Eq. 1.5 and 1.6, $\psi_{Ch,S}(x)$ is the semiconductor surface potential and C_G is the total gate capacitance. Fig. 1.7b shows a simplified model of the most important potentials and capacitances between gate and source. The semiconductor capacitance, C_{semi}, corresponds to the derivative of the channel charge in respect to $\psi_{Ch,S}$. C_G is the series connection of the geometrical oxide capacitance density, C_{ox}, and C_{semi}.

The current flow in a MOSFET can be modeled with two most significant parameters; the average channel charge area density, Q_{Ch}, and the average carrier velocity, v_{Avg}. For a long-channel FET, the current can be described by:

$$\frac{I_{DS}}{W} = Q_{Ch} v_{Avg}. \quad (1.7)$$

In Eq. 1.7, the device current I_{DS} is normalized with the width of the channel, W. Q_{Ch} can be calculated by integrating Q_{Ch}(x) over the channel. If $Q_{S,i} = Q_S(0)$ and $Q_{D,i} = Q_D(L_G)$, then:

$$Q_{Ch} = \frac{1}{L_G}\int_0^{L_G}\left[Q_{S,i}\left(1 - \frac{x}{L_G}\right) + Q_{D,i}\left(\frac{x}{L_G}\right)\right]dx = \frac{Q_{S,i}+Q_{D,i}}{2}. \quad (1.8)$$

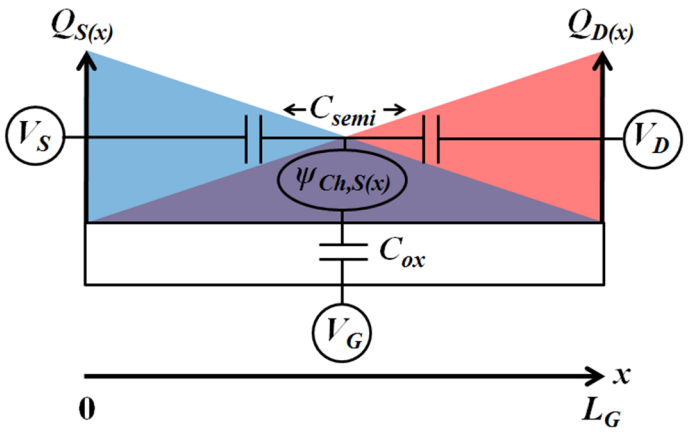

Fig 1.8. *A schematic illustration of a MOSFET channel with indicated voltage terminals, V_G, V_S, and V_D, channel surface potential, $\psi_{S,Ch}$, oxide capacitance, C_{ox}, the semiconductor capacitance, C_{semi}, and the channel charge density associated with the source, Q_S, and drain, Q_D. The illustrated device is in equilibrium (V_S = V_D), where the charge is conformally distributed along the channel.*

In equilibrium ($V_S = V_D$), the charge area density equals the average in every point along the channel and thus, $Q_{Ch} = Q_{S,i} = Q_{D,i}$. It is the difference in charge area density between Q_S and Q_D in combination with injection of carriers that give rise to a steady state drift current, as explained in section 1.3. Fig. 1.8 shows Q_S(x) and Q_D(x) from $x = 0$ to $x = L_G$ for a channel in equilibrium.

The second factor of the right-hand-side of Eq 1.7, v_{Avg}, can be expressed as:

$$v_{Avg} = \mu_{eff} \frac{V_{Acc}}{L_G}. \quad (1.9)$$

In Eq. 1.9, μ_{eff} is the effective mobility and V_{Acc} is the carrier acceleration voltage corresponding to potential difference between the drain and source contact. Using Eq. 1.5 and 1.6, V_{Acc} can be translated into an equivalent charge difference between $x = 0$ and $x = L_G$:

$$V_{Acc} = \frac{Q_{S,i} - Q_{D,i}}{C_G}. \quad (1.10)$$

Inserting Eq. 1.10 in Eq. 1.9 gives:

$$v_{Avg} = \mu_{eff} \frac{Q_{S,i} - Q_{D,i}}{C_G L_G}. \quad (1.11)$$

By inserting Eq. 1.8 and 1.11 into 1.7, the device current can be expressed as:

$$\frac{I_{DS}}{W} = \mu_{eff} \frac{Q_{S,i} + Q_{D,i}}{2} \frac{Q_{S,i} - Q_{D,i}}{C_G L_G}. \quad (1.12)$$

With a simple rearrangement, Eq. 1.12 can be altered into:

$$\frac{I_{DS}}{W} = \mu_{eff} \frac{1}{2 C_G L_G} (Q_{S,i}{}^2 - Q_{D,i}{}^2). \quad (1.13)$$

As described in section 1.4.1, V_T corresponds to the gate potential at which a first layer of carriers are accumulated (or inverted) at the channel surface. V_T can be calculated using the potentials shown in Fig. 1.7, or it can be estimated from a measurement. To simplify the modelling and avoid band structure calculations, $Q_{S,i}$ can be described with an approximate MOS-function expression [11], using an estimated V_T instead of $\psi_{Ch,S}$:

$$Q_{S,i} = C_G \varphi_T \left(\ln\left(1 + \exp\left(\frac{(V_G - V_T)/\gamma - V_S}{\varphi_T}\right)\right)\right). \quad (1.14)$$

The right-hand-side of Eq. 1.14 describes the potential difference both in sub-V_T and above-V_T operation, where the term $Q_{S,i}$ converges to $C_G(V_G\text{-}V_T\text{-}V_S)$ high above

threshold and to $C_G(\varphi_T)$ for well below, where the thermal voltage, $\varphi_T = k_BT/q$. The γ-parameter is a combination of a Fermi function, F_f, and the inverse sub-threshold slope (SS) factor, n;

$$n = 1 + C_D / C_i, \tag{1.15}$$

where C_D is C_{semi} in depletion and C_i is the parallel parasitic capacitance induced by interface energy traps (more about energy traps in chapter 4). The γ-parameter and F_f are given by:

$$\gamma = nF_f + (1 - F_f), \tag{1.16}$$

$$F_f = \frac{1}{1+\exp((V_G - V_T)/\varphi_T))}. \tag{1.17}$$

F_f mitigates a transition between sub-V_T and above-V_T operation as γ takes the value of n at sub-V_T and unity at above-V_T. Regarding the depletion capacitance, in in an ideal transistor there is a sharp transition of C_{semi} towards zero when decreasing the gate potential below V_T. This is the case for a small C_D/C_i ratio, where n thus is close to unity. However, if there is a large trap density around the conductance band edge, n will be considerably large than 1 and it will thus require a large negative voltage, relative to V_T, to deplete the channel such that it is considered off. This would then compromise the constraints of the voltage envelope, leakage and possibly the switching time constant.

The description of $Q_{D,i}$ can done in the same manner as for $Q_{S,i}$. Using Eq.

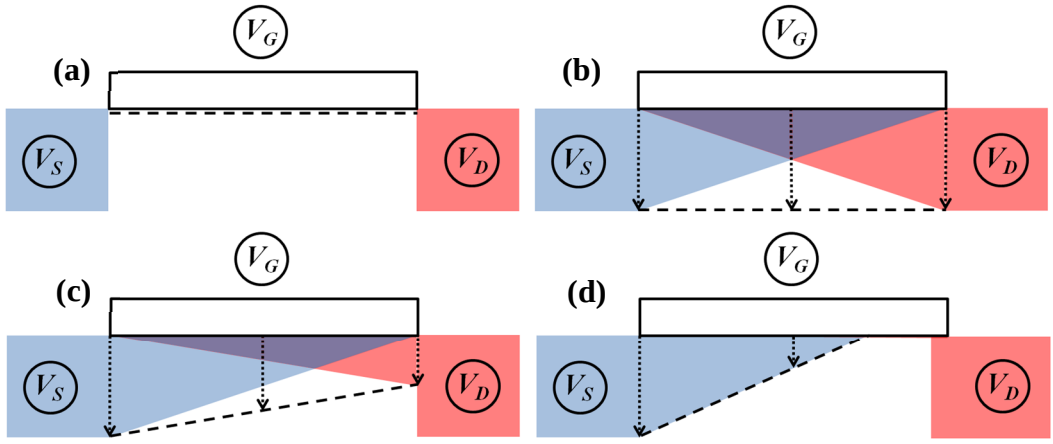

Fig 1.9. *Schematic illustrations of different biasing conditions of a planar MOSFET. The positive channel area charge density axis points downwards in the figures and the charge density at x = 0, x = L_G/2, and x= L_G are indicated. The blue area corresponds to Q_S, while the red area corresponds to Q_D; the striped line is the superposition of the two.* **(a)** *V_G is below V_T.* **(b)** *V_G is above V_T and V_S = V_D.* **(c)** *V_G is above V_T and V_D > V_S.* **(d)** *V_G is above V_T and V_{DS} is beyond pinch-off.*

1.14 and an equivalent expression for $Q_{D,i}$, for operation above V_T, Eq. 1.13 can be rewritten as:

$$\frac{I_{DS}}{W} = \mu_{eff}\frac{C_G}{2L}((V_{GS}-V_T)^2-(V_{GD}-V_T)^2). \qquad (1.18)$$

For a FET in saturation, $Q_{D,i}$ goes towards zero (pinch-off) and thus the current is effectively controlled by V_{GS}^2, and this is true for devices with a low V_{DS}/L_G ratio. The pinch-off behavior can also be described by introducing a saturation voltage, V_{sat}, where $V_{DS} = V_{sat} = V_{GS}\text{-}V_T$ for $V_{DS} > V_{GS}\text{-}V_T$. The superposition of $Q_S(x)$ and $Q_D(x)$ for 4 different biases, including pinch-off, is shown in Fig 1.9a, b, c, and d, respectively.

1.4.3 **Velocity Saturation**

For a FET with $L_G >> \lambda_{mp}$, where the current is reduced by scattering along the channel, there are other effects than pinch-off that causes saturation. When the acceleration field reaches a critical level, the carrier velocity will saturate due to carrier-lattice interactions. To account for velocity saturation in the classical model, Eq. 1.13 must first be split into the individual contributions, the carrier density and the carrier velocity, respectively, as in Eq. 1.12. For above-V_T operation, the normalized drain current is given by:

$$\frac{I_{DS}}{W} = \mu_{eff}\frac{C_G}{2L}((V_{GS}-V_T)+(V_{GD}-V_T))\cdot ((V_{GS}-V_T)-(V_{GD}-V_T)). \qquad (1.19)$$

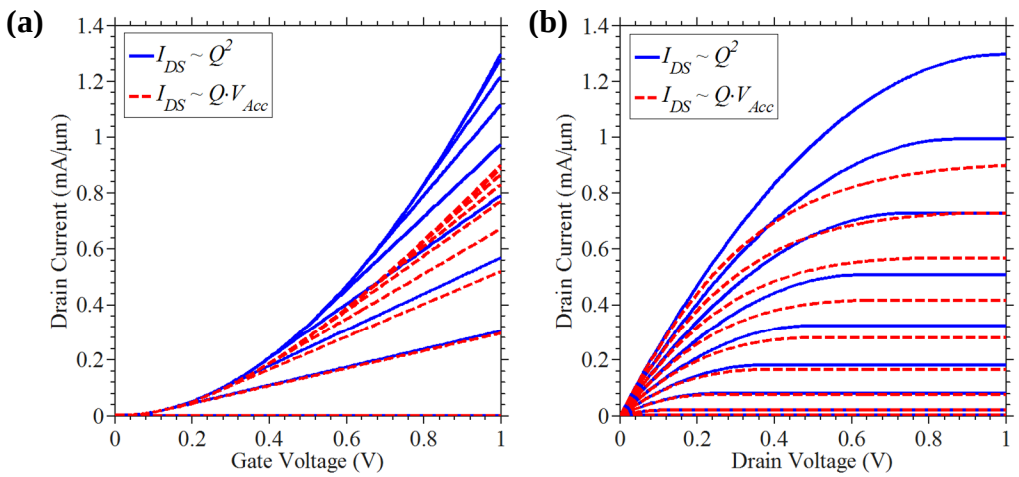

Fig. 1.10 *Simulated data of an InAs NW MOSFET with 600 nm channel length, illustrating the impact of velocity saturation (the red striped curves compared to the blue solid curves). The data is normalized to the nanowire circumference. It should be noted that the graphs show intrinsic performance without any added series resistance.* **(a)** *Transfer characteristics.* **(b)** *Output characteristics.*

Using Eq. 1.7 and 1.9, Eq. 1.19 can be rewritten as:

$$\frac{I_{DS}}{W} = \frac{C_G}{2}\left((\mathrm{V}_{GS} - V_T) + (\mathrm{V}_{GD} - V_T)\right)\left(\mu_{eff}\frac{V_{Acc}}{L_G}\right). \quad (1.20)$$

The current saturation in Eq. 1.20 can now be modeled by introducing a saturation function of V_{Acc}:

$$V_{Acc} = \frac{V_{DS}}{1+V_{DS}/V_{Sat}}. \quad (1.21)$$

In Eq. 1.21, V_{Sat} is the corresponding voltage that when multiplied with μ_{eff} gives the saturation velocity. A device that is biased in velocity saturation will have a linear current dependence, $I_{DS} \sim V_{GS}$, differing from what was shown for the square dependence, $I_{DS} \sim V_{GS}^2$, for a device that is saturated solely by pinch-off. The transfer characteristics, I_{DS} as a function of V_{GS}, and the output characteristics, I_{DS} as a function of V_{DS}, are plotted in Fig. 10a and b, respectively, for the described MOSFET model, with and without velocity saturation.

The concept of saturation velocity is applicable for intermediately long channel devices, however, for semi-ballistic devices, if the transition time between the point of injection and point of unlikely return (explained in chapter 2) is short enough compared to the time scale for carrier-lattice interactions, velocity can overshoot (not restricted by saturation), and thus velocity saturation can be disregarded. A consequence of this is that in semi-ballistic transport, the current is primarily affected by scattering in the beginning of the channel. For such a device, the velocity of interest is instead at the point of injection.

1.4.4 **Short-Channel Effects**

For a more complete FET description, the effects of drain-induced barrier lowering (DIBL) and channel-length modulation can also be introduced. DIBL is where the drain potential impacts the curvature of the conduction band structure and thereby change the gate-barrier height, effectively lowering the threshold value proportionally to V_{DS} and a δ-parameter:

$$\mathrm{V}_T = \mathrm{V}_{T0} + \delta V_{DS}. \quad (1.22)$$

Another effect with similar impact on the transistor characteristics is the channel length modulation, where the channel is shortened due to expansion of the drain region beyond pinch-off. This is modeled with a λ-parameter that is zero below pinch-off and some measured or estimated value above:

$$\mathrm{I}_{DS} = \mathrm{I}_{DS}(1 + \lambda(V_{DS} - V_{Sat})). \quad (1.23)$$

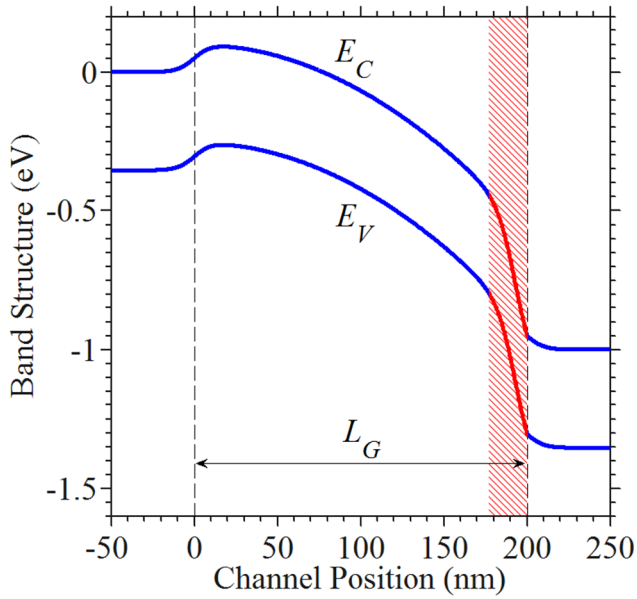

Fig 1.11. *Band energy diagram indicating in red the high-field region. The effects of impact ionization, where free electrons can knock out bound electrons, in turn creating a carrier avalanche, occur if the channel is biased steep enough.*

A MOSFET in pinch-off is shown in Fig. 1.6d and 1.9d. The two effects described in Eq. 1.22 and 1.23 are referred to as short-channel-effects (SCE). The later of the two will, however, have a minor impact in semi-ballistic devices as the channel resistance is only vaguely dependent on the channel length (explained in chapter 2). Another SCE present in narrow bandgap devices is impact ionization (IMIO), a phenomenon where high momentum carriers may knock out bound valance electrons and thus create a hole-electron pair. If the high field region is of considerable length, both the original and the knocked-out carrier may in turn knock out another pair and this process may go on, creating an avalanche effect with an exponentially increasing current, possibly physically damaging the device. The region of IMIO is indicated in Fig. 1.11. Modeling of IMIO can be accomplished by introducing a critical field, F_C, above which impact ionization occurs. F_C is directly related to the semiconductor bandgap energy, E_G, as $F_C = E_G/(qL_{ii})$, where L_{ii} is the length of high-field region where the IMIO occurs. As a rough estimate, it is possible calculate the field strength by simply using the potential difference between the drain and gate potential, V_{DG}, times a fitting parameter α [12]. The description of the total device current, including the contribution of the IMIO, can thus be expressed as;

$$I_{DS} = I_{DS}(1 + \beta e^{(\alpha V_{DG})/(F_C L_{ii})}), \qquad (1.24)$$

where β is another fitting parameter. By exchanging F_C with $E_G/(qL_{ii})$, 1.24 can be rewritten as:

$$I_{DS} = I_{DS}(1 + \beta e^{q(\alpha V_{DG})/E_G}). \quad (1.25)$$

As with other tunneling dependent processes, due to the finite response time in band-to-band tunneling (BTB) of carriers, IMIO will not affect high frequency gate potential switching, but can still have a detrimental effect in RF applications as it may increase circuit DC power consumption.

1.5 **CMOS**

Logic circuits are relying on its fundamental operator, the inverter, with which any other logic can be constructed. To fabricate an inverter, two complementary transistors are needed; the n-type and the p-type. The idea is that a high voltage, representing a digital 1, switches the n-type transistor on and the p-type transistor off, while a low voltage, representing a digital 0, switches the n-type transistor off and the p-type transistor on. An inverter consist of a pair of complementary transistors in series, with the output placed in between them, and with their gate controls connected in parallel from a common input, shown in Fig. 1.12. Si complementary MOSFETs (Si-CMOS) also has relatively good performance for analog RF applications and with its unrivaled cost efficiency, most commercial radio circuits of today are made with Si-CMOS. Besides the term CMOS, these devices are also referred to using its subgroups, n-type MOSFETs (nMOS) and p-type MOSFETs (pMOS). Lately, with the progress of III-V nMOS, it has been a widespread effort to develop good III-V pMOS devices for CMOS integration. There are however numerous of associated problems. One major issue is that even though III-Vs have high electron mobilities compared to Si, hole mobilities in III-Vs are at best comparable to SiGe. So far, no work has been published demonstrating III-V pMOS performance coming anywhere close to that of reported III-V nMOS. If the pMOS cannot come reasonably close to that of the nMOS, it will consume too much area, as the channel has to be made wider to compensate for the lower on-current, and hence, compromise on the important integration density.

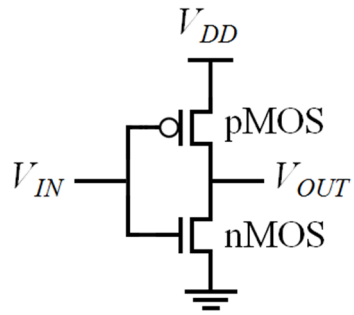

Fig 1.12. *A schematic layout of an inverter circuit, implementing a complementary MOSFET (CMOS) pair.*

CHAPTER 2

2 Modern MOSFET Modeling

Just remember, once you're over the hill you begin to pick up speed.
- Arthur Schopenhauer

A compact model can provide a way to predict the performance of an electronic device. It can also be used to extract intrinsic properties or explain certain behaviors of a measured device. Computation speed is an important aspect and a compact model is often the preferred choice for computer heavy investigations, such as how a modeled device behaves as a component in a larger network. At the other end of the spectrum, a device model can be constructed by a cross-sectional sheet of atoms, where quantum mechanics is applied to calculate the interactions between the electron orbitals of each atom and its nearest neighbors [13]. The result is the position and curvature of accessible energy bands, which can then be used for calculations with classical physics. Atomistic modeling gives an accurate description at the lower end of the nanoscale. This type of model also offers the possibility of prediction when the structure of interest is drastically altered. The disadvantage is that the model becomes increasingly complex and computer intense with the number of atoms. For a relatively large structure, where the positions of the energy sub-bands do not significantly differ from those of an infinitely large structure, the use of averages of quantum probabilities, converging with Newtonian derived analytical expressions, can be a better alternative. Using a model built upon averages can also offer the possibility of prediction for moderate structure alterations, as long as the trends of the limiting factors, with regards to scale, are accounted for.

2.1 The Virtual Source Model

Modern high performance FET devices are scaled to a point where its behavior has departed from that of the classical FET model. In an effort to give better understanding and get more accurate modeling, the analytical description has been reassessed by identifying the transport bottlenecks associated with short electron traveling distances [14] [15] [16] [17] [18]. As described in chapter 1, for channels that correspond to a few λ_{mp} or less, it is necessary not only to consider the drift velocity but also the velocity of injection, as the latter is a strongly limiting factor

when the conduction approaches ballistic or semi-ballistic behavior. The name of the virtual source (VS) model implies that the point of most significance to the conduction is that corresponding to the highest peak in the potential landscape, going from the source into the channel, see Fig. 2.1. This point is referred to as the top of the barrier or the virtual source. It has been shown that by knowing the average velocity which an electron is injected at the virtual source, v_{inj}, the saturation current of a short channel FET can be accurately modeled [Lundstrom]. While the current in a long channel device is dependent on the transport characteristics along the entire channel, for a semi-ballistic or ballistic channel, velocity can overshoot (not being affected by scattering causing velocity saturation), and is then only limited by the slowest average velocity, which is at the point of injection.

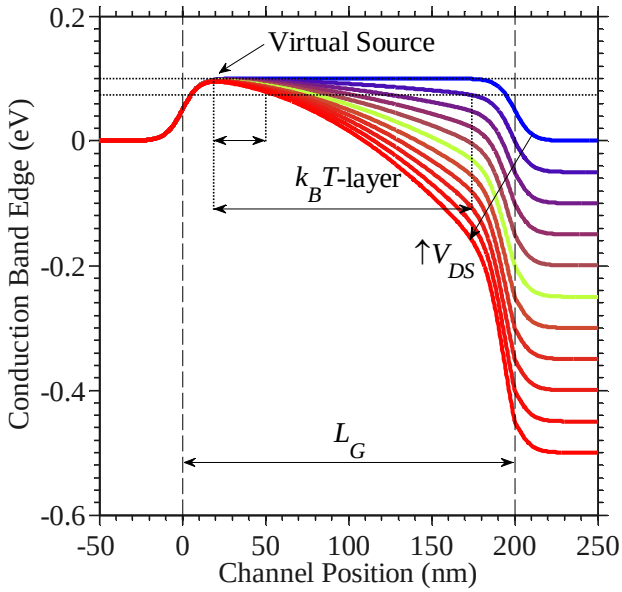

Fig. 2.1. *Energy band diagram showing the relative energy depth and slope of the bottom of conduction band for a range of different drain biases. The two dashed lateral lines indicate the thermal energy, k_BT, at room temperature, referenced to the point of the virtual source. By increasing V_{DS}, the point beyond which a carrier cannot be backscattered to the source moves to the left. The saturation of the drain current occurs when the k_BT-layer-width is shorter than λ_{mp}.*

To account for the behavior below saturation, some models choose to include a back scattering probability component to deal with scattering in the beginning of the channel. Scattering leads to a randomization of travel direction and carriers traveling in the opposite direction effectively reduce the current. However, only carriers that scatter in the beginning of the channel, both elastically and inelastically [14], can gain sufficient energy to make it back to the source. To model this, the notion of the first k_BT-layer is used. The k_BT-layer is the region after the virtual source point where, under a drain bias, the curvature of potential in the

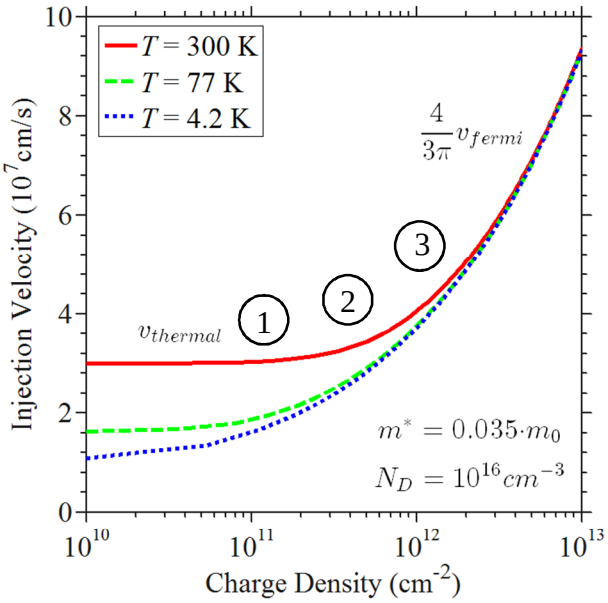

Fig. 2.2. *Injection velocity versus area charge density for a ballistic InAs 2D-channel. In the simulation, a channel doping of 10^{16} cm^{-3} and an average effective mass of $0.035 \cdot m_0$ was used. The indicated positions along the red curve (T = 300 K), circled 1, 2, and 3, correspond to a source Fermi level position as shown in Fig. 2.3a, b, and c, respectively.*

channel has dropped an equivalent amount to that of the available thermal energy. This is illustrated in Fig. 2.1. Using the k_BT-layer width approach, the effective injection velocity can be modeled as the product of the average unilateral velocity of carriers at the point injection multiplied with a zero-pole-function of the reflection probability R:

$$v_{inj} = v_{avr}\left(\frac{1-R}{1+R}\right). \qquad (2.1)$$

In Eq. 2.1, v_{avr} is the average unilateral velocity at the virtual source, which is described by:

$$v_{avr} = v_{thermal}\left\{\frac{\mathcal{F}_{1/2}(\eta)}{\mathcal{F}_0(\eta)}\right\}. \qquad (2.2)$$

In Eq 2.2, $\mathcal{F}_k$ is the Fermi-Dirac integral of order k (see Eq. A2), $\eta = (E\text{-}E_C)/(k_BT)$, E is the Fermi level position from bulk to surface, and $v_{thermal}$ is the thermal velocity given by:

$$v_{thermal} = \sqrt{\frac{2k_BT}{\pi m^*}}. \qquad (2.3)$$

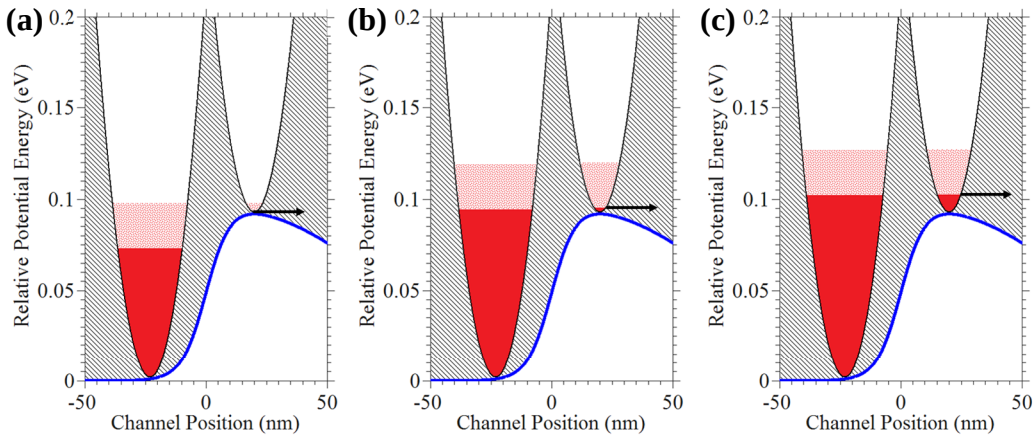

Fig. 2.3. *The k-vector along the channel direction is shown at the source contact and at the virtual source, for 3 different positions of the source Fermi level, E_F. The black arrow indicates the mean relative energy of the injected carriers.* **(a)** *E_F is below the top of barrier and the mean energy corresponds to the thermal energy.* **(b)** *E_F is slightly above the top of the barrier and the mean energy is also slightly above the thermal energy.* **(c)** *E_F is significantly above the top of the barrier and the mean energy corresponds roughly to the Fermi energy.*

For a channel barrier position relatively high above the source Fermi level, injected carriers are non-degenerate and the expression in the bracket of Eq. 2.2 is close to unity, hence the carrier velocity equals $v_{thermal}$. When the channel barrier is lowered such that source Fermi level is slightly above the barrier, degenerate carriers start getting injected and as the barrier is further lowered, the average carrier velocity is going to approach the Fermi velocity, v_{Fermi}, times a scalar;

$$v_{avr} = \left(\frac{4}{3\pi}\right) v_{Fermi}. \qquad (2.4)$$

The average velocity at the beginning of a ballistic InAs 2D-channel is shown in Fig. 2.2, for 3 different temperatures.

A carrier that is scattered in a point beyond the k_BT-layer width is unlikely to travel back to the source and exit, and can thus be disregarded. For a semi-ballistic transistor, at the bias point where λ_{mp} starts to exceed the k_BT-layer width, the injection velocity is going to saturate. This corresponds to when the expression in the parenthesis of Eq. 2.1 goes towards 0.5, or R ≈ 0.3 [14]. After reaching saturation, the reflection constant is only going to slowly reduce with increased drain potential as only a small portion of the additional applied field will fall over the channel region close to the source. If the v_{avr} is equal to $v_{thermal}$, in saturation, v_{inj} is roughly half of $v_{thermal}$.

A way to model backscattering is to have a V_{DS}-dependence on the injection velocity, where the corresponding saturation voltage, V_{Sat}, is a function of the channel mobility (and thus λ_{mp}), coherent with the k_BT-layer approach [Ant 2009]. With altered definitions, the current in ballistic and semi-ballistic devices can then be modeled in a similar fashion as to the classical FET description,

$$\frac{I_{DS}}{W} = Q_{VS} v_{inj}, \tag{2.5}$$

where I_{DS} is instead a product of Q_{VS}, the charge area density at the virtual source, and v_{inj}. To get a full description for all regions of operation it is possible to embed the voltage dependences into Q_{VS} and v_{inj}, similar to the model of the classical FET. At the virtual source, the carrier density is weakly dependent on the band bending induced by V_{DS} and can thus be neglected. Further, the V_{GS} dependence of v_{inj} can also be disregarded, albeit with a more complicated reasoning; at a high V_{GS}, any further lowering of the barrier will increase the carrier diffusion probability but it will also increase the k_BT-layer width such that the two effects can be considered to cancel each other out [Lundstrom 2002].

2.1.1 A Compact Virtual Source Model

Following the derivation proposed by Khakifirooz et al [2009], Q_{VS} can be written as;

$$Q_{VS} = C_G n\varphi_T \log\left(1 + \exp\left(\frac{V_{GS} - (V_T - \alpha\varphi_T F_f)}{n\varphi_T}\right)\right), \tag{2.6}$$

where the α-constant is a fitting parameter which relates to the shift in reference voltage of the threshold in the sub-threshold and above-threshold region. α is found empirically to be around 3.5 [2009]. The rest of the parameters and constant have the same meaning as in Chapter 1, except F_f, the Fermi function, which is here expanded compared to Eq. 1.17:

$$F_f = \frac{1}{1+\exp\left(\frac{V_{GS} - (V_T - \alpha\varphi_T/2)}{\alpha\varphi_T}\right)}. \tag{2.7}$$

The introduction of the Fermi function (ranging between 1 and 0) is a way to mathematically accomplish a smooth transition from sub-V_T to above-V_T operation.

The description of the v_{inj}, it can be modeled as;

$$v_{inj} = v_{avr} \frac{V_{Acc}}{V_{Sat}}, \tag{2.8}$$

where the two voltages have the same meaning as for the classical FET model albeit with modified descriptions. V_{Acc} is given by;

$$V_{Acc} = \frac{V_{DS}}{\left(1+\left(\frac{V_{DS}}{V_{Sat}}\right)^{\beta}\right)^{1/\beta}}, \quad (2.9)$$

where a β-constant is introduced as a fitting parameter. When β is equal to 1, Eq. 2.9 is identical to Eq. 1.21. For higher values of β, the current will saturate faster and for a β value going towards infinity, there will be a very sharp transition where $V_{Acc} = V_{Sat}$ for $V_{DS} = V_{Sat}$. Empirically, β is found to be around 1.8 for n-type channels [2009]. To complete the model, the description of V_{Sat} is given by:

$$V_{Sat} = \frac{v_{inj}L_G}{\mu_{eff}}\left(1 - F_f\right) + \varphi_T F_f. \quad (2.10)$$

2.1.2 Material and Structural Considerations

To fit the model to a transistor, it is important to look at the specifics of the actual structure and materials. A real transistor has for example channel access series resistances at both the source, R_S, and the drain, R_D. When fitting the analytical model in a network of parasitic elements, the voltages and currents must all add up and this can be accomplished with the use of an iterative method. Beyond series resistance, other additions and modifications should also be considered to successfully capture measured characteristics. Examples of what could be added to the transistor model to accurately describe an InAs NW-FET are given in the following segments.

2.1.2.1 Channel Capacitance

For a NW-FET, the geometric coaxial gate oxide capacitance, C_{coax}, is given by;

$$C_{coax} = 2\pi\epsilon_0\epsilon_{ox}L_G / \log\left(\frac{2t_{ox}+D_{NW}}{D_{NW}}\right), \quad (2.11)$$

where ε_0 is the permittivity of vacuum, ε_{ox} is the relative oxide permittivity, D_{NW} is the nanowire diameter, and t_{ox} is the oxide thickness. C_{coax} approaches the planar oxide capacitance, for which $C_{ox} = \varepsilon_0\varepsilon_{ox}/t_{ox}$ per unit area, for $D_{NW} > 100\cdot t_{ox}$. As described in chapter 1, the total gate capacitance, C_G, is the series coupling of C_{ox} and C_{semi} (see Fig. 1.7b), where C_{semi} is referred to as the quantum capacitance, being related to the density of states. In the simplistic case of a planar 2D-channel, under the assumption of a large gate bias, the channel is in strong accumulation/inversion (depending on the type of channel doping), and C_{semi} is much larger than C_{ox}, thus C_G equals C_{ox}. For InAs, however, the density of states is relatively low. Studies of C_G of InAs NW FETs report a discrepancy between the geometric and the measured on-state (V_G-$V_T = V_D$) capacitance of about a factor of 2 [19] [20] indicating that C_{semi} is about the same size C_{coax} in strong

accumulation/inversion. To accurately model C_G, a parameter, $\varsigma = C_{coax}/(C_{coax} + C_{semi})$, can be introduced to account for the effect of C_{semi} such that $C_G = \varsigma C_{coax}$.

2.1.2.2 Mobility

As discussed in chapter 1, mobility is material specific, and relates the drift velocity to the applied electric field. III-V materials can have very high bulk mobilities (μ_{bulk} > 10,000 cm^2/Vs) compared with other semiconductors due to its relatively low effective mass for electrons. The way a materia1 is integrated will, however, affect the mobility. Interface defects, defects in bordering materials, and channel surface roughness will result in an effective channel mobility, μ_{eff}. According to the Matthiessen's rule, μ_{eff} can be determined by a summation expression, where the mechanism with the lowest corresponding mobility is the most dominating:

$$\mu_{eff} = \left(\frac{1}{\mu_1} + \frac{1}{\mu_2} + \frac{1}{\mu_3} + \cdots\right)^{-1}. \qquad (2.13)$$

Fig. 2.4a shows an example of a profile of the effective mobility of InAs NW as a function of distance to the NW surface. The data is simulated using a simple model. If the scattering related to the surface is higher than in the bulk NW, the average mobility will decrease with the NW diameter. Fig. 2.4b shows the mean of the cross-sectional area integration of the profile in Fig. 2.4a. If the conduction at the surface and in the core of NW can be separated [21], it can be meaningful not

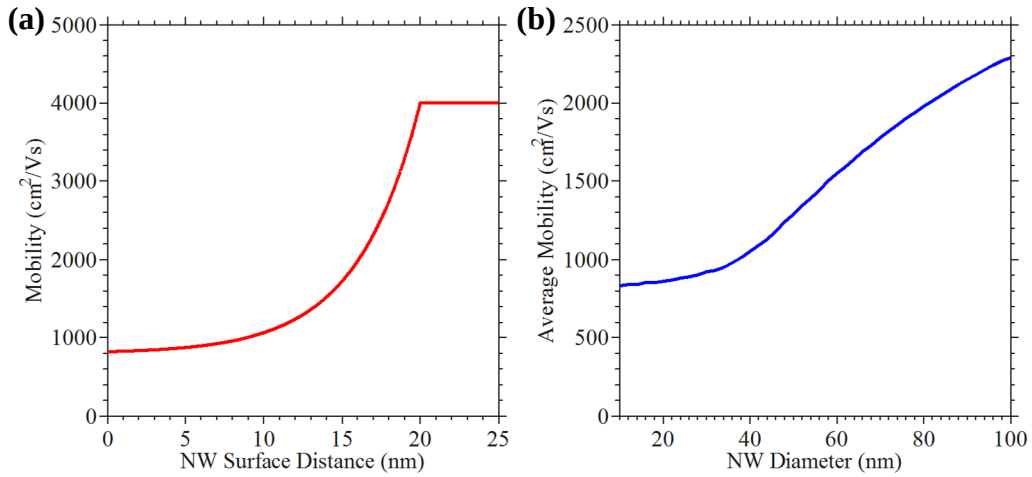

Fig. 2.4. (a) *Effective mobility versus the distance of the transport channel to the NW surface. The saturation at a large distance represents the bulk NW mobility. The data is constructed as an example of how the mobility could vary between the centre and the surface of a NW.* **(b)** *The average effective mobility as function of NW diameter, calculated by integrating the NW cross-sectional area and the mobility profile in a.*

only to have a measure of the average mobility, but also the mobility as a function of the distance to the surface. A buried channel device exploits the potential gain in having the channel separated from the oxide interface with an intermediate semiconductor layer, thus having a higher μ_{eff}.

2.1.2.3 Doping and Parasitic Conductance

An important device parameter is the doping concentration, where precise doping profiles are essential for good device performance. Ideally, doping in the access regions should be as high as possible to form good Ohmic contacts and minimize series resistance. For an n-type device, the channel region should either be p-doped or kept intrinsic, the later meaning as low doped as possible to maximize channel control. Less than ideal doping profiles may be used for reduced fabrication complexity, such as homogenously doped devices, where channel control is sacrificed for reduced series access resistance. In worst case, a highly doped channel might result in increased scattering or an inability to fully deplete due to charge screening. Considering a NW FET, screening could effectively result in a parasitic conduction path in the center of the NW, with little control by the surface potential. The maximum depletion width, W_{max}, which is the distance from the surface that can effectively be depleted of charge, is given by [10]:

$$W_{max} = 2\sqrt{\varphi_T \epsilon_0 \epsilon_s \log\left(\frac{N_D}{n_i}\right) / (qN_D)}. \qquad (2.14)$$

In Eq. 2.14, N_D is the donor doping concentration and n_i is the intrinsic carrier concentration. A parasitic channel can be modeled as a parallel resistance, R_{par}, and for a NW FET it is given by the channel length divided by the sheet area and multiplied with the resistivity:

$$R_{par} = \frac{4L_G}{\pi(D_{NW}-2W_{max})^2}\frac{1}{qN_D\mu_{bulk}}. \qquad (2.15)$$

In Eq. 2.15, μ_{bulk} is the bulk NW mobility, where the conduction is separated from the surface at a distance corresponding to the flat level mobility in Fig. 2.4a. A large parasitic conductance path is very detrimental for the transistor performance as it limits both the ability to turn a device off as well as lowers the maximum voltage gain. The voltage gain is proportional to g_m/g_d, where g_d is the output conductance, dI_{DS}/dV_{DS} (read more about g_m and g_d in chapter 3).

2.1.2.4 Sub-threshold Slope

A lot of the complexity associated with the development of high speed MOSFETs lies in improving the interface between the semiconductor channel surface and the applied dielectric film. Many different types of studies are conducted in an effort to

find the right surface preparation methods, deposition parameters and temperature treatments. High-*k* films, dielectrics with high electrical permittivity compared to traditional SiO_2 (used by the Si-CMOS-industry until recently as it could be formed as a native oxide), offer the possibility of thicker films not compromised by tunneling, yet with good electrostatics. High-*k* films are, however, also associated with higher concentrations of defects than the SiO_2. Energy traps, spatially distributed at the interface and throughout the dielectric films, screen the gate potential and thus affect the ability of steep device current turn-off. The channel control is especially compromised by traps at an energy depth corresponding to the region around the conduction band edge. The impact of traps can be modelled with the SS-parameter, *n*, as explained in chapter 1, and that can be taken directly from DC measurements. When fitting a device model, it is of importance to be able to distinguish between a parasitic conduction path and the effect of *n* in the sub-threshold region. For an accurate RF model, the frequency dependence of C_i needs to be characterized. The characterization of trap density is further explained in chapter 4.

2.1.2.5 Ballistic Transport

When applying the VS model, v_{inj} is one of the fitting parameters. A measure of ballisticity is when the effective injection velocity approaches that of Eq. 2.2, thus, when being equal, there occur no backscattering and R in Eq. 2.1 is close to zero. Although there occur no scattering in a ballistic channel, carrier acceleration is not infinite. As the transition time through the channel is relatively short ($< L_G/v_{avr}$), a carrier has limited time to acquire a drift velocity [22]. The current characteristics of this effect strongly resemble that of scattering and can in fact be modeled with a nonphysical ballistic pseudo mobility, μ_B. The effective, apparent, channel mobility can be calculated by summing up all the scattering and pseudo-scattering mechanisms using Eq. 2.13 [15]. For short L_G III-V devices, μ_B is considerably lower than μ_{bulk}, and thus μ_B is dominating the device characteristics. Using the derivation proposed by Wang et al. [15], for a unidirectional thermal injection velocity, μ_B can be expressed as:

$$\mu_B = qL_G/(\pi m^* v_{thermal}). \qquad (2.16)$$

For a 1D-channel, the conduction band cannot accurately be modeled as continuous, but is better described as split into distinct modes, sub-bands, where effect is increasing with decreasing channel width. In accordance with Landauer-Büttiker formalism [23], for ballistic transport, each mode has the fundamental quantized conduction of $G_q = 2q^2/h$, which corresponds to a fundamental quantized contact resistance, $R_q \approx 12.9$ kΩ. The limit in conduction originates from a finite transmission probability, entering and exiting the sub-bands due to reduction in available states going from the contact to the channel [23]. When modeling a low-dimensional ballistic transistor, the number of accessible modes (significantly close

to the bulk band edge), can be estimated from the on-resistance, R_{on} (more about R_{on} in chapter 3). In this case, R_{on} will consist of the quantum contact resistance in series with the channel resistance due to the ballistic pseudo mobility. Assuming that the channel resistance at low V_{DS} is relatively small, the number of available modes, M, can be determined using $R_{on} \sim 12.9/M$ kΩ. Knowing M, the total quantized contact resistance can be added in the transistor model as intrinsic series resistances located at both the source, $R_{S,i}$, and the drain, $R_{D,i}$. The size of $R_{S,i}$ and $R_{D,i}$ each corresponds to half the total quantized contact resistance, or approximately 12.9/(2·M) kΩ. To get a good modelling accuracy of I_{DS} in operation slightly above V_T, however, the sub-band energy separation should also be

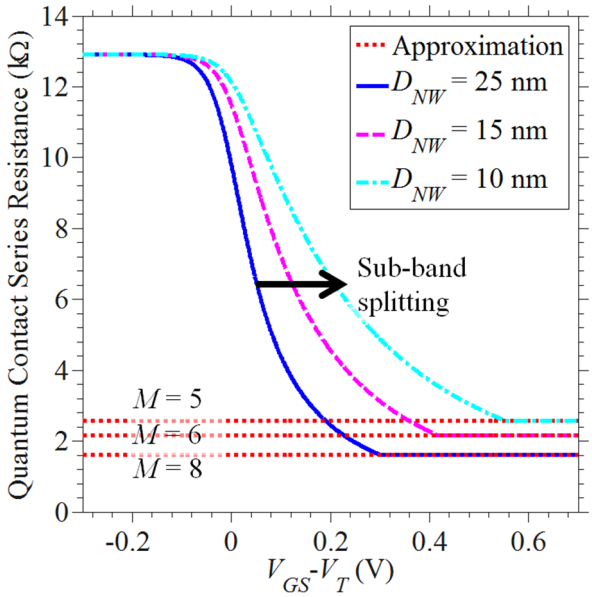

Fig. 2.5. *The plot is showing a quadratic approximation of the quantum contact resistance. The depicted function models individual threshold voltages for each sub-band, estimated from reference* [24]*, showing the V_{GS}-dependence of the series resistance. As an approximation, the quantum contact resistance can be set at a constant value; this will, however, give a poor fitting for the region slightly above threshold, and this region is expanding to higher V_{GS} with decreasing nanowire diameter. At a relatively large $V_{GS,}$ the number of conducting sub-bands, M, correspond to 5, 6, and 8 for a D_{NW} of 10, 15, and 25 nm, respectively.*

included. Each sub-band can be seen as a parallel channel, where each channel has a different associated threshold voltage. This behavior can be described by introducing a gate potential dependence of M in a confined region above V_T, where M goes from 1 to the value of determined from R_{on}. In the sub-V_T region, I_{DS} has little dependence on the series resistance as V_{DS} is replaced with φ_T. A quadratic model for the band separation for ballistic NW-FETs with D_{NW} of 25, 15, and 10 nm are shown in Fig. 2.5. It can be seen that for the NW-FET with D_{NW} = 25 nm, having V_{GS}-dependence on M is necessary for at least the first 100 mV above V_T.

For the thinner channels with larger sub-band separations, a variable *M* is important over a larger bias range. Beyond the implementation of a V_T-function through a variable *M*, a function accounting for the increasing effective mass associated with the increasing energy level at the edge of each sub-band should also be included [24]. Taking the average effective mass for all of the sub-bands, however, works for transistors where the band separation if relatively small.

2.2 Studies of VS Model Application

To show validity of the presented compact VS model and the material and structural considerations, data fittings are presented for measured characteristics of fabricated semi-ballistic NW-FETs as well as electrical characteristics from atomistic simulations of ballistic NW-FETs.

2.2.1 Fabricated InAs NW FETs

Fig. 2.6a, b, and c show data fittings of output characteristics of fabricated NW-FETs consisting of single NWs. The devices are fitted for different parasitic series resistances but with the same injection velocity (v_{inj} = 1.7·10^7 cm/s) and mobility (μ_{eff} = 1300 cm^2/V·s). A more though investigation of the intrinsic components and

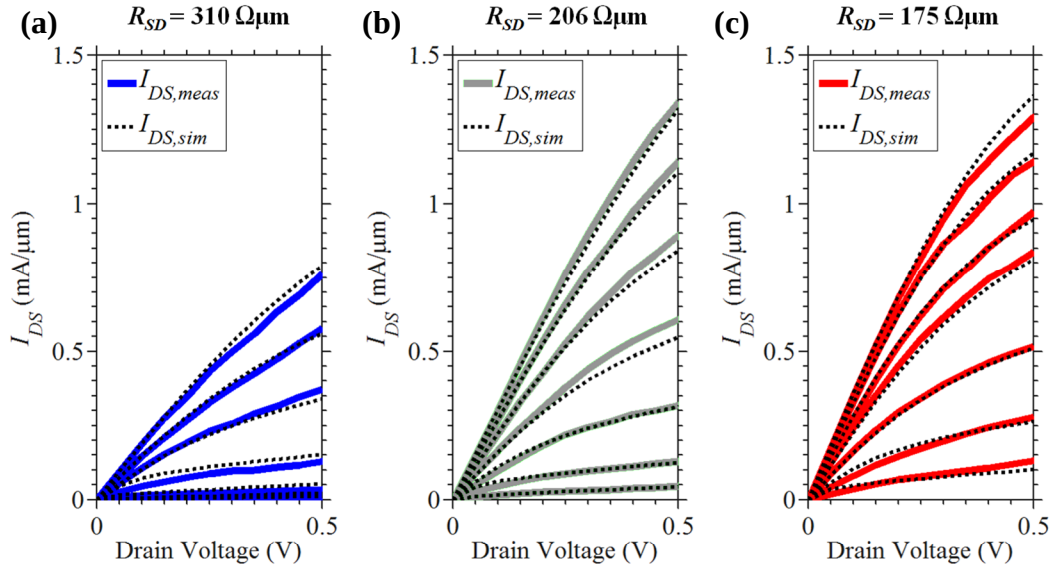

Fig. 2.6. *Measurements of single nanowire MOSFETs with D_{NW} = 45 and L_G = 200 nm, $I_{DS,meas}$, fitted with a virtual source model, $I_{DS,sim}$. All devices are fitted with an injection velocity of 1.7·10^5 cm/s and an effective mobility of 1300 cm^2/V·s. There is possibly a variation in doping incorporation, nanowire diameter, vertical segment lengths, or a combination, which cause a difference in determined parasitic source and drain series resistance, R_{SD}.* **(a)** *NW-FET with R_{SD} = 310 Ωμm.* **(b)** *NW-FET with R_{SD} = 206 Ωμm.* **(c)** *NW-FET with R_{SD} = 175 Ωμm.*

the impact of series resistance is explored in paper IV. Due to a high and homogenous doping profile and relatively large diameter, the devices show a large parallel parasitic conduction in the center of the wires. Modelling of parallel parasitic conduction is described in section 2.1.2.3. A study of the center conduction is also presented in paper VII in regards of the low frequency noise characteristics, also described in in chapter 4.

2.2.2 **Atomistic Modeled Ballistic InAs NW FET**

Fig. 2.7a, b, and c show data fittings of the output characteristics of ballistic nanowire transistors with 10, 15, and 25 nm diameter, respectively. The original data has been generated through atomistic tight binding simulations [24]. The simulations where performed with $\mu_{bulk} >> \mu_B$ and no access resistances. For InAs NW FETs, sub-band splitting is prominent for D_{NW} < 20 nm [24]. To model the behavior of the different channel widths, the number of available sub-bands was determined at low V_{DS} accrrding to the method described in section 2.1.2.5. For D_{NW} = 15 nm, the resistance corresponds to 6 accessible modes, meaning that the finite sub-band conduction introduces a series resistance of R_{series} = 12.9 kΩ/6 = 2.2 kΩ giving $R_S = R_D = R_{series}/2$ = 1100 Ω. From the fitting, a mobility of 1200 cm^2/Vs was deduced and this agrees well with calculations of the ballistic pseudo mobility. In Fig. 2.7a, showing the fitting for a D_{NW} = 10 nm NW-FET, it can be seen that it is less accurate than for the 15 and 25 nm NW-FETs. The explanation in part lies in

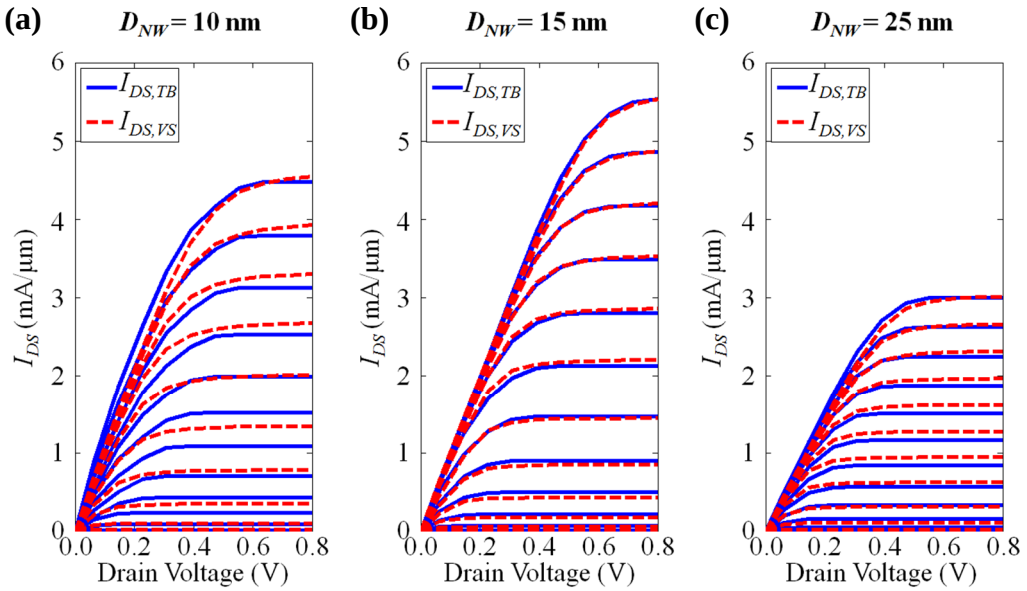

Fig. 2.7. *Tight binding simulations of ballistic NW-FETs, $I_{DS,TB}$* [24]*, fitted with a virtual source model, $I_{DS,VS}$, using the quantum series resistance function shown in Fig. 2.5. For D_{NW} = 10 nm, the fitting is worse than for the larger diameters, partly due to a large sub-band splitting where an average effective electron mass is no longer sufficient.* **(a)** *D_{NW} = 10 nm, L_G = 14 nm.* **(b)** *D_{NW} = 15 nm, L_G = 20 nm.* **(c)** *D_{NW} = 25 nm, L_G = 34 nm.*

that the quadratic quantum contact series resistance function does not capture the conduction increase in quantized steps, which is the case for large sub-band separation. Also, the quantum capacitance can in this case no longer be modelled by a fixed value at high accumulation [17], as it increases in steps in accordance with the quantum conductance. Using an average effective electron mass also becomes less accurate, impacting the injection velocity.

2.2.3 **Data Fitting Evaluation**

It can be concluded that a compact model can capture the behavior of semi-ballistic and ballistic NW-FETs, however, for channel widths of 10 nm or below, the accuracy may not be sufficient. In Fig. 2.8, the determined injection velocities are plotted against L_G. There are two observable trends. The first is that injection velocity and normalized conduction increases when the channel width and gate length is decreased from 25 to 15 nm. The other is the opposite trend going from 15 to 10 nm. As the channel width is decreased, the average effective mass of injected electrons increases and thus lowers the average injection velocity (Eq. 2.2). A deeper analysis of the change in transport with nanowire diameter is given in reference [24]. The change of v_{inj} with L_G is discussed in the next section, also benchmarking the determined values to other technologies.

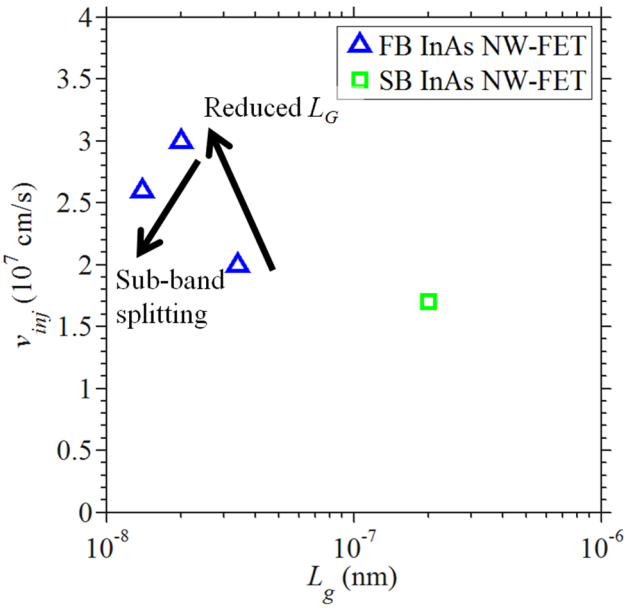

Fig. 2.8. *Injection velocity, v_{inj}, as a function L_G. For the full-ballistic (FB) transistors, there are two observable trends. As L_G reduces, v_{inj} increases from the thermal velocity towards the Fermi velocity. However, as the nanowire diameter decreases, sub-band splitting increases, resulting in both an increase in average effective electron mass and v_{inj} again going towards the thermal velocity. As a reference, the fitted value for the measured semi-ballistic (SB) transistors is added to the graph.*

2.3 **Benchmarking**

To contrast the determined values from the compact VS model, different values for III-V high electron mobility transistors (HEMTs) [25] as well as Si-CMOS (Intel, 45 nm node) [18] are given in table 2.1. In the table, T_{ch} indicates the thickness of the channel and is given for the HEMTs while DNW is. The non- confined channel widths of the HEMT structures do in part explain the higher injection velocities obtained compared to the values determined for the NW-FETs. As discussed in section 2.2.3, for an increasing average effective electron mass, the thermal velocity is reduced. One reason for an increase in injection velocity with reduced L_G is the associated increase in DIBL, where carriers with lower effective mass can be injected as the channel barrier is lowered. This is also a contributing factor for the higher injection velocities achieved in the HEMT structures as compared to the NW-FETs, which have significantly better electrostatics [2].

Comparing the injection velocities determined for the III-V devices with the values determined for Si-CMOS, at a comparable L_G, v_{inj} is about a factor of 2-3 higher at a V_{DD} that is 1/2.

TABLE 2.1. BENCHMARK OF INJECTION VELOCITIES FOR DIFFERENT TECHNOLOGIES

Material	Technology	V_{DD} (V)	D_{NW} / T_{ch} (nm)	L_G (nm)	v_{inj} (10^7 cm/s)
InAs [25]	HEMT	0.5	5	30	3.7
$In_{0.7}Ga_{0.3}As$ [25]	HEMT	0.5	5	30	3.2
$In_{0.53}Ga_{0.47}As$ [25]	HEMT	0.5	5	30	2.8
Strained-Si [18]	nMOS	1.0	-	30	1.35
Strained-Si [18]	pMOS	1.0	-	30	1.03
InAs [21] Paper IV	NW-FET	0.5	45	200	1.7
InAs	NW-FET	0.5	25	34	2.0
InAs	NW-FET	0.5	15	22	3.0
InAs	NW-FET	0.5	10	14	2.6

CHAPTER 3

3 High-speed Technologies

But anybody who steps into the lane beside you is the biggest competition because they made it to the finals.
- Usain Bolt

The international technology roadmap for semiconductors (ITRS) is an industry organization that defines future targets for integrated electronics to coherently guideline the development [26]. One of the specified targets for high performance RF front-end technologies (ITRS 2012) is the transistor electron injection velocity, which should be at a level of $v_{inj} = 3 \cdot 10^7$ cm/s in the year 2018. This requirement hints that a transition to III-V based transistors in this particular segment is needed, which is in line with the progress in the field, with technologies reaching maturity.

The advantage of III-V materials concerning electron transport in MOSFETs was described in chapter 2. This chapter will give a brief overview and benchmark of current high-end III-V technologies.

3.1 Transistor Metrics

In order to make relevant comparisons between different technologies, the most important performance metrics are considered. Due to the intrinsic transport properties of III-Vs, less drive voltage is needed compared to Si-based devices, and III-V technology is commonly benchmarked at a V_{DS} of 0.5 V. Also, conductance metrics, series resistances, and capacitive loads, are often normalized with respect to the channel width. For a nanowire, the width is set as the circumference.

At least 7 major transistor metrics can be determined From DC- and RF-measurements and these are presented for both bipolar transistors (BJTs, see section 3.2) and field-effect transistors (FETs, see section 3.3-3.5) in the following. For simplicity, n-type devices are assumed if not stated otherwise.

Transconductance - g_m

The transconductance, g_m, is a measure of the change in device current with a change in gate or base potential. g_m is generally proportional to the device mobility, however, for ballistic and semi-ballistic devices, it is instead highly dependent on

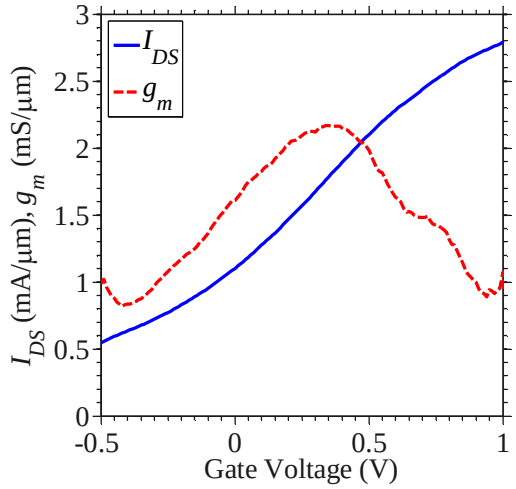

Fig. 3.1. *Transfer characteristics, normalized to the NW circumference D_{NW} = 45 nm, measured for an InAs NW-FET consisting of a single NW. The data is measured for V_{DS} = 1.5 V, corresponding to an estimated 6 V/μm average field strength at maximum g_m.*

the injection velocity. Measured data of I_{DS} and the determined g_m for a NW-FET is shown in Fig. 3.1.

Output conductance - g_d

The output conductance, g_d, a measure of the change in current with the drain or collector potential in saturation, where it is desired to have as low value as possible. Short channel effects, high-field effects, and parallel parasitic conduction paths lead to higher g_d. The amplification of a transistor is limited by the intrinsic voltage gain, defined as g_m/g_d.

Access Resistance – R_{access}

The parasitic resistance in the access regions to the channel, R_{access}, is electrically in series with the channel and is detrimental to the device performance as it reduces the channel field strength and consumes power. Without separating the parasitic resistance and the channel resistance, values can instead be given in the on-resistance, R_{ON}, which is a measure of the total device resistance in either strong accumulation (FET) or saturation mode (BJT). For a FET device, R_{access} is the sum of the parasitic resistance at the source contact, R_S, and at the drain contact, R_D. For a BJT device, R_{access} is instead the sum of the resistance at the emitter contact, R_E, and at the collector contact, R_C. Two other important parasitic resistances are the gate contact resistance associated with the FET, R_G, and the base contact resistance associated with the BJT, R_B.

Sub-threshold Swing - SS

The sub-threshold swing, SS, is a measure of the amount of voltage needed to change the device current in the sub-threshold region by one decade. The minimum SS is set by the thermal voltage and corresponds to 60 mV/decade at room temperature. Most fabricated devices are, however, measurably affected by electron screening, for example due to the presence of interface states, which degrades the SS performance. The SS is important as it sets the limit, within a given voltage envelope, of how much current is consumed in the transistor off-state, or reversibly, how high above threshold it is possible to operate in the on-state for a fixed off-state-current.

Load Capacitance – C_L

The load capacitance, C_L, is the total capacitance that loads a transistor. By knowing C_L and g_m, a crude estimate of the highest operational frequency can be calculated as $g_m/(2\pi \cdot C_L)$. Minimization of C_L is thus critical to reach high-frequencies. Usually, there is a distinction between the intrinsic and the extrinsic capacitances, where the former is due to accumulation of charge in the channel. The extrinsic capacitance, also referred to as the parasitic capacitance, is instead attributed to parasitic charge accumulation between the different electrodes. For ultra-scaled devices, the extrinsic capacitances may constitute the larger part of C_L. For the FET, the intrinsic parts of C_L roughly correspond to the gate capacitance, C_G, which is the capacitance between the gate node and the source and drain nodes, C_{GS} and C_{GD}, respectively. For the BJT, the intrinsic parts of C_L roughly correspond to capacitance between the junction and the collector, C_{JC}, and the junction and the emitter, C_{JE}.

Cut-off Frequency – f_T

A transducer is an amplifier that can transform for example a voltage to a current. The frequency at which the transducer gain equals unity is called the cut-off frequency, f_T. For an FET, f_T can be estimated as the following:

$$f_T = \frac{1}{2\pi}\left(\frac{C_G}{g_m} + C_{GD}(R_S + R_D)\right)^{-1}. \qquad (3.1)$$

The expression for estimating f_T for a BJT will not be given here as it requires a deeper analysis of the different intrinsic components.

Maximum Oscillation Frequency - f_{max}

A unilateral amplifier has the same type of signal at both the input and the output. The maximum oscillation frequency, f_{max}, is a measure of when the unilateral power gain reaches unity. For a FET, f_{max} can be estimated by the following expression:

$$f_{max} = \frac{1}{2}\left(\frac{f_T}{2\pi C_{GD}(R_S+R_G)+g_d(R_S+R_G)/f_T}\right)^{1/2}. \quad (3.2)$$

For a BJT, f_{max} is instead estimated by:

$$f_{max} = \frac{1}{2}\left(\frac{f_T}{2\pi R_B C_{JC}}\right)^{1/2}. \quad (3.3)$$

f_T and f_{max} are important as they indicate the maximum bandwidth. The two metrics can be determined from measurements of two-port scattering parameters. By measuring the scattering parameters and using similar expressions to Eq. 3.1-3.3, it also possible extract the intrinsic components, such as g_m, which are otherwise determined from DC characteristics. By comparing the values acquired from RF and DC measurements, information about the impact of different physical effects and their corresponding time constants can be obtained [27].

3.2 Technology Overview

The following sections will give a brief overview of a set of different high speed technologies, explaining the different features and to some extent listing the specific advantages and drawbacks.

3.2.1 The Heterojunction Bipolar Transistor (HBT)

A diode can be fabricated by having two adjacent regions of a semiconductor doped with opposite charge. The bipolar transistor (BJT) can be understood as a series connection of two diodes with either the anodes (p-doped) or the cathodes (n-doped) in a common node, the base (B), depending on if it is an n-type (n-p-n) or a p-type (p-n-p) transistor, respectively. Bipolar refers to the fact that the transistor makes use of both hole and electron diffusion.

In contrast to the FET, the mechanism for current control for a BJT comes from the injection of carriers at the base, where a change in charge concentration, in an exponential relation to the B-E voltage, V_{BE}, changes the relative band position of the base. For an n-p-n transistor, the current between the collector (C) and the emitter (E), I_{CE}, corresponds to the base current, I_B, multiplied by a factor β. The multiplication comes from a difference in doping concentration between B and E, where E is higher doped.

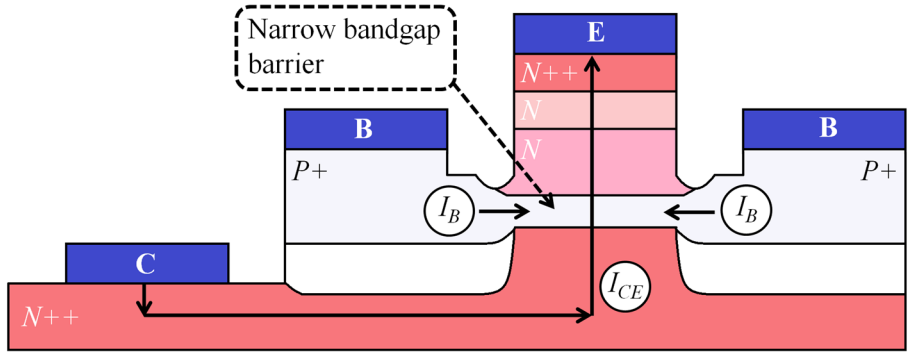

Fig. 3.2. *A schematic illustration of a HBT structure indicating the current paths as well as the narrow bandgap barrier situated in the base region.*

A heterojunction BJT (HBT) has an enhanced multiplication from a difference in barrier height between the conduction band edge and valence band edge [28]. For an n-p-n device, the holes in the valance band which are injected at B, face a higher barrier going towards E than the electrons at E going into B. A HBT is fabricated with a hetero-interface, where materials with different bandgaps are epitaxially grown on top of each other. To avoid the formation of a discontinuity in the conduction band at the E-B-junction, a graded material transition can be implemented, where there is a gradual change of the material composition going from E to B. One of the advantages of an HBT over a BJT is that the difference in barrier height in the case of the former can be used to increase the doping in the base, thus reducing the base contact resistance.

The carrier transport through the base relies on diffusion, where a high concentration of electrons injected at the E-B-junction diffuse towards the B-C-junction, where the electron concentration is lower. For this reason it is important that the thickness of the base is within λ_{mp}. A way to enhance the carrier transport though the base is to implement a graded base doping profile with a built in slope in the conduction band. This will increase the efficiency by drift assisted transport. Another device enhancement is the implementation of a two heterojunctions, called double-HBT (DHBT), improving the breakdown voltage.

Electrons injected from into the base, diffuse over the barrier and then drift towards C, accelerated by the potential difference between E and C, V_{CE}. A schematic illustration of a HBT is shown in Fig. 3.2. The output characteristics of a BJT are reminiscent of a FET and circuits designed for a similar function do not differ to a large extent between the two architectures. One of the advantages of the HBT compared to the FET is that the critical feature sizes are made with epitaxy rather than lithography [28]. As current is conducted vertically, through the stacked layers, it is the area of the emitter that restricts the current flow, where W_E corresponds to the cross-sectional emitter width illustrated in Fig. 3.2. For an FET, conduction capacity increases with the channel width it but does not scale with the channel length (although the drive voltage scales). The exponential dependence on

I_{CE} on the base-emitter voltage, V_{BE}, also means that the transconductance is substantially larger than for FETs, which have linear or quadratic current dependence. Furthermore, in contrast to the surface conduction associated with MOSFETs, HBTs conduct mostly in the bulk of the material and are thus less affected by surface effects. However, high-frequency noise tends to be worse for HBTs than for FETs, where the noise level, attributed to shot noise, is proportional to the current (see chapter 4). Also, an increase of R_B, in an effort of reducing the area footprint, has a larger detrimental impact on the switching time constant than an increase in R_G when scaling a FET device.

3.2.2 The High-Electron Mobility Transistor (HEMT)

The high-electron mobility transistor (HEMT) is similar to a MOSFET but differ in the aspect that instead of a dielectric film separating the gate and the channel, a semiconductor with a wide bandgap is used as a separation layer. The advantage is that the epitaxial material interface to the channel can be made with a very low density of defects, thus minimizing surface scattering. Further, to avoid scattering related to doping incorporation, referred to as impurity scattering, the channel is grown intrinsic and doping is instead introduced in the separation layer. Thus, the gate potential in a HEMT device only affects the channel charge concentration indirectly, where a higher gate potential increases the charge concentration in the separation layer and more carriers diffuse [28]. A high enough gate potential will, however, also form a layer of conductive charge in the separation layer, which will then deteriorate the performance. To avoid conductance in the separation layer, doping is low enough such that the chemical potential at room temperature is kept significantly below the conduction band edge. An abrupt junction between the separation layer and the channel results in the formation of a barrier at the channel interface and a quantum well (QW) in the channel. For a narrow bandgap III-V channel material and a steep band-bending in the QW, the close to intrinsic chemical potential will be above the conduction band edge and carriers in the

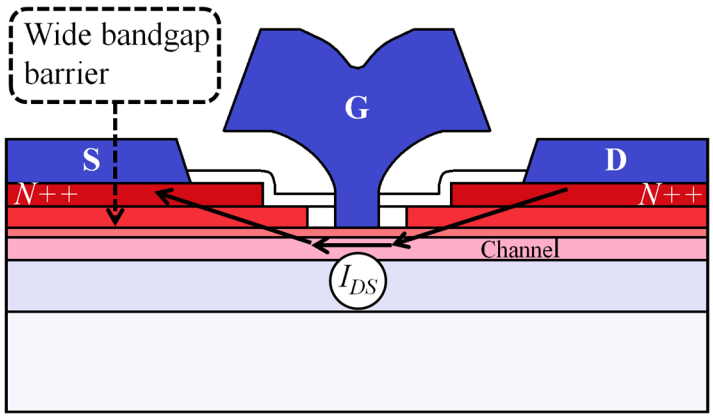

Fig. 3.3. *A schematic illustration of a HEMT structure, indicating the current path as well as the wide bandgap barrier material situated above the channel.*

separation layer can tunnel and diffuse to available states in the channel. The thickness of the QW is commonly ranging between 5-10 nm. A schematic illustration of a HEMT is shown in Fig. 3.3.

There are different categories of HEMTs. A pseudomorphic HEMT (pHEMT) implements very thin layers of different materials, significantly reducing crystal defects and trap formation, which are found in much higher densities for thicker layers of merged materials with large lattice mismatch [28]. Another way to overcome the problems associated with large material lattice mismatch is to have an intermediate buffer layer. A buffer layer can have a graded content of a certain element such that the top and bottom of the buffer matches both the material over and under it. This type of HEMT is called metamorphic HEMT (mHEMT). Common HEMT channel materials are InGaAs, GaAs, InP, GaN, and InAs.

The advantage of using a HEMT architecture is that very high mobilities can be achieved while the possibly low effective mass and high concentration of carriers in the QW ensure high injection velocities. One drawback of HEMTs is that the height of the barrier between the gate metal and the separation layer is relatively low compared to a MOSFET, resulting in problems concerning gate leakage [29]. Also, the design has a large foot-print and has possibly reached the end of channel length scaling where performance deteriorates for L_G shorter than about 30 nm due to insufficient gate control. In this aspect, the MOSFET architecture is more attractive as it offers the possibility to make 3D-channels and surrounding gates.

3.2.3 The Metal-Oxide-Semiconductor Field Effect Transistor (MOSFET)

The physical structure and operation of MOSFETs is described to some extent in chapter 1 and will therefore be left out in this section. The following text will instead focus on enhancement methods as of late.

The development of high performance III-V MOSFETs have long been held

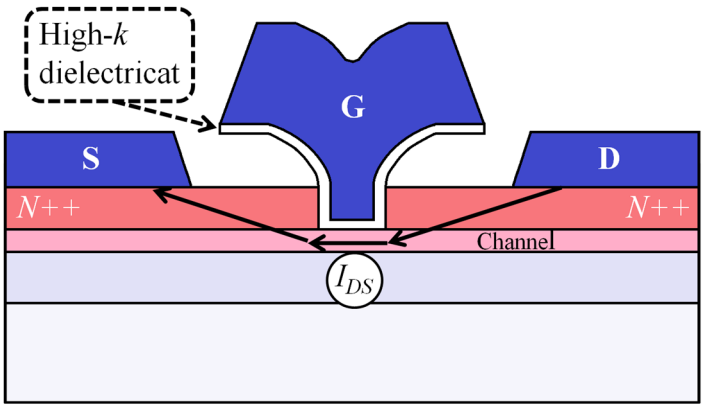

Fig. 3.4. *A schematic illustration of a MOSFET structure indicating the current path as well as the high-k dielectric barrier situated above the channel.*

back by the lack of good integration methods of dielectric films. Recent progress of high-*k* integration have, however, changed the situation and recent reports of III-V MOSFETs show normalized conductance values [30] [31] comparable to state-of-the art HEMTs [32]. A schematic illustration of a high-performance MOSFET is shown in Fig. 3.4.

The strength of the MOSFET architecture comes from having an isolation barrier, where relatively very low leakage currents can be achieved [29]. By adding the gate stack (the gate metal and the dielectric film) in one of the final fabrication procedures, the channel contacts can be optimized to reduce parasitic series resistance. The MOSFET is also associated with a relative freedom in forming the shape of the channel where there are various ways to fabricate a 3D-channel. Some methods involve pre-patterning and epitaxial growth and others make use of post-patterning and etch-back procedures; both directions have showed promising results. Since 2013, commercial available Si-FinFETs are found in Intels 22 nm node, demonstrating a maturity in high-density integration of non-planar FET technologies. As the electrostatics improve substantially from a surrounding gate [2], MOSFET architectures are predicted to be able to be scaled down to very small feature sizes. The possibility of adding capping layers in between the channel and the dielectric film is a way to reduce the impact of surface states. A QW-FET is using similar material stacks as HEMTs to benefit from both the transport advantage of a HEMT and the isolation advantage of a FET [33].

3.2.4 **The InAs Nanowire MOSFET (NW-FET)**

InAs HEMTs have been demonstrated with very high injection velocities and have shown very good performance metrics concerning both normalized conductance

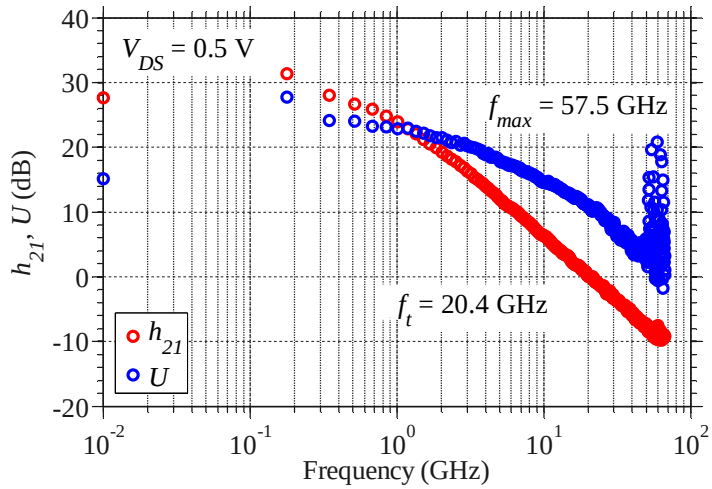

Fig. 3.5. *Unilateral gain, U, and transducer gain, h_{21}, for a multi-nanowire FET with 120 NWs, D_{NW} = 35 nm and L_G = 200 nm. The data was obtained for V_{DS} = 0.5 V. The data has been de-embedded according to the procedures described in paper IV.*

and frequency performance, thus demonstrating InAs as a good channel material. The possibility to make 3D channels make nanowires a good candidate for prolonged scaling towards very small dimensions [34].

The nanowire offer the possibility of enhanced electrostatics relative to both planar and double gate, and competitive conduction metrics for III-V NW-FETs have been reported [35] [36] [37] [38] [39]. Simulation studies of InAs NWs show intrinsic RF performance of several THz [24] and accounting for the parasitic elements in optimized realistic layouts, C_L is predicted to be below the IRTS target for corresponding nodes [40]. In Fig. 3.5 is plotted data for both the unilateral gain and the transducer gain as a function of frequency. The best multi-nanowire FET show f_T and f_{max} of 20 GHz and 57 GHz, respectively, at V_{DS} = 0.5 V. At the same drive voltage, a g_m of 0.2 mS/μm was determined. This is an improvement of previously reported data for InAs NW-FETs presented in paper IV [38], where f_{max} and g_m was reported at 30 GHz and 0.3 mS/μm, respectively, at V_{DS} = 1.0 V. As g_m in the measurements is similar for the two compared devices, the enhanced performance can instead be explained with an increase in the intrinsic voltage gain, from about 5 to 25, due to a reduction in parasitic parallel conduction related to reducing D_{NW} from 45 nm to 35 nm. Also, a reduction in size of the contact pads corresponds to a reduction of normalized extrinsic capacitances to about 2.0 fF/μm, down from 5.3 fF/μm. With a decrease in nanowire diameter, but an increase in the doping flow during the nanowire growth, the parasitic resistance has remained at about the same with R_{ON} ~ 3000 Ωμm.

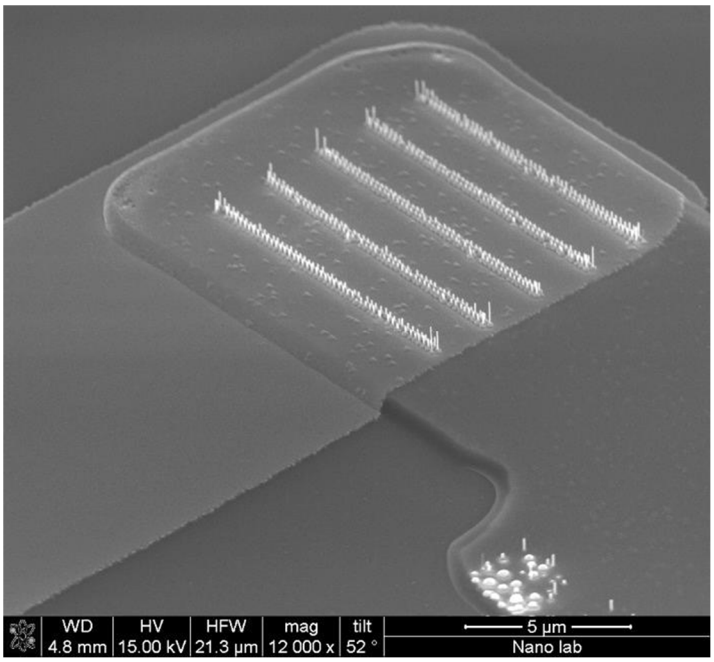

Fig. 3.6. *A multi-nanowire FET where the gate pad has been defined using UV-lithography. The visible lines are arrays of double rows of NWs.*

To enhance the performance it is important to further increase the doping incorporation while also reducing the channel width to maintain good electrostatics. It is also vital to further reduce the size of the overlapping pads [38], possibly by implementing stripes that contact the arrays separately, where a multi-nanowire FET consist of several arrays, as shown in Fig. 3.3. The data in Fig. 3.1 shows relatively high conductance values for a single-nanowire FET with g_m = 2.1 mS/μm at V_{DS} = 1.5 V. The same device show g_m = 1.37 mS/μm at V_{DS} = 0.5 V. The single-nanowire FET has been fitted with a VS model (Fig. 2.6c), where R_{access} was determined to 175 Ωμm. The low values of R_{access} was achieved with a high doping level, where there is possibly a difference in doping incorporation between single- and multi-nanowire FETs [38]. If the access resistance and the extrinsic plate capacitances can be substantially reduced for multi-nanowire FETs, it becomes necessary to also reduce L_G in order to reduce the drive voltage (and the intrinsic C_G). The current gate process is limited to a length about 200 nm. Replacing the sputtered gate process (which has shown good stability but has proven difficult to scale) to an evaporated gate process, successfully implemented in devices dating a few years back [41] [42], may be an alternative.

3.3 **Benchmarking**

In table 3.1, g_m and RF performance metrics are benchmarked for devices from each of the presented categories. As mentioned in the section about the HBT, it is difficult to make a relevant comparison between the conduction of a HBT and a FET. In the table, only g_m values for FETs are given. The data is sorted in falling order of *fT*. It is common to plot *fT* versus f_{max} as devices can be optimized for one or the other, however, it is especially difficult to achieve high values of both metrics. The highest performance regarding both *fT* and *fmax*, of the included papers, is achieved with HEMT architectures [32] [43]. The performance of the DHBTs is, however, not falling long behind [44] [45]. For MOSFETs, the RF performance is steadily improving [33] [46]. The nanowire-based devices are currently at an early stage in the development phase concerning RF performance. The advantageous scaling properties of non-planar MOSFETs are, however, promising for future implementations if NW-FETs can be implemented in optimized structures [40].

TABLE 3.1. BENCHMARK OF PERFORMANCE METRICS FOR DIFFERENT TECHNOLOGIES

Material	Technology	L_G / W_E (nm)	g_m *(mS/μm)*	f_T (GHz)	f_{max} (GHz)
$In_{0.7}Ga_{0.3}As$ [32]	HEMT	40	2.5[c]	688	800
InP [44]	DHBT	130	-	520	1100
InP [45]	DHBT	250	-	430	1030
InP [43]	HEMT	55	2.3[b]	385	1200
InAs [33]	QW-MOSFET	100	1.7	245	355
$In_{0.53}Ga_{0.47}As$ [46] Paper XII	MOSFET	55	1.5	244	292
InAs[a]	NW-MOSFET	200	0.2	20	57
InAs [30]	QW-MOSFET	70	2.7	-	-
InAs [31]	Tri-gate MOSFET	20	2.1	-	-
InAs [39]	NW-MOSFET	20	1.7	-	-
InAs [38] Paper IV	NW-MOSFET	200	1.4	-	-

a) Reported in section 3.2.4.
b) V_{DS} = 1 V
c) Estimated

CHAPTER 4

4 Noise

The worst wheel of the cart makes the most noise.
- Benjamin Franklin

One of the most important parameters regarding the performance of electronic circuits is the amount of generated noise. Noise limits the smallest detection level as there is a lower limit of possible signal-to-noise-ratio (SNR) if a signal is to be accurately recognized. All electronic components that consume energy at a signal conversion, such as voltage to current, will dilute the signal with thermal noise proportionally to the conversion factor.

Thermal noise is referred to as white noise and is spread out over all frequencies. The amount of white noise in a circuit is thus dependent on the circuit bandwidth (BW) and noise power is thus often given in per Hz. In addition to the white noise, there is pink noise or low-frequency noise (LFN), which originates from different electron transport processes and that characteristically have a $1/f$-dependence. In a spectrum of the noise power density, the point at which the contribution from LFN is lower than the white noise is called the corner frequency, f_C. In RF applications, it is desirable that the lowest used frequency is higher than f_C, as the LFN otherwise can introduce excessive circuit noise due to frequency up-scaling when signals are converted. Scaling of transistor dimensions and the introduction of new materials is typically coupled with an increase in the level of generated LFN and impose future challenges to avoid compromising the circuit noise performance.

In order to optimize the noise performance in electronic devices, the different noise sources must be identified. The most common sources of noise:

- **Johnson-Nyquist** - thermal, white noise
- **Shot Noise** - transport limitation, white noise
- **Generation-Recombination (GR)** - trapping, $1/f$-noise
- **Random Telegraph Signal (RTS) Noise** - GR-noise, few traps
- **Mobility Fluctuations** - scattering, $1/f$-noise

Noise is characterized by measuring power, voltage, or current as a function of frequency. The result is expressed in spectral noise density (SND), referred to the

noise current, S_I (A^2/Hz), the noise voltage, S_V (V^2/Hz), or the noise power, S_P (W^2/Hz). For a resistor, S_I can be converted to S_V by a multiplication of the squared resistance. For a transistor, however, it is the equivalent input gate voltage SND, S_{Vg}, which is of interest, where S_{Vg} equals S_I a divided by the squared transconductance, $S_{Vg} = S_I/g_m{}^2$.

The noise sources, classified as either LFN or white noise, will be explained further in section 4.1 and 4.2.

4.1 Low-Frequency Noise

The origin of 1/*f*-noise in transistors has been a long-going debate. The two underlying processes that are involved in causing a 1/*f*-noise characteristic are number fluctuations, where the number of carriers in the channel vary, and mobility fluctuations, where the velocity of the carriers vary. The lack of a complete theoretical description for mobility fluctuations still makes it an area of uncertainty, however, it is widely accepted that both number fluctuations and mobility fluctuations cause LFN. In Fig. 4.1, simulated curves of number and mobility fluctuations together make up the total noise spectrum. The shape of the total noise curve illustrates how the different noise contributions tend to dominate in different bias regions, where number fluctuations have a relatively larger impact on the noise level around threshold, where g_m is at its peak. On the contrary, mobility fluctuations have a larger contribution to the total noise in the sub-threshold region and at biases fairly high above threshold. The data in the graph is plotted in units of normalized current noise spectral density, $S_I/I_{DS}{}^2$, a relative measure that is often used when comparing different technologies.

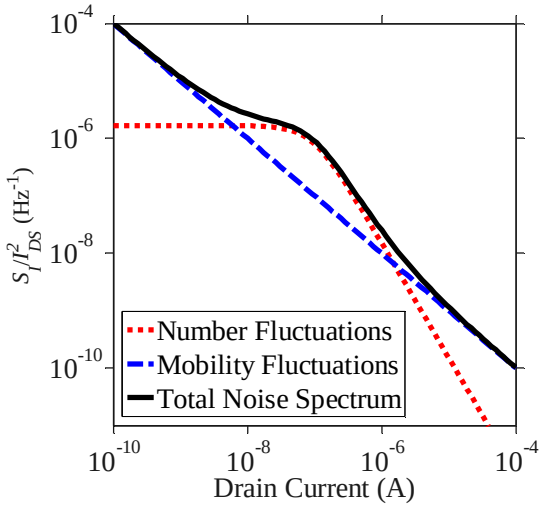

Fig. 4.1. *Normalized current noise spectral density, $S_I/I_{DS}{}^2$, plotted against drain current. The dashed and the dotted lines represent simulated number fluctuations and mobility fluctuations, respectively, while the solid line is the superposition of the two. For the plotted data, the frequency is set at constant f = 10 Hz.*

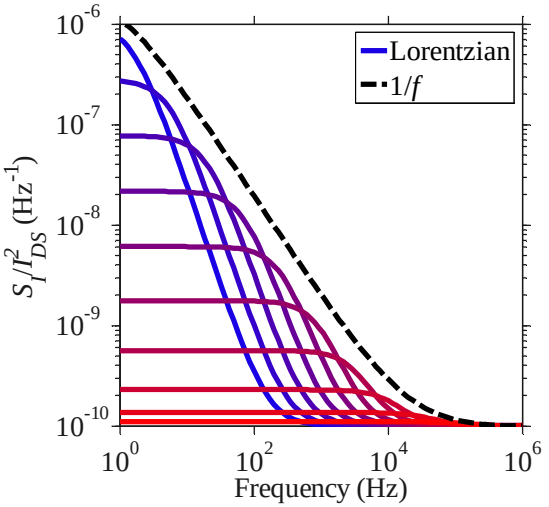

Fig. 4.2. *The superposition of evenly distributed energy border traps, where the total frequency response has the shape of a 1/f-curve.*

4.1.1 **Number Fluctuations**

Number fluctuations are mostly associated with MOSFETs, where border energy traps in the dielectric film causes a variation of carriers in the channel over time. For a trap at a certain spatial depth, the tunneling time constant, τ, is exponentially dependent on the tunneling distance, z:

$$z = \lambda_a \cdot \log_n\left(\frac{1}{2\pi f \tau_0}\right). \quad (4.1)$$

In Eq. 4.1, λ_a is the tunneling attenuation length and τ_0 is a time constant, usually set as 100 ps [47]. λ_a can be calculated as

$$\lambda_a = \left(\frac{4\pi}{h}\sqrt{2m^*\Phi_B}\right)^{-1}, \quad (4.2)$$

where Φ_B is the channel-to-barrier energy height. For a Si-SiO_2 interface, λ_a is about 1 Å [47]. For an InAs-HfO_2 interface, assuming that m^* = 0.14 [48] and Φ_B = 2.3 eV [49], λ_a can calculated to about 1.5 Å. The frequency noise response from a single trap situated at a certain depth will have a Lorentzian-shape. If there are few carriers in a channel with traps at a specific spatial depth, a measurement in the time-domain can resemble that of a telegraph signal, with the current jumping between two distinct levels, hence it is referred to as random telegraph signal (RTS) noise.

If there is an even distribution of the trap density regarding the dielectric film depth and there are, relatively, many channel carriers, the frequency trap response

can look like the simulated 1/*f*-noise curve in Fig. 4.2. The superposition of Lorentzian-curves together form a 1/f^{γ}-shaped total noise curve, where γ is the frequency exponent which typically ranges between 0.7-1.3 [47].

To model noise originating from number fluctuations, the following expression can be used;

$$S_I = \left(\frac{q^2 k_B T \lambda N_t}{f^{\gamma} W L_G {C_{ox}}^2}\right) {g_m}^2, \tag{4.3}$$

where N_t is the trap volume density, often given in the units of $cm^{-3}eV^{-1}$. By using $S_{Vg} = S_I/{g_m}^2$ and rewriting Eq. 4.3, N_t can be expressed as:

$$N_t = \frac{f^{\gamma} W L_G {C_{ox}}^2 S_{Vg}}{q^2 k_B T \lambda}. \tag{4.4}$$

4.1.2 **Mobility Fluctuations**

Mobility fluctuations are variations in mobility over time, related to different scattering mechanisms. According to the empirical Hooge model, the LFN originating from mobility fluctuations can be expressed as:

$$S_{ID} = \frac{q \alpha_H {I_{DS}}^2}{f^{\gamma} W L_G Q}. \tag{4.4}$$

In Eq. 4.4, Q is the channel charge density and α_H is the Hooge-parameter. The physical meaning of α_H has been under debate. The parameter has been found to be material specific and bias dependent, and it has been proposed that it can be modeled as a summation of specific constants associated with each scattering mechanism [47] [50]:

$$\alpha_H = (\mu_{eff}/\mu_{latt})^2 \alpha_{latt} + (\mu_{eff}/\mu_{imp})^2 \alpha_{imp} + (\mu_{eff}/\mu_{SR})^2 \alpha_{SR}. \tag{4.5}$$

In Eq. 4.5, the mobilities μ_{latt}, μ_{imp}, and μ_{SR}, and the Hooge-constants, α_{latt}, α_{imp}, and α_{SR}, are associated with the lattice, impurities, and surface roughness scattering, respectively. More scattering mechanisms could be added for higher accuracy. As different scattering mechanism dominate in different bias regions, α_H will also vary and different values may be determined for sub-V_T and above-V_T.

To model the noise characteristics of a device with both number fluctuations and mobility fluctuations, it is possible to combine Eq. 4.3 and 4.5:

$$\frac{S_I}{{I_{DS}}^2} = \left(\frac{q^2 k_B T \lambda N_t}{f^{\gamma} W L_G {C_{ox}}^2}\right)\left(1 + \frac{\alpha_H \mu_{eff} C_{ox} I_D}{g_m}\right)^2 \frac{{g_m}^2}{{I_{DS}}^2}. \tag{4.6}$$

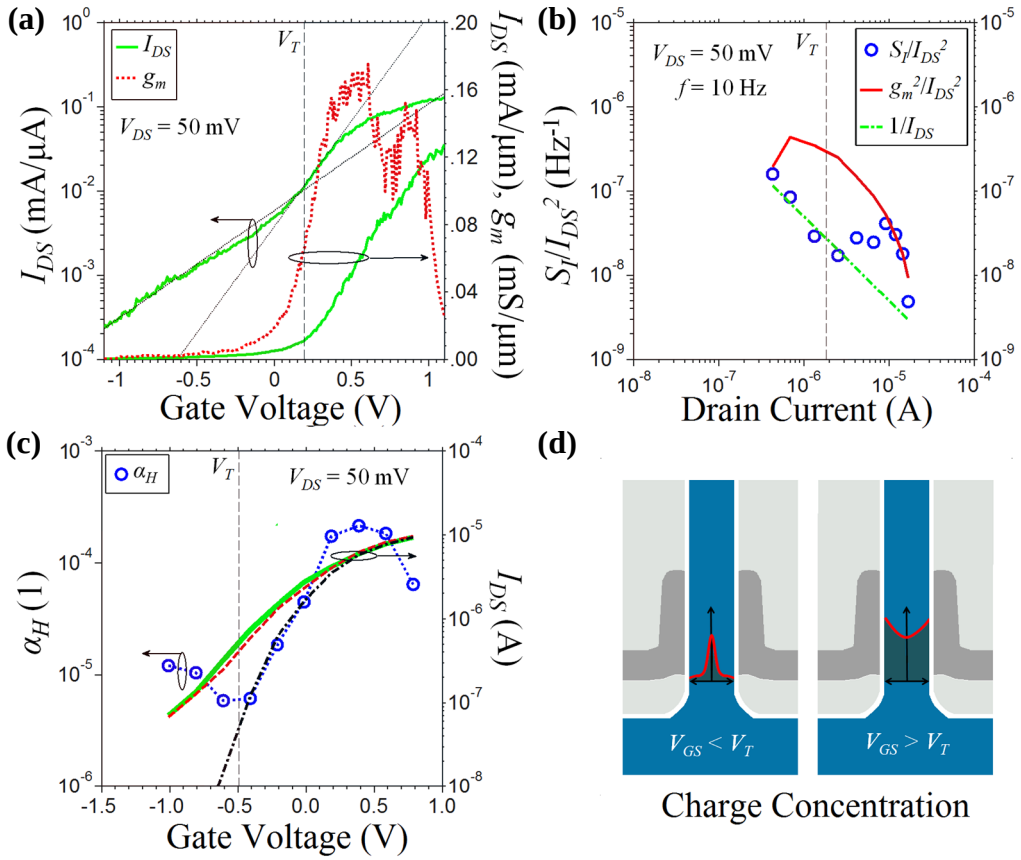

Fig. 4.3. (a) *The transfer characteristics plotted on both linear and logarithmic scale for a measured InAs NW-FET consisting of a single NW.* **(b)** *Normalized current noise spectral density plotted against the I_D for the same device as in (a).* **(c)** *Extractions of α_H-parameter plotted against V_{GS} for the same device as in (a). The green line is measured I_{DS} while the dashed red line and the black dashed-dotted line are simulations with and without core conduction, respectively.* **(d)** *Schematic spatial carrier concentration in a NW for sub-V_T (left) and above-V_T (right), illustrating how the carrier concentration in the center of a wide and highly doped NW channel can have a weak coupling to the gate potential.*

4.1.3 **Measurement of 1/*f*-noise in InAs nanowire MOSFETs**

Fig. 4.3a shows the transfer characteristics of a homogenously doped single NW-FET with D_{NW} = 45 nm. The output characteristics of the same transistor are shown in Fig. 2.6b. In Fig. 4.3a, there is a distinct shift in the slope of I_{DS} around V_T due to a leakage in the core of the NW, which is only weakly influenced by the gate potential [38]. A corresponding shift can be observed in Fig 4.3b, which is showing the measured S_I/I_{DS}^2 versus I_{DS}. Fitting g_m^2/I_{DS}^2 and $1/I_{DS}$ curves, which represent the trend for number fluctuations and mobility fluctuations, respectively, the

measured data seems to correlate with $1/I_{DS}$ in sub-V_T and g_m^2/I_{DS}^2 above-V_T. Fig. 4.3c shows the extracted α_H versus V_{GS} and the values that should be considered are those in sub-V_T since they seem to originate from mobility fluctuations. Fig. 4.3d shows a schematic spatial carrier concentration in a wide and highly doped NW-FET channel in sub-V_T and above-V_T. This data is in line with what was reported in reference [21].

4.1.4 **Technology Comparison**

Characterization of LFN is often made in order to improve certain fabrication aspects. However, the measurements also give metrics for technology comparison where there are industry targets specific for different technologies in order to comply with future integration requirements. In table 4.1, different non-planar technologies [51] [52] [53] are benchmarked in order to contrast the determined values obtained from paper V and VII. Values are given for the conventional $f = 10$ Hz. From an integration standpoint, the channel area normalized input referred gate voltage SND, $W \cdot L_G \cdot S_{Vg}$, is perhaps the most important metric for comparison. In the ITRS (2012) concerning multi-gate FET CMOS technology for high performance logic [26], it is stated that in 2015, devices should exhibit equal or better noise performance than 58 $\mu m^2 \mu V^2/Hz$ (at 1 Hz). The values for InAs NW-

TABLE 4.1. BENCHMARK OF NOISE PROPERTIES FOR DIFFERENT TECHNOLOGIES

Dielectric film	HfO_2 **[42]** Paper V	Al_2O_3 / HfO_2 **[21]** Paper VII	HfO_2 **[52]**	SiO_2 **[51]**	SiO_2 **[53]**
Technology	NW n-MOSFET	NW n-MOSFET	p-FINFET	NW n-MOSFET	NW n-MOSFET
Material	InAs	InAs	Strained SiGe	Si	InAs
W (nm)	$40 \cdot \pi$	$45 \cdot \pi$	$30 \cdot 10{,}000$	$10 \cdot \pi$	$30 \cdot \pi$
L_G (nm)	35	200	100	55-123	2000
EOT (nm)	1.5	1.8	1.3	3.5	-
N (cm^{-3})	$\sim 0.1 \cdot 10^{18}$	$\sim 1 \cdot 10^{18}$	-	-	-
μ_{eff} (cm^2/Vs)	-	1300/4000	200	-	500
I_{ON} (mA/mm)	104	670	-	-	-
R_{ON} (Ωmm)	3.7	0.33	-	-	-
g_m (mA/mm)	227	1190	-	-	-
SS (mV/decade)	110	500	-	-	-
S_{Id}/I_{DS}^2 (Hz^{-1})	$7.3 \cdot 10^{-7}$	$5 \cdot 10^{-9}$	$1.2 \cdot 10^{-8}$	$4 \cdot 10^{-8}$	-
N_t ($cm^{-3}eV^{-1}$)	$1.5 \cdot 10^{20}$	$6 \cdot 10^{19}$	$2 \cdot 10^{18}$	$1.7 \cdot 10^{18}$	-
$W \cdot L_G \cdot S_{Vg}$ ($\mu m^2 \mu V^2/Hz$)	5700	60	6	50	-
α_H	$5 \cdot 10^{-3}$	$5 \cdot 10^{-5}$	$\sim 10^{-5}$		$5 \cdot 10^{-4}$

FETs in paper VII [21] are about a factor 10-13 higher than the ITRS target, which is fairly close although further optimization of the interface between the nanowire surface and the dielectric film is still needed. High-*k* on Si have shown promising results with extracted N_t values down to $2 \cdot 10^{17}$ $cm^{-3}eV^{-1}$, more than two orders of magnitude lower than for the InAs NW-FETs. Although the results are not transferable to InAs, they indicate that under the right conditions, it would be possible to substantially reduce the trap concentration.

The extracted values of α_H are interesting for evaluation as they can impose a lower limit of the level of LFN for a specific material. For a meaningful comparison of α_H, however, it is important that the parameter extraction is correct such that the measured noise originates from mobility fluctuations. If α_H was to be determined from the data in Fig. 4.1, it would only be correct in the higher and the lower end of the graph along the x-axis. The noise level in the middle of the graph is instead associated with number fluctuations and would therefore yield a higher value of α_H than the value that is associated with scattering.

It has been argued from an empirical stand-point that the lower bound of α_H for InAs systems may have been reached [53] since a similar lowest value has been extracted in several studies ($\alpha_H \sim 5 \cdot 10^{-4}$) [53] [54]. With the theory for the physical origin of mobility fluctuations not fully in place, however, there is currently no theoretical support of a specific number. Our study suggests that a lower value can be achieved when the scattering associated with the surface is substantially reduced ($\alpha_H \sim 4 \cdot 10^{-5}$) [21].

4.2 **High-Frequency Noise**

At frequencies above f_C and without LFN up-scaling, only thermal noise and shot noise contributes to the measured noise level. While thermal noise is proportional to the device resistance, the level of shot noise is instead proportional to the current flow through the device. Shot noise is associated with uncorrelated electron transport over a barrier and corresponds to $2qI$ for a device current I [47]. Devices constructed around transport over barriers, such as diodes and bipolar transistor devices, can have shot noise levels on par with the corresponding thermal noise of the equivalent resistance of the device [47]. For FET devices, however, thermal noise is typically the dominating source.

A resistor has a corresponding generated thermal noise current or thermal noise voltage equal to $I_N^2 = 4k_BT\gamma_eR$ A^2/Hz and $V_N^2 = 4k_BT\gamma_e/R$ V^2/Hz, respectively. The excessive noise factor, γ_e, is a material and structural specific parameter that for ideal devices is found to be about 2/3. Examples of ideal devices are metal strip resistors and non-aggressively scaled semiconductor transistors with heavy effective electron masses, like long-channel Si-FETs. For short channels and light effective mass materials, however, the γ_e–factor can be several times higher. The noise figure (NF) is a measure of the ratio between the SNR at the output compared to the SNR at the input, and is coupled with γ_e of the implemented technology. If a

signal is amplified without any addition of noise, NF = 0 dB. Recent studies of simulated semi-ballistic III-V MOSFETs, with L_G = 50 nm, suggest that competitive noise performance can be achieved with NF = 0.7 dB at 60 GHz [55]. Also, by having a buried channel design, the level noise can be further reduced.

4.2.1 **Noise Figure for an Amplifier Stage**

To be able to quantify the NF in a specific circuit, it can be helpful to construct an equivalent circuit where the different noise contributions are added as separate sources. An example of an equivalent amplifier circuit is shown in Fig. 4.4, where R_s is the source output resistance and R_{in} is the amplifier input resistance. The noise power on the input is given by [56]:

$$N_i = k_B T \ [W/\mathrm{Hz}]. \tag{4.6}$$

The noise power on the output, referred to the input, is the summation of the input noise power and the noise power generated by the amplifier:

$$N_{oi} = \frac{e_{n,s}{}^2 + e_{n,in}{}^2 + i_{n,in}{}^2 R_S}{4R_S} [W/\mathrm{Hz}]. \tag{4.7}$$

In Eq. 4.7, $e_{n,s}$ is the corresponding noise voltage induced by R_s, $e_{n,in}$ is the noise voltage induced by R_{in}, and $i_{n,in}$ is the equivalent noise current generated by the amplifier and that accounts for a noise voltage addition to the output corresponding to $i_{n,in} \cdot R_s$. By replacing $e_{n,s}{}^2$ with the expression for the noise voltage amplitude, Eq. 4.7 can be rewritten as:

$$N_{oi} = \frac{4k_B T_o R_S + e_{n,in}{}^2 + i_{n,in}{}^2 R_S}{4R_S} [W/\mathrm{Hz}]. \tag{4.8}$$

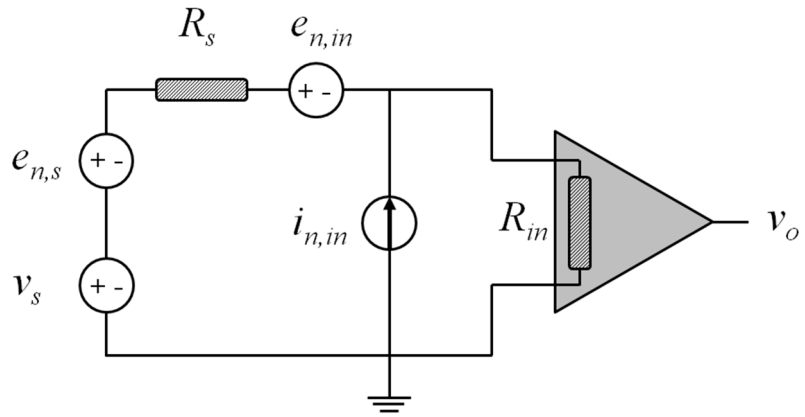

Fig. 4.4. *The equivalent circuit of an amplifier stage including the different noise sources. In the schematic, R_{in} is the input resistance of the amplifier, R_S is the output resistance of the source, v_o is the output voltage, v_s is the source voltage, $e_{n,s}$ is the equivalent noise voltage due to R_S, $e_{n,in}$ is the equivalent noise voltage due to R_{in}, and $i_{n,in}$ is the equivalent noise current due to R_{in}, which will add a corresponding noise voltage to the output with an amplitude of $i_{n,in} \cdot R_s$.*

The noise factor, *F*, which when given in dB corresponds to the NF, is calculated as the noise power at the output divided by the noise power at the input:

$$F = \frac{N_{oi}}{N_i} = 1 + \frac{{e_{nd}}^2 + {i_{nd}}^2 R_S}{4k_B T_o R_S}. \quad (4.9)$$

4.2.2 Friis Formula

Transceivers are often designed as a cascade of different amplifier stages with different purposes. For a chain of amplifiers it can be shown that for an input stage with sufficient gain, the noise in the following stages will have a minor impact on the total NF. Friis formula for noise [56] specifies the relation of *F* between the output of the last stage and input of the first stage,

$$F_{OUT} - F_{IN} = F_1 + \frac{F_2 - 1}{G_1} + \frac{F_3 - 1}{G_1 G_2}, \quad (4.10)$$

where G_k and F_k are the gain and the noise factor of the stage *k*, respectively. Eq. 4.10 represents a chain of 3 stages but can be expanded to an infinite number of stages where the contribution to the overall *F* of each consecutive stage can be calculated as $(F_m - 1) / \prod_1^{m-1} G_k$, where m is the index of the m:th stage. Fig. 4.5 shows a cascade of three amplifier stages. If *G* is known for all the stages and *F* is known for the input and output stage, then F_2 of the device under test (DUT) is given as:

$$F_2 = 1 + G_1 \left(F_{OUT} - F_{IN} - F_1 - \frac{F_3 - 1}{G_1 G_2} \right). \quad (4.11)$$

Friis formula is often used to determine the NF of an un-characterized amplifier stage in a chain of stages from a measurement of the total noise power, using a noise source with a known SNR at the input stage.

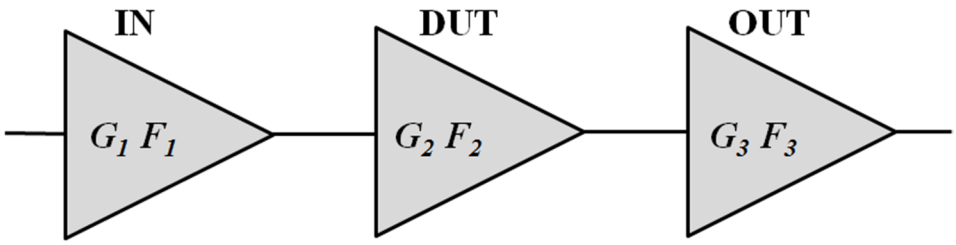

Fig. 4.5. *A chain of amplifier stages where each stage has the associated noise factor F_k and the gain G_k. The middle stage is here referred to as device under test (DUT).*

CHAPTER 5

5 RF Circuits

It's not true I had nothing on, I had the radio on.
- Marilyn Monroe

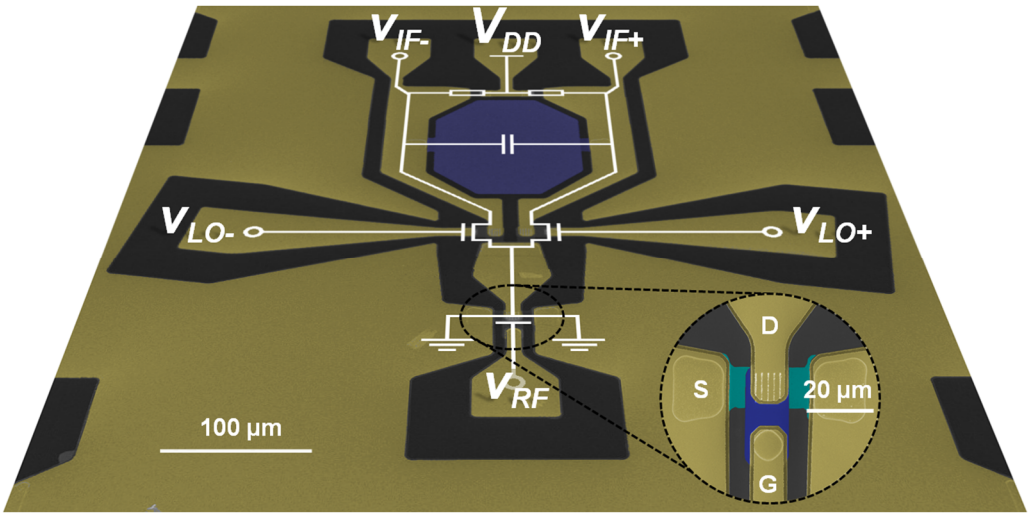

Fig. 5.1. *A fabricated active single balanced, differential mixer with a schematic circuit layout drawn on top. The inset is showing the RF NW-FET at a larger magnification where G, S and D corresponds to the gate contact, the source contact, and drain contact, respectively.*

Electronics will continue to expand into our daily lives as the industry is aiming to connect 50 billion devices until the year 2020 [1]. Radio frequency (RF) circuits are needed to transmit data and constitute an important component of future electronics. One way to understand how electronics works is to identify the hierarchy of distinct abstraction levels; there are physical units limited by transport properties (transistors etc.), circuits that connect different physical units into functional units, and there as systems that can perform complete tasks. For good implementations, the co-dependence on the strengths and weaknesses associated with each level has to be considered. A big part of circuit design is devoted to make up for shortcomings in transistor performance. As Si- and III-V-technologies have different shortcomings, it is likely that more future research will focus on how to

optimize the implementation of III-V technology, which comparably is far less explored. As a III-V technology demonstration, we have fabricated a mixer circuit, implementing InAs NW-FETs. The fabricated circuit is shown in Fig 5.1 and in paper I and paper II, investigations of DC and RF characteristics of circuits like the one shown, are reported.

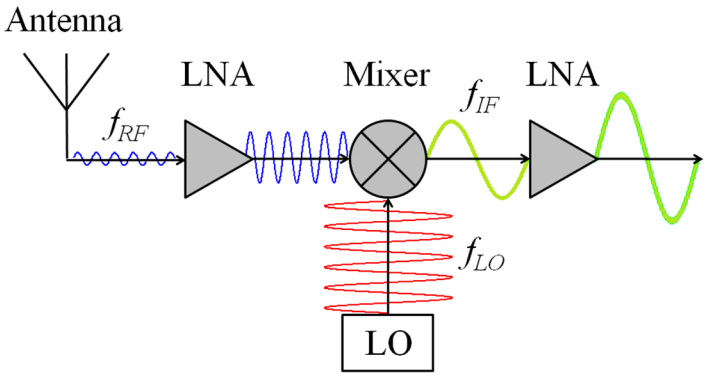

Fig. 5.2. *A radio front-end with a chain of sub-circuits consisting of an antenna, an input LNA, a mixer, and an output LNA. The blue curve represents the incoming RF signal with frequency f_{RF}, which is amplified at the input LNA stage. At the mixer stage, the RF signal is down-converted to an intermediate frequency, f_{IF}, represented in green, as it is multiplied with the LO signal, represented in red, with frequency f_{LO}. The gain of the mixer is less than that of the output LNA stage, which amplifies the IF signal further.*

A circuit can consist of a single transistor, using two of the three fundamental nodes as input and output, and be seen as with a single purpose such as amplification. Going up in architectural complexity, by connecting a chain of single purpose sub-circuits, they can collaboratively perform more complex functions, for example as a front-end transceiver, which can transmit and receive radio signals. A signal flow chart of a front-end sub-circuit chain is shown in Fig. 5.2. In the chain there is one input low-noise-amplifier (LNAs), one output LNA, one mixer, one antenna, and one local oscillator (LO). The signal received at the antenna is amplified, down-converted, and then amplified again. To not only receive a signal but also interpret the information, other sets of circuits have to be added, such as circuitry for signal processing. A collection of circuit sets that together can perform a complete task could be referred to as a system.

Common for analog circuit design is that it is striving to minimize noise and maximize gain while keeping the power consumption at a minimal. For performance optimization, it is important to know the application space of a circuit as there is no ultimate design; with every enhancement there is a drawback. For example, with increased circuit bandwidth, the thermal noise power becomes larger and inflicts negatively on the SNR-level (see chapter 4). The following sections will focus on different aspects of a specific sub-circuit in the RF front-end; the active differential mixer.

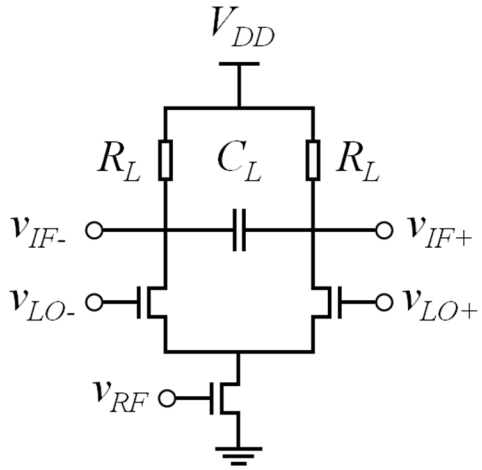

Fig. 5.3. *A schematic layout of an active single balanced differential mixer. V_{DD} is the drive voltage, R_L is the load resistance, C_L is the load capacitance, $v_{IF}\pm$ are the voltages at the ±IF output nodes, $v_{LO}\pm$ are the voltages at the ±LO input node, and v_{RF} is the voltage at the RF input node.*

5.1 **The Mixer Circuit**

The mixer circuit can convert an input RF signal by multiplication it with a LO signal, producing intermediate frequency (IF) signals which are the sum and difference of the two. In Fig. 5.3, a schematic layout of an active single-balanced differential mixer is shown. This is the same circuit layout as for the fabricated circuit shown in Fig. 5.1. The circuit can be divided to three different parts; a transconductance stage (RF input), a switching stage (LO differential inputs with 180° phase difference), and a transimpedance stage (IF differential output). An active mixer design, in contrast to a passive mixer, consumes a direct current (DC) power but has the advantage of contributing with a positive gain. The low-frequency (low-*f*) differential voltage conversion gain, G_{VC}, is set by the multiplication of the conversion factors of the first and third stage [56];

$$G_{VC} = \frac{2}{\pi} g_{m,RF} R_L, \qquad (5.1)$$

where $g_{m,RF}$ is the transconductance of the RF input transistor and R_L is the resistance of the load resistors. The differential design reduces the common signal contribution, such as RF-IF leakage, as the resulting output signal is the difference between the left and right circuit branch. At each output IF node, R_L is coupled in parallel with a load capacitance, C_L, forming a low-pass RC-filter and thus reducing the impact of the leakage of both the RF and LO signals.

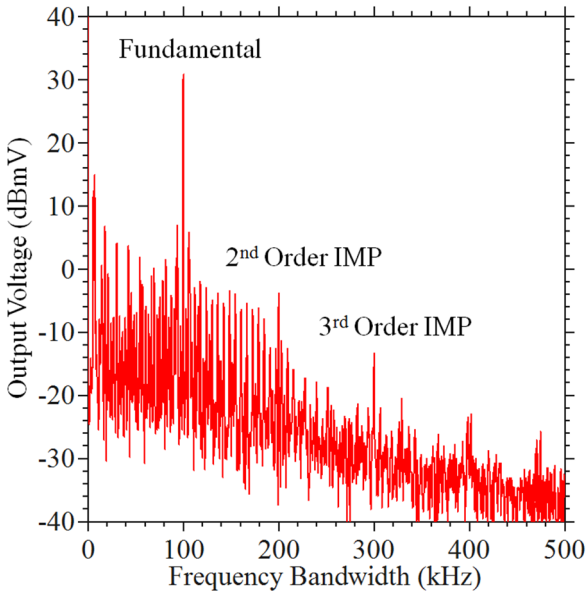

Fig. 5.4. *Sampled output data from a measurement of an active mixer showing a frequency spectrum of 0-500 kHz. The fundamental (1st), 2nd, and 3rd order intermodulation product (IMP) peaks are indicated.*

The converted signals at the IF output are called inter modulation products (IMP), where there is a 1st order tone (the fundamental) and overtones. A measurement of the IF output voltage is shown in Fig. 5.4. In this single-tone test, the RF frequency, f_{RF}, was set at 1.0001 GHz and the LO frequency, f_{LO}, was set at 1.0000 GHz. The 1st order IMP can be seen as the highest peak, at 100 kHz, where the 200 kHz and 300 kHz overtone peaks are also visible. The 3rd overtone peak increases with the cube of the input power and one important measure is the theoretical crossing of the amplitude of the 1st and 3rd IMP, called the input third order intercept point (IIP3). During operation it is vital to be sufficiently below the IIP3. Another important measure is the linearity versus input power, where the output may go into compression for large input signals. Depending on if the compression is related to the RF input or the IF output, it is called input referred or output referred compression, respectively. A third important measure is the highest RF frequency bandwidth (BW), which is related to the capacitive loads in the high frequency nodes. The highest operational frequency is thus highly dependent on the transistor architecture of the RF and LO transistors. At a high enough RF frequency, the mixer gain starts to roll off, however, a mixer can operate at a RF frequency in the roll-off as long as the gain is high enough. Using an inductor-resistor-capacitor (LRC) network, a frequency saddle point can be created in the roll-off slope. A LRC-network will also suppress frequencies that are higher or lower than the desired RF frequency, thus reducing possible disturbances at other frequencies. The IF BW is often kept at a minimal needed as the thermal noise

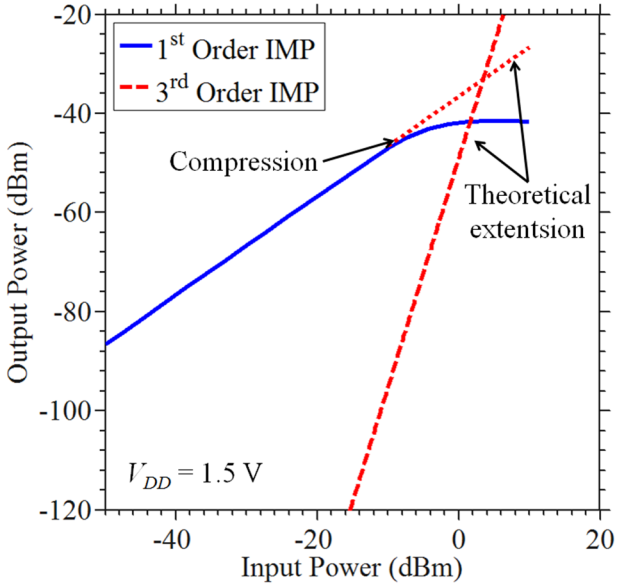

Fig. 5.5. *Output power plotted versus input power, showing both compression of the 1st order IMP as well as the theoretical crossing of the 1st and 3rd order IMP.*

power at the output scales with the IF BW. If the mixer operates at a RF frequency in the frequency gain roll off, linearity requirements impose a limit on the maximum IF BW.

5.1.1 **Mixer Noise Figure**

The current noise spectral density, S_I, of a MOSFET can be described by [47]:

$$S_I = 4k_B T\gamma_e g_m + \frac{K_1 {g_m}^2}{f} + \frac{K_2 I_{DS}}{f} [A^2/Hz^{-1}]. \qquad (5.2)$$

In Eq. 5.2, the constants K_1 and K_2 are introduced as a simplified mean of modelling the 1/*f*-noise associated with number fluctuations and mobility fluctuations, respectively. The physical constants and the transistor parameters that are included in K_1 and K_2 are given in chapter 4, which also explains γ_e.

Ignoring the 1/*f*-noise contribution, the gate input referred voltage spectral density, S_{Vin}, for a MOSFET is given by:

$$S_{Vin} = \frac{4k_B T\gamma_e}{g_m} [V^2/Hz^{-1}]. \qquad (5.3)$$

Referring the noise to the input is often preferable as the value can then be directly compared to the source impedance. For a common source amplifier, which is a transistor in series with a resistor, R_L, S_{Vin} is given by:

$$S_{Vin} = 4k_BT\left(\frac{\gamma_e}{g_m} + \frac{1}{{g_m}^2 R_L}\right) [V^2/Hz^{-1}]. \quad (5.4)$$

Eq. 5.4 can be compared to S_{Vin} for an ideal active single balanced differential mixer [56]:

$$S_{Vin} = \pi^2 k_BT\left(\frac{\gamma_e}{g_{m,RF}} + \frac{2}{{g_{m,RF}}^2 R_L}\right) [V^2/Hz^{-1}]. \quad (5.5)$$

Ignoring the second term in the parenthesis, there is a factor $\pi^2/4$ difference due to the frequency conversion. For a non-ideal mixer circuit, there are other contributions in addition to those given in Eq. 5.5 as well, related to finite LO transistor switching speeds, parasitic series resistances, and parasitic capacitive loads. Depending on the type of transistors and the frequency bandwidths, 1/*f*-noise may be the largest non-ideal noise contributor. The LO mixer pair will be the two main sources of 1/*f*-noise as the noise from the RF transistor is up-converted. The 1/*f*-noise voltage spectral density present at the IF output, $V_{N,IF}^2$, originating from the 1/*f*-noise in each of the LO switching pair transistors, $V_{N,LO}^2$, can be expressed as [56]:

$${v_{N,IF}}^2(f) = 2\left(\frac{I_{DD}R_L}{\pi v_{LO}}\right)^2 {v_{N,LO}}^2(f) [V^2/Hz^{-1}]. \quad (5.6)$$

Diving the expression in Eq. 5.6 by the mixer gain given in Eq. 5.1, yields an expression for the noise voltage spectral density instead referred to the input, $S_{Vin}(f)$:

$$S_{Vin}(f) = 2\left(\frac{I_{DD}}{g_{m,RF}v_{LO}}\right)^2 {v_{N,LO}}^2(f) [V^2/Hz^{-1}]. \quad (5.7)$$

In Eq. 5.7, v_{LO} is the LO switch peak voltage amplitude. Combining Eq. 5.7 and Eq. 5.5 gives the total input referred noise voltage spectral density at the output, considering both the thermal noise and 1/*f*-noise. Ideally MOSFETs have very small gate currents and noise current source, shown in Fig. 4.4, can be ignored. The total noise power spectral density at the output can thus be expressed as:

$$N_{IF} = \frac{4k_BTR_S + S_{Vin} + S_{Vin}(f)}{4R_S} [W/Hz^{-1}]. \quad (5.8)$$

In Eq. 5.8, R_S is the source resistance. The dual sideband (DSB) noise factor can be expressed as:

$$F_{DSB} = 1 + \frac{S_{Vin} + S_{Vin}(f)}{4k_BTR_S}. \quad (5.9)$$

The single sideband (SSB) noise factor is up to two times higher than the DSB noise factor, or measured in noise figure, 3 dB higher. The difference between the two definitions is that for the DSB, the RF signal carriers an IF signal at a frequency both below and above the LO frequency. For the SSB, the signal is only present on one side and as the image frequency is down converted as well, the noise power is doubled.

5.1.2 **Mixer Simulations**

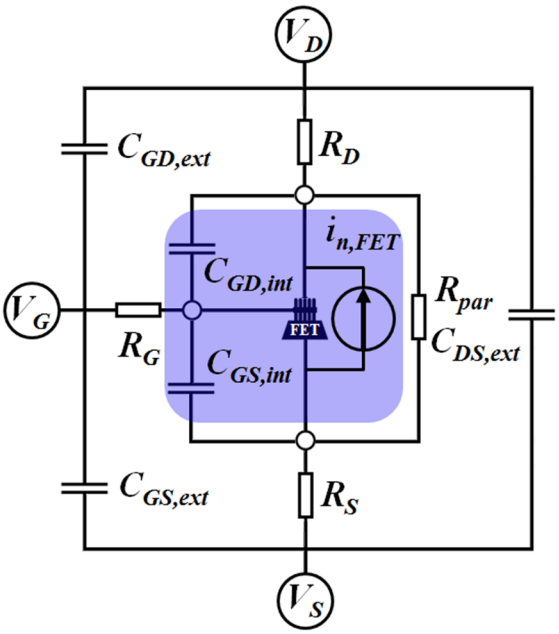

Fig. 5.6. *The Cadence model used in RF circuit simulations, consisting of a RC-network built around a Verilog-A intrinsic model. The intrinsic modelled components (highlighted in blue) include capacitances, $C_{GD,int}$ and $C_{GS,int}$, a noise current, $i_{n,FET}$, and the conductance model. The added extrinsic components include capacitances, $C_{GD,ext}$, $C_{GS,ext}$, and $C_{DS,ext}$, and resistances, R_S, R_D, R_G, and R_{par}.*

A Verilog-A MOSFET model has been implemented around the extracted transport properties of the device plotted in Fig. 2.6b (D_{NW} = 45 nm, L_G = 200 nm), using the described compact virtual source (VS) model in chapter 2. To optimize the model device for circuit performance, the output conductance related to parasitic nanowire core conduction was reduced to 1/20, assuming a lower doping in the channel region can be accomplished. Devices with D_{NW} = 28 nm show parasitic conductance values of about 1/50. Also, to increase high frequency performance for the implemented model, L_G is reduced to 60 nm in order to decrease the intrinsic gate capacitance. Due to the reduction of L_G, the injection velocity was assumed to increase about 10 % [57]. Parasitic series resistances are kept as it was

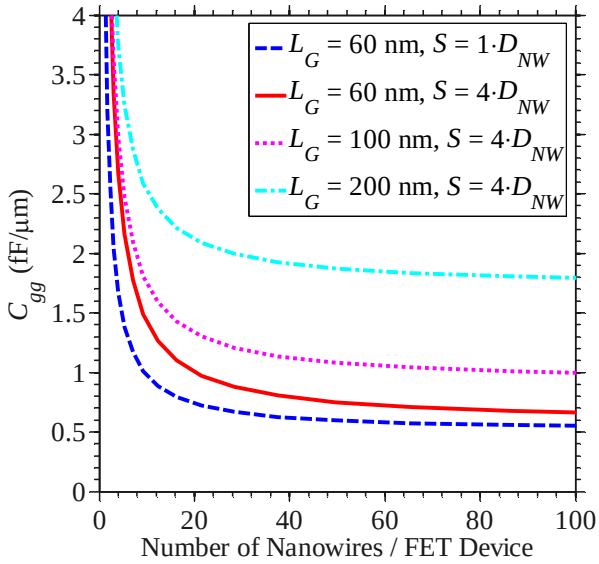

Fig. 5.7. *Total gate capacitance, C_{GG}, versus the number of nanowires per FET device, illustrating the difference with regards to NW spacing, S, as well as L_G. The values for the extrinsic capacitances correspond to the proposed InAs NW-FET 50 nm node found in reference* [40].

extracted from measurement. The performance enhancements at V_{DS} = 0.5 V, attributed to the modifications, can be summarized as; g_m increase from 1.3 to 1.7 mS/μm, the intrinsic voltage gain increased from 3.8 to 6.8, and the intrinsic gate

TABLE 5.1. BENCHMARK OF MIXER PERFORMANCE OF SIMUALTED CIRCUIT AT DIFFERET BIASES

V_{DD} (V)	0.9	1.5
DC Power (mW)	0.43	0.84
Low-f G_{VC} (dB)	8	12
1-dB-BW (GHz)	60-100	60-120
3-dB-BW (GHz)	60-150	60-180
6 dB G_{VC} (GHz)	100	340
0 dB G_{VC} (GHz)	450	1200

capacitance decreased from 1.87 to 0.54 fF/μm. The voltage dependence of the intrinsic capacitance is accomplished with the charge model from the VS model, thus coherent with the modeling of the conduction [18].

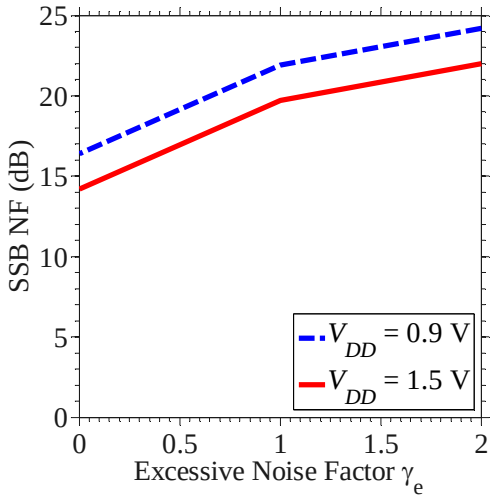

Fig. 5.8. *The single sideband (SSB) noise figure (NF) for two different circuits biases and varying the excessive noise factor γ_e.*

The extrinsic resistances and capacitances are taken from investigations of realistic layouts of InAs nanowire MOSFET integration, with a proposed roadmap, given in reference [40]. Values for the 50 nm node are used. A schematic RC-network layout is shown in Fig 5.6. The dependence on L_G and the distance between nanowires in the array for the total C_{GG} is shown in Fig. 5.7. The Cadence mixer circuit layout used in the simulations is shown in the Appendix A2. In table 5.1 is a comparison of performance metrics between two different circuit biases. In Fig. 5.8, the SSB noise figure is plotted for different values of γ_e. In the noise model, the measured level of V_{GS} dependent 1/*f*-noise from the study in paper VII is included. The simulated total noise power indicates that most of the noise originates from the load resistors and the parasitic series resistances. A more thorough study of the noise in the mixer circuit, using the here presented model, is given in reference [58].

5.2 **Benchmark**

To contrast the simulated performance values and the measured performance values reported in paper I and II, a benchmark against high frequency Si-technology implementations [59] [60] [61] is given table 5.2. In the table, p_{LO} is the LO switching power, P1dB is the 1 dB compression point, and G_{VC} is the voltage conversion gain.

There exist a great number of circuit enhancements that can be implemented to reduce noise and power consumption. Two of the benchmarked Si circuits have inductive loads, which produce significantly less thermal noise than resistors while also consuming smaller amounts of DC power. One of the mixers has implemented a double balanced design, which increases port isolation but reduces gain and increases noise. The optimal circuit design, however, can only be achieved

knowing the target application as there are advantages and drawbacks associated with each design. A set of different circuit enhancement methods for the modeled InAs NW-FETs are investigated in reference [58].

TABLE 5.2. BENCHMARK OF MIXER PERFORMANCE OF DIFFERENT TECHNOLOGIES

	This work	Paper I and II	D.-H. Kim, J.-S. Rieh 2012 [59]	D.-H. Kim, J.-S. Rieh 2012 [60]	O. Inac, et al 2011 [61]
Evaluation	Simulated	Measured	Measured	Measured	Measured
Technology	III-V NW FET	III-V NW FET	Si-CMOS	SiGe BiCMOS	SOI CMOS
Node (nm)	-	-	90	180	45
L_G (nm)	60	200	-	-	-
Balancing	Single	Single	Single	Double	Single
Load (R/L)	R	R	jωL	R	jωL
RF (GHz)	340	5 (14)	138	135	145
BW1dB (GHz)	60-120	0.1-1	-	-	-
BW3dB (GHz)	60-180	0.1-2	-	-	145–161
V_{DD} (V)	1.5	1.5	1.2	3	1.5
DC Power (mW)	0.84	3.8	0.79	3.9	10
pLO (dBm)	6	-3	10	10	2.8
G_{VC} (dB)	6	6	3.7	11.5	-4
SSB NF (dB)	22	-	-	-	-
P1dB (dBm)	-7	-20	-6	-20	-
IIP3 (dBm)	3	0	-	-	> 16
LO-RF (dB)	-78	-	-	-31	-

CHAPTER 6

6 Nanofabrication

If an elderly but distinguished scientist says that something is possible, he is almost certainly right; but if he says that it is impossible, he is very probably wrong.
- Arthur C. Clarke

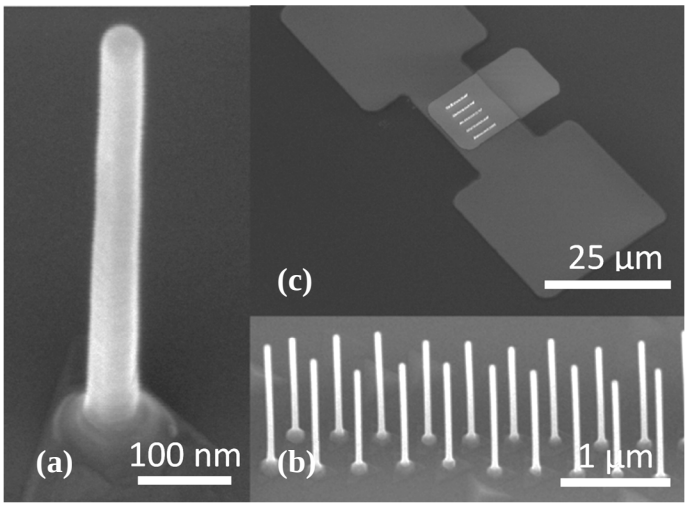

Fig. 6.1. *Three different SEM-images showing structures at different magnification.* **(a)** *A single nanowire.* **(b)** *An array of seeded nanowires.* **(c)** *Several rows of arrays where the mesa structure has been defined as well as the gate pad.*

The definition of nanofabrication could be argued as processes involving the formation of sub-100-nm material feature sizes. The semiconductor industry has since long relied on processes that can produce thin films (a film with a thickness measured in µm or nm). Methods such as thermal evaporation, molecular beam epitaxy, and atomic layer deposition (ALD), are able to control film thicknesses down to individual atom-layers. Among the difficulties associated with nanofabrication are the processes of adding and removing material at defined locations. Other challenging aspects are preparation of surfaces and control of deposition conditions in order to make good materials and material interfaces.

An important evaluation tool in fabrication development is the scanning-electron-microscope (SEM), which uses scattered electrons rather than photons to image structures down to a few nm. In Fig. 6.1a, b and c, SEM images of nanowires are shown at different magnification. This chapter will give a brief overview of the process flow in the fabrication of NW-FETs. A more detailed description is given in paper IV. The last section of the chapter will describe the fabrication of the bottom electrode separation layer in greater detail, comparing two different methods.

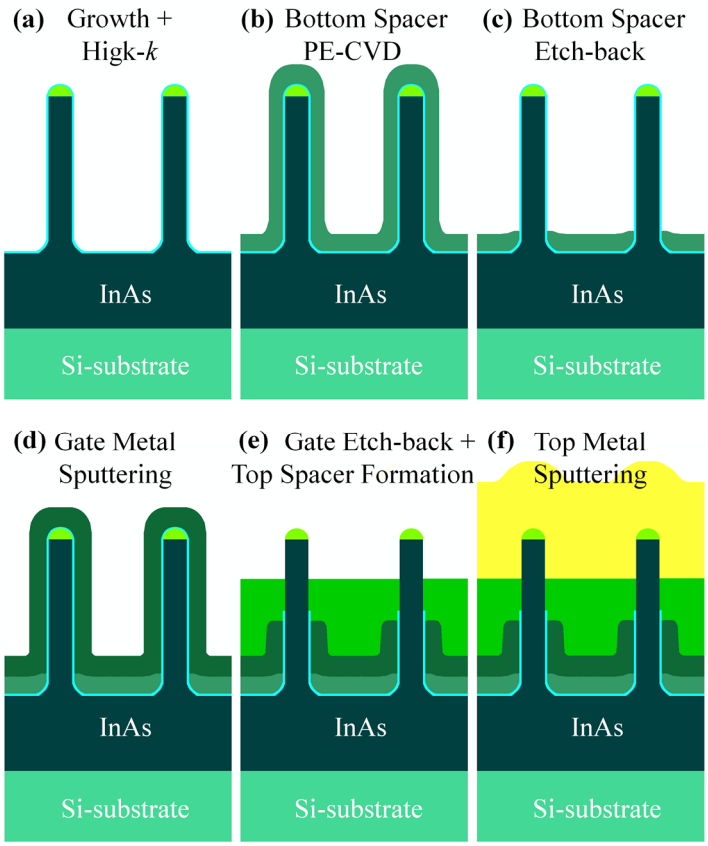

Fig. 6.2. *Schematic fabrication flow of NW-FETs, starting with* **(a)** *nanowire growth,* **(b)** *bottom spacer PE-CVD,* **(c)** *bottom spacer etch-back,* **(d)** *gate metal sputtering,* **(e)** *gate etch-back and top spacer formation,* **(f)** *and ending with top metal sputtering.*

6.1 **The Vertical Challenge**

One way to selectively add or remove material is to use films that can be patterned and easily removed, such as spin-on polymers. In a research laboratory, Electron-beam-lithography (EBL) can typically make line widths down to 10-20 nm, while a UV-lithography system may have a resolution of about 500 nm. Once the polymer

is exposed, developed and baked (where cross-polymerization makes the polymer more etch-resistant), it can act as either a protective mask in an etch-procedure or be used in a lift-off procedure where the material on the mask is removed along with the mask using a solvent.

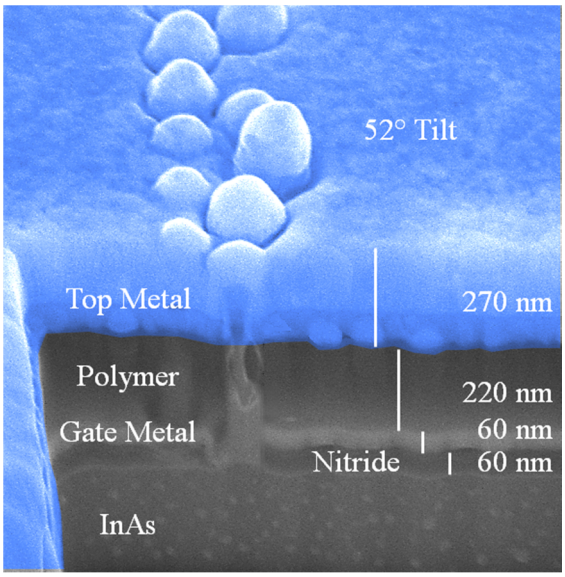

Fig. 6.3. *A SEM image showing the different layers in a fabricated NW-FET. The image was taken after milling through a nanowire array using an ion-beam.*

Nanowire transistors can be fabricated using either vertical or lateral channels. For a lateral channel, the critical dimensions, such as L_G, are limited by the resolution of the lithography system, and using EBL, good precision can be achieved. When instead processing a channel in the vertical direction, the accuracy in determining the layer thicknesses are dependent on the control and timing of etching procedures. Both the vertical and the lateral channel approach have advantages and drawbacks. The process flow of fabricating our vertical NW-FETs is shown in Fig. 6.2a-f. EBL lithography is only used in one of the initial fabrication steps, where the size and placement of gold dots are defined. The dots later seed the growth of nanowires in an epitaxial growth reactor. The InAs nanowires are grown on a 300-nm-thick InAs contact layer, which is etched-out into source mesas using UV-lithography and a wet-etch procedure. An etched-out source mesa structure is shown in Fig. 6.1c, where the bottom segments of the NWs are to be connected to source metal electrodes with dual via-holes, one in each opposite direction of the device area, thus reducing the parasitic source resistance. An ALD process is used to deposit the dielectric high-*k* film. The bottom separation layer between the source mesa and the gate pad electrode is made with plasma-enhanced chemical-vapor-deposition (PE-CVD) together with an etch-back process that is described in greater detail in section 6.2. The gate

metal is sputtered and defined with use a polymer film which is back-etched and then used as a metal etch-mask. The top separation layer, separating the gate and drain electrode pads, is fabricated using an etched-back polymer film that is permanent baked (200 °C). The top metal is sputtered. Apart from the initial EBL step to define the nanowires, all of the lithography is done with UV, such as mesa structures and contact pads. A cross-sectional view of a fabricated NW-FET is shown in Fig. 6.3.

The described method for NW-FET fabrication differs in several aspects from the method used to fabricate the studied NW-FETs in paper V. The reason for the transition, where the procedures where changed into the fabrication method

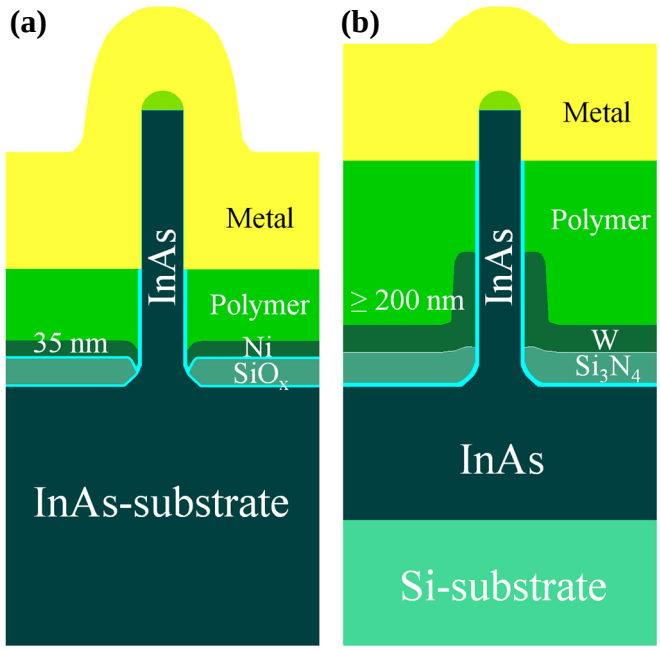

Fig. 6.4. *Schematic structures of NW-FETs fabricated on* **(a)** *an InAs-substrate and* **(b)** *on a Si-substrate.*

described in paper III, was in part to make it compatible with RF implementations. In the shift, the InAs substrate was replaced with a semi-insulating InP-substrate that was overgrown with an InAs contact layer. This enabled the fabrication of isolated mesa structures, thus substantially reducing the parasitic load. The other reason for a transition was to improve the process stability where the sputtered gate process was found to be more stable than the process involving an evaporated gate. The downside of the sputtered gate, however, is the limited possibility of making a short L_G. In papers I, II, IV, VI, and VII, fabrication procedures were developed further, and the InP-substrate was replaced with a Si-substrate. The bottom polymer spacer was also replaced from an organic film to a PE-CVD film. A schematic comparison of the InAs-substrate transistor structure and the Si-substrate transistor structure is shown in Fig. 6.4a and b.

6.2 **Contact Spacer Formation**

Spacers, material films used for separating different electrical contact layers, are made in a balancing act between the different magnitudes of parasitic elements. Making too thick spacers adds excessive series resistance while too thin spacers will bring down the high frequency performance due to large parasitic capacitances. For research purposes, using an organic spacer has the advantage of ease of appliance with a spin-on procedure followed by an oxygen plasma etch to define the thickness, where the resulting film also has a low permittivity. However, the control of the thickness is at best within 50 nm due to unevenness from the spin-on as well as unevenness from the thickness plasma etching. Organic materials also have issues with stability over time as they have tendencies of absorbing moisture. Alternative materials are low-*k* oxides and nitrides such as SiO_2 and Si_3N_4. These materials have been widely used for electrical/optical structure integration, and ways of applying and etching [62] [63] are well understood. This section will describe two procedures to fabricate the bottom spacer layer using a plasma-enhanced chemical-vapor deposition (PE-CVD) process followed by either a wet-etch or a dry-etch procedure.

6.2.1 **Formation of the SiO_2 spacer**

The PE-CVD deposition of SiO_2 results in a film thickness ratio between vertical and lateral layers at 3:5. Etch rates of SiO_2 were established for two different concentrations of hydrofluoric acid (HF) and H_2O solutions; a 1:10 solution of

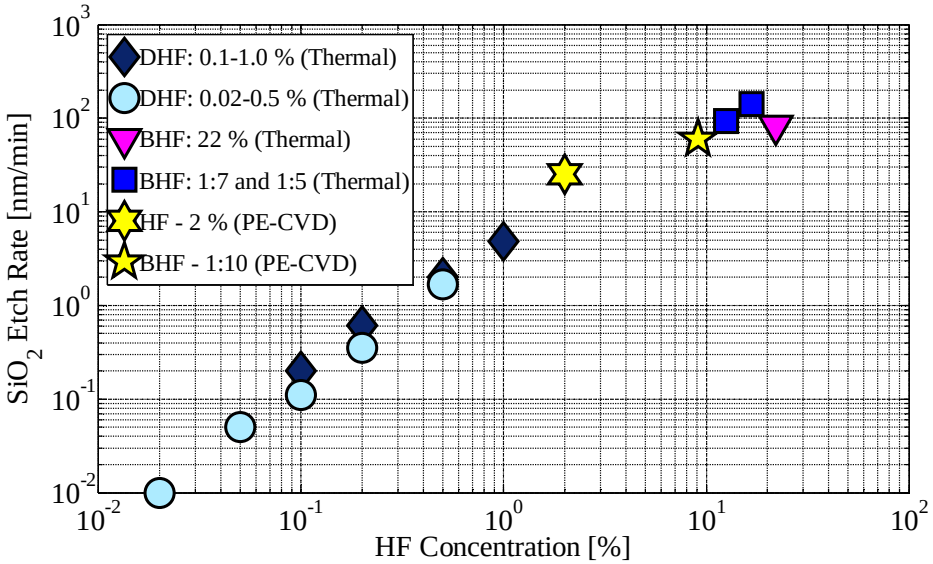

Fig. 6.5. *Etch rates of SiO2 for different concentrations of HF. The included studies were performed with a thermal oxide while this study used a PE-CVD film.*

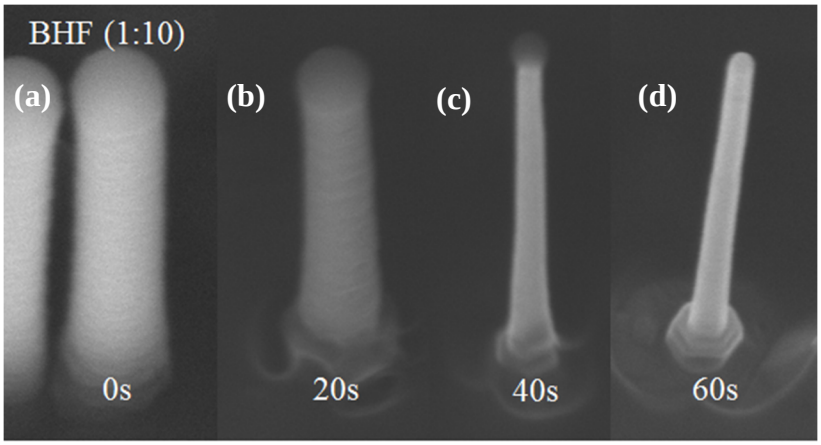

Fig. 6.6. *SEM images of nanowires covered with a SiO_2 film with a lateral thickness of 150 nm.* **(a)** *After PE-CVD.* **(b)** *After 20 s BHF etching.* **(c)** *After 40 s BHF etching.* **(d)** *After 60 s BHF etching.*

buffered HF (BHF) with an etch rate of 100 nm/min, and a 1:50 solution HF with an etch rate of 25 nm/min. Etch rates reported in different studies [64] [65] [66] [67] are plotted together with the here determined etch rates in Fig. 6.5. It is expected that the etch rate for HF drops at concentrations above about 10 % due to a decrease in pH [65], hence the use of BHF. Although the deposited SiO_2 film is amorphous and the wet-etch is inherently isotropic, the SiO_2 is etched at a faster rate on the NW structures as compared to the lateral layer. One reason for this could be the decrease in etchant flow close to the lateral surface. In combination with a thinner deposited vertical film, the oxide on the sides of the NWs is removed after an etch time for which less than half of the lateral layer thickness has been etched. Images of NWs with 150 nm SiO_2 lateral thickness deposited are shown in Fig. 6.6a, b, c, and d, after 0s, 20s, 40s, and 60s, respectively. Depending on the NW base geometry, cavities of different severity are observed, possibly indicating that the etchant flow is increased around the NW base as compared to the lateral surface further away. When reducing the deposited film thickness and size of the growth-parameter-related NW base, the size of the cavity is reduced. A cavity may pose a detrimental impact on analogue device performance due to excess parasitic capacitance.

6.2.2 Formation of the Si_3N_4 spacer

The PE-CVD deposition of Si_3N_4 results in a film thickness ratio between vertical and lateral layers at 8:9. In order to only remove the film deposited on the sides of the NWs, a protective etch-mask is needed. To selectively dry-etch Si_3N_4 deposited on the sides of the NWs, a spun-on polymer can be applied to protect the lateral film. The thickness of the polymer can be set with an accuracy of about 150 nm

(over a 10x10 mm^2 sample) using a high pressure 50 W oxygen-plasma reactive ion etch (RIE). The nitride film can then be etched in a following step using a 140 W SF_6/Ar-plasma with 45 and 10 sccm gas flows, respectively, under a 185 mTorr pressure. The dry-etch will have a mixed anisotropic and isotropic behavior due to the simultaneous chemical and physical etching with high gas flows and high pressure, and high acceleration power. The Ar gas is added to physically etch redeposited material. For a lateral film thickness of 90 nm, the etch-time is set between 25-50s depending on resist-layer-thickness and the NW height. From observations of several etch-tests, it can be concluded that the etch-rate of the nitride-film deposited on the NWs, which is etched trough openings in the resist, is substantially lower than for the film etehc on wires that are sticking up above the polymer film. This phenomenon is known as RIE-lag which is accounted for when etching high aspect ratio trenches [68]. Factors that will impact on the trench etch rate are many, including absolute size of openings and the rate of material redeposition.

In one test, two samples with arrays of NWs were processed side by side. Arrays of NWs on these samples were inspected before and after Si_3N_4 deposition and in both instances it was found that the NW length was around 1 μm and that the vertical Si_3N_4 film thickness was about 80 nm. In Fig. 6.7a, a NW array is shown after Si_3N_4 deposition. In Fig. 6.7b, an array on one of the samples is shown after 30s dry-etch. In Fig. 6.7c, an array on the other sample is shown after 40s dry-etch. The under- and over-etch, respectively, illustrate the narrow process-window for the etch-time and the need for a correlation to the NW length and polymer-film-thickness. A first step in a more thorough investigation would be to determine the etch-rate of the Si_3N_4 film that is situated on the NW, above and beneath the resist-edge, respectively. Studies found in literature are can give also give insight, however, they tend to be more indicative of parameter trends rather than absolute values as etch chambers are inherently unique and will give different results.

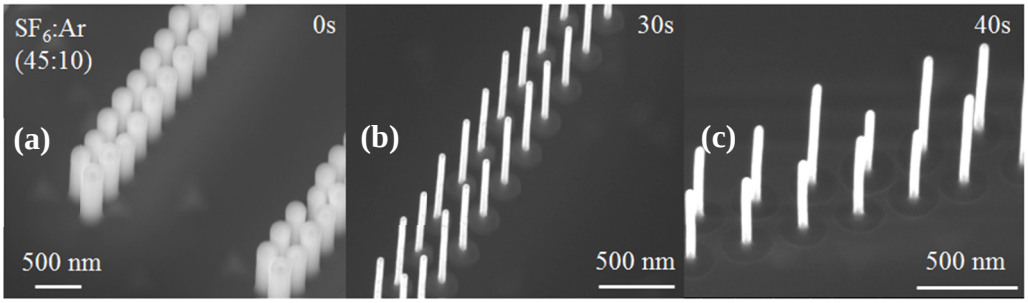

Fig. 6.7. *SEM images of arrays of nanowires covered with a PE-CVD Si_3N_4 film with a lateral thickness of 90 nm.* **(a)** *After PE-CVD.* **(b)** *After 30 s plasma-etch.* **(c)** *After 40 s plasma-etch.*

6.2.2.1 Method Discussion

Depending on the material system, one or the other of the two presented methods may be more suitable for the fabrication of a spacer layer for a vertical channel. The RIE procedure may inflict surface damage but can be the better choice when high control of the deposited film close to the structure is prioritized, or when there is a risk for the wet-etchant to react with other material films. The wet-etch procedure has the advantage of simplicity as no etch-mask is needed, giving a relatively higher reproducibility. Methods for reducing the faster etching of the deposited SiO_2 film close to the base of nanowire structure should be investigated to optimize the process for better RF performance.

CHAPTER 7

7 Perspectives

7.1 Technology and Evolution

It is not the strongest of the species that survives, nor the most intelligent that survives. It is the one that is the most adaptable to change.
- Charles Darwin

The civilized society along with modern medicine have taken us humans onto a path of evolutionary self-mastery. In our domain, the survival of the fittest equivalent is found in the competitive nature of the human ambition, today embodied in the form of states and corporations, where adaptability of trends and technological development is vital to stay ahead. New technological wonders seem to arise in the cross-roads of our divided ways of looking at things, where broader perspectives yield better solutions. With the global development, it has become prominent that we live on a planet with limited resources and technologies that offer ways to improve efficiency will become increasingly important.

The incredible progress in electronics has given those connected access to vast amounts of information and brought us all closer together. Technology is, however, still at a level where there is a clear divide between us and our gadgets. As technology becomes more sophisticated, the visible boundaries will likely blur out; in the digital world we have already experienced how technology can become extensions of our selves. Looking back at the technological development in the last 200 years, it seems indeed fair to state; only our imagination limits the possible.

Things do not change; we change.
- Henry David Thoreau

7.2 **Future Possibilities**

Small minds are concerned with the extraordinary, great minds with the ordinary.
- Blaise Pascal

The starting process in making great innovations often begins by looking at ordinary situations and figuring out how to improve that experience. The advanced technology available in electronics today makes it easier than ever before to create new ways of doing anything. In the modern world, every idea fiercely competes with those created by anyone and anywhere. Likewise, the world is open like never before to embrace a great idea with intense rapidity. Some of today's largest companies were founded not long ago by people having good ideas, the right amount of knowhow and not the least perfect timing. Products like the touchpad was not a meaningful concept until the power efficiency of electronics had come to the right level. Timing is always a crucial component in doing anything successful. Future market adoption of high speed materials in optimized structures may lead to new, yet unthought-of, imaginative ways of interacting with technology.

With my background from the nanoscience I have had the opportunity to dig deeper than the just the circulating buzz. Nanotechnology has a tremendous future potential, promising on a whole; smarter, greener, faster. The nanowire is a prime example of a nanodevice due to its dimensionality and the self-assembly of sorts. Nano is all about large surface/volume ratios. With its small dimensions and thereby large surface ratio, the nanowire is a good candidate for things such as nanosensors [69]. In the cross-roads of surfaces and electronics comes the field of bioelectricity where feature size and biologically preferential surfaces are a key component in probing living matter [70].

Nanotechnology has the capability to offer improvements in all areas of the natural sciences. The openness and sharing of knowledge within the research community may assist in technology exportation to unlikely destinations. A technology not turning the tables in one specific area might make all the difference in another.

All the forces in the world are not so powerful as an idea whose time has come.
- Victor Hugo

7.3 **At the end of the Road**

Still round the corner there may wait, a new road or a secret gate.
- J. R. R. Tolkien

Speculating about the distant future is a risky business; what can be said is that III-V materials and non-planar technologies are very likely an expanding base for

integrated electronics. These technologies might, however, only have a window of opportunity as something different altogether, such as graphene (or materials with similar properties), may eventually replace all types of electronics [71]. EU has recently launched a 9 billion SEK project, financing the collaborative research around graphene [72]. The era of III-Vs may, however, come to an end before graphene technology has come to dominate the market. Recent studies suggest that to maintain channel control for really ultra-scaled devices (sub 8 nm), materials with relatively heavy electron masses, like Si and opposite to III-V materials, may be preferential, as they have a lower tunneling probability through a thin stopping barrier [73]. This could either suggest a minimal feature size, keeping III-Vs, or that Si could actually reappear, if it would ever leave the high performance market scene.

A specific technology is only worth as much as the advantage in a possible application. If a technology is not cost efficient within existing demands, the solution is either not to pursue that technology, or to innovate new and different demands. An example of this is bendable electronics, which initially might not be competitive regarding performance, but instead has a unique feature. There are also technologies developed to target specific existing demands. Quantum computing is not thought of as a replacement of general computing, but rather as a specialized solution for parallel computations, promising an exceptional improvement [74]. On a whole, it may not be just one technology, like with Si-CMOS, but a diversity of technologies that best represent the most likely future scenario for integrated electronics.

All things that have a beginning must also have an end. At the end of the road, one would perhaps realize that it was all about the journey and as a scientist, one is not only permitted but, occasionally, encouraged to explore the bumps along the way. Before ending this introductory part, I would just like to mention one of my favorite childhood characters, Winnie the Pooh, whom I think is as good as a scientist role model as any other curious free-thinker.

Promise me you'll always remember: You're braver than you believe, and stronger than you seem, and smarter than you think.
- A. A. Milne

References

[1] Dave Evans, April 2011. [Online]. Available: http://www.cisco.com/web/about/ac79/docs/innov/IoT_IBSG_0411FINAL.pdf.

[2] C. P. Auth and J. D. Plummer, "Scaling Theory for Cylindrical, Fully-Depleted, Surrounding-Gate MOSFET's," *IEEE,* vol. 18, no. 2, pp. 74-76, February 1997.

[3] T. Tanaka, K. Tomioka, S. Hara, J. Motohisa, E. Sano and T. Fukui, "Vertical Surrounding Gate Transistors Using Single InAs Nanowires Grown on Si Substrates," *Applied Physics Express,* vol. 3, no. 2, p. 025003, January 2010.

[4] S. G. Ghalamestani, S. Johansson, B. M. Borg, E. Lind, K. A. Dick and a. L.-E. Wernersson, "Uniform and position-controlled InAs nanowires on 2 " Si substrates for transistor applications," *Nanotechnology,* vol. 23, no. 1, p. 015302, December 2011.

[5] A. Dodabalapur, "Organic and polymer transistors for electronics," *Materials Today,* vol. 9, no. 4, pp. 24-30, April 2006.

[6] G. E. Moore, "Cramming more components onto integrated circuits, Reprinted from Electronics," *Electronics,* vol. 38, no. 8, p. 114, April 1965.

[7] 2012. [Online]. Available: http://www.nvidia.com/content/PDF/kepler/NVIDIA-Kepler-GK110-Architecture-Whitepaper.pdf.

[8] S. Herculano-Houzel, "The human brain in numbers: a linearly scaled-up primate brain," *Frontiers in Human Neuroscience,* vol. 3, no. 31, pp. 1-11, November 2009.

[9] S. Koswatta, S. Koester and W. Haensch, "On the Possibility of Obtaining MOSFET-Like Performance and Sub-60-mV/dec Swing in 1-D Broken-Gap Tunnel Transistors," *IEEE Trans. Electron Devices,* vol. 57, no. 12, pp. 3222-3230, December 2010.

[10] S. M. Sze, Physics of Semicondcutor Devices, USA: John Wiley & Sons, Inc., 1969.

[11] G. T. Wright, "Threshold modelling of MOSFETs for CAD of CMOS-VLSI," *Electronics Letters,* vol. 21, no. 6, pp. 223-224, March 1985.

[12] A. G.Chynoweth, "Ionization Rates for Electrons and Holes in Silicon," *Physical Review,* vol. 109, no. 5, pp. 1537-1540, March 1958.

[13] J. H. Davies, The physics of low-dimensional semiconductors : an introduction, New York: Cambridge University Press, 1998.

[14] M. Lundstrom and Z. Ren, "Essential physics of carrier transport in nanoscale MOSFETs," *IEEE Trans. Electron Devices,* vol. 49, no. 1, pp. 133-141, January 2002.

[15] J. Wang and M. Lundstrom, "Ballistic transport in high electron mobility transistors," *IEEE Trans. Electron Devices,* vol. 50, no. 7, pp. 1604-1609, July 2003.

[16] A. Rahman, J. Guo, S. Datta and M. Lundstrom, "Theory of Ballistic Nanotransistors," *IEEE Trans. Electron Devices,* vol. 50, no. 9, pp. 1853-1864, September 2003.

[17] A. Khakifirooz, O. Nayfeh and D. Antoniadis, "A Simple Semiempirical Short-Channel MOSFET Current–Voltage Model Continuous Across All Regions of Operation and Employing Only Physical Parameters," *IEEE Trans. Electron Devices,* vol. 56, no. 8, pp. 1674-1680, August 2009.

[18] L. Wei, O. Mysore and D. Antoniadis, "Virtual-Source-Based Self-Consistent Current and Charge FET Models: From Ballistic to Drift-Diffusion Velocity-Saturation Operation," *IEEE Trans. on Electron Devices,* vol. 59, no. 5, pp. 1263-1271, May 2012.

[19] S. Roddaro, K. Nilsson, G. Astromskas, L. Samuelson, L.-E. Wernersson, O. Karlström and A. Wacker, "InAs nanowire metal-oxide-semiconductor capacitors," *Appl. Phys. Lett.*, vol. 92, p. 253509, June 2008.

[20] A. C. Ford, J. C. Ho, Y.-L. Chueh, Y.-C. Tseng, Z. Fan, J. Guo, J. Bokor and A. Javey, "Diameter-Dependent Electron Mobility of InAs Nanowires," *Nano Lett.*, vol. 9, no. 1, pp. 360-365, 2009.

[21] K.-M. Persson, B. G. Malm and L.-E. Wernersson, "Surface and core contribution to 1/f-noise in InAs nanowire," *Appl. Phys. Lett.*, vol. 103, p. 033508, July 2013.

[22] M. Shur, "Low ballistic mobility in submicron HEMTs," *IEEE Electron Device Lett.*, vol. 23, no. 9, pp. 511-513, September 2002.

[23] F. Elsholz, "Landauer Büttiker Formalism," Univesity of Rochester, Rochester, NY, USA, 2002.

[24] E. Lind, M. Persson, Y.-M. Niquet and L. -E. Wernersson, "Band Structure Effects on the Scaling Properties of [111] InAs Nanowire MOSFETs," *IEEE Trans. Electron Devices*, vol. 56, no. 2, pp. 201-205, February 2009.

[25] D. Kim, J. del Alamo, D. Antoniadis and B. Brar, "Extraction of virtual-source injection velocity in sub-100 nm III–V HFETs," in *IEEE International Electron Devices Meeting (IEDM)*, Baltimore, MD, USA, 2009.

[26] "International Technology Roadmap for Semiconductors," 2012. [Online]. Available: http://www.itrs.net/Links/2012ITRS/2012Tables/FEP_2012Tables.xlsx. [Accessed 30 January 2014].

[27] S. Johansson, M. Berg, K.-M. Persson and E. Lind, "A High-Frequency Transconductance Method for Characterization of High-k Border Traps in III-V MOSFETs," *IEEE Trans. Electron Devices*, vol. 60, no. 2, pp. 776-781, February 2013.

[28] W. Liu, Fundamentals of III-V Devices, New York: John Wiley & Sons, 1999.

[29] J. del Alamo, "Nanometer-scale InGaAs Field-Effect Transistors for THz and CMOS technologies," in *ESSCIRC*, Bucharest, Romania, 2013.

[30] J. Lin, X. Zhao, T. Yu, D. A. Antoniadis and J. A. d. Alamo, "A New Self-aligned Quantum-Well MOSFET Architecture Fabricated by a Scalable Tight-Pitch Process," in *IEEE International Electron Device Meeting (IEDM)*, Washincton DC, USA, 2013.

[31] S. H. Kim, M. Yokoyama, R. Nakane, O. Ichikawa, T. Osada, M. Hata, M. Takenaka and S. Takagi, "High performance sub-20-nm-channel-length extremely-thin body InAs-on-insulator Tri-gate MOSFETs with high short channel effect immunity and Vth tunability," in *IEEE International Electron Devices Meeting (IEDM)*, Washington DC, USA, 2013.

[32] D.-H. Kim, B. Brar and J. del Alamo, "fT = 688 GHz and fmax = 800 GHz in Lg = 40 nm In0.7Ga0.3As MHEMTs with gm_max > 2.7 mS/μm," in *IEEE International Electron Devices Meeting (IEDM)*, Washington DC, USA, 2011.

[33] T. -W. Kim, R. J. W. Hill, C. D. Young, D. Veksler, L. Morassi, S. Oktybrshky, J. Oh, C. Y. Kang, D. Kim, J. del Alamo, C. Hobbs, P. D. Kirsch and R. Jammy, "InAs quantum-well MOSFET (Lg = 100 nm) with record high gm, fT and fmax," in *Symposium on VLSI Technology (VLSIT)*, Honolulu, HI, USA, 2012.

[34] K. H. Yeo, S.-D. Suk, M. Li, Y.-y. Yeoh, K. H. Cho, K.-H. Hong, S. Yun, M. S. Lee, N. Cho, K. Lee, D. Hwang, B. Park, D.-W. Kim, D. Park and B.-i. Ryu, "Gate-All-Around (GAA) Twin Silicon Nanowire MOSFET (TSNWFET) with 15 nm Length Gate and 4 nm Radius Nanowires," in *International Electron Devices Meeting (2006)*, San Francisco, CA, USA, 2006.

[35] X. Jiang, Q. Xiong, S. Nam, F. Qian, Y. Li and C. M. Lieber, "InAs/InP radial nanowire heterostructures as high electron mobility devices," *Nano Lett.*, vol. 7, no. 10, pp. 3214-3218,

September 2007.

[36] K. Takei, S. Chuang, H. Fang, R. Kapadia, C.-H. Liu, J. Nah, H. S. Kim, E. Plis, S. Krishna, Y.-L. Chueh and A. Javey, "Benchmarking the performance of ultrathin body InAs-on-insulator transistors as function of body thickness," *Appl. Phys. Lett,* vol. 99, no. 10, p. 103507, September 2011.

[37] A. Dey, C. Thelander, E. Lind, K. A. Dick, B. Borg, M. Borgstrom, P. Nilsson and L. -E. Wernersson, "High-Performance InAs Nanowire MOSFETs," *IEEE Electron Device Letters,* vol. 33, no. 6, pp. 791-793, June 2012.

[38] K.-M. Persson, M. Berg, M. B. Borg, J. Wu, S. Johansson, J. Svensson, K. Jansson, E. Lind and a. L.-E. Wernersson, "Extrinsic and Intrinsic Performance of Vertical InAs Nanowire MOSFETs on Si Substrates," *IEEE Trans. Electron Devices,* vol. 60, no. 9, pp. 2761-2667, September 2013.

[39] J. J. Gu, X. W. Wang, H. Wu, J. Shao, A. T. Neal, M. J. Manfra, R. G. Gordon and P. D. Ye, "20-80 nm channel length gate-all-around nanowire MOSFETs with EOT=1.2nm and Lowest SS = 63mV/dec," in *IEEE International Electron Device Meeting,* San Francisco, CA, USA, 2012.

[40] K. Jansson, E. Lind and L. -E. Wernersson, "Performance Evaluation of III–V Nanowire Transistors," *IEEE Electron Devices,* vol. 59, no. 9, pp. 2375-2382, September 2012.

[41] C. Thelander, L. FrobergFroberg, C. Rehnstedt, L. Samuelson and L. -E. Wernersson, "Vertical Enhancement-Mode InAs Nanowire Field-Effect Transistor With 50-nm Wrap Gate," *IEEE Electron Device Letters,* vol. 29, no. 3, pp. 206-208, March 2008.

[42] K.-M. Persson, E. Lind, A. Dey, C. Thelander, H. Sjöland and L.-E. Wernersson, "Low-Frequency Noise in Vertical InAs," *IEEE Electron Dev. Lett.,* vol. 31, no. 5, pp. 428-430, May 2010.

[43] R. Lai, X. Mei, W. Deal, W. Yoshida, Y. -M. Kim, P. -H. Liu, J. Lee, J. Uyeda, V. Radisic, M. Lange, T. Gaier, L. Samoska and A. Fung, "Sub 50 nm InP HEMT Device with Fmax Greater than 1 THz," in *IEEE International Electron Devices Meeting (IEDM),* Washington DC, USA, 2007.

[44] M. Urteaga, R. Pierson, P. Rowell, V. Jain, E. Lobisser and M. Rodwell, "130nm InP DHBTs with ft >0.52THz and fmax >1.1THz," in *IEEE Device Research Conference (DRC),* Santa Barbara, CA, USA, 2011.

[45] M. Urteaga, M. Seo, J. Hacker, Z. Griffith, A. Young, R. Pierson, P. Rowell, A. Skalare, V. Jain, E. Lobisser and M. Rodwell, "InP HBTs for THz frequency integrated circuits," in *Compound Semiconductor Week (CSW) / International Conference on Indium Phosphide and Related Materials (IPRM),* Berlin, Germany, 2011.

[46] M. Egard, L. Ohlsson, M. Arlelid, K.-M. Persson, B. Borg, F. Lenrick, R. Wallenberg, E. Lind and L. -E. Wernersson, "High-Frequency Performance of Self-Aligned Gate-Last Surface Channel In0.53Ga0.47As MOSFET," *IEEE Electron Device Lett.,* vol. 33, no. 3, pp. 369-371, March 2012.

[47] M. V. Haartman and M. Östling, Low-Frequency Noise in Advanced MOS Devices, Dordrecht: Springer, 2007.

[48] S. Monaghan, P. Hurley, K. Cherkaoui, M. Negara and A. Schenk, "Determination of electron effective mass and electron affinity in HfO2 using MOS and MOSFET structures," *Solid-State Electronics,* vol. 53, no. 4, pp. 438-444, April 2009.

[49] N. Li, E. S. Harmon, J. Hyland, D. B. Salzman, T. P. Ma, Y. Xuan and P. D.Ye, "Properties of InAs metal-oxide-semiconductor structures with atomic-layer-deposited Al2O3 Dielectric," *Appl. Phys. Lett.,* vol. 92, p. 143507, March 2008.

[50] M. Tacano, "Hooge Fluctuation Parameter of Semiconductor Microstructures," *IEEE Trans.*

Electron Device, vol. 40, no. 11, pp. 2060-2064, 1993.

[51] C. Wei, Y. Jiang, Y.-Z. Xiong, X. Zhou, N. Singh, S. Rustagi, G. Q. Lo and D.-L. Kwong, "Impact of Gate Electrodes on 1/f Noise of Gate-All-Around Silicon Nanowire Transistors," *IEEE Electron Device Lett.*, vol. 30, no. 10, pp. 1081-1083, October 2009.

[52] B. Rajamohanan, I. Ok, S. Mujumdar, C. Hobbs, P. Majhi, R. Jammy and S. Datta, "Correlated Flicker Noise and Hole Mobility Characteristics of (110)/<110> Uniaxially Strained SiGe FINFETs," *IEEE Electron Device Lett.*, vol. 33, no. 9, pp. 1237-1239, 2012.

[53] M. R. Gao and X. P. A. Sakr, "Temperature dependence of the low frequency noise in indium arsenide nanowire," *Appl. Phys. Lett.*, vol. 93, p. 203503, November 2008.

[54] M. Tacano, M. Ando, I. Shibasaki, S. Hashiguchi, J. Sikula and T. Matsui, "Dependence of Hooge parameter of InAs heterostructure on temperature," *Microelectronics Reliability*, vol. 40, no. 11, pp. 1921-1924, November 2000.

[55] M. Shi, J. Saint-Martin, A. Bournel, D. Querlioz, N. Wichmann, S. Bollaert, F. Danneville and P. Dollfus, "Monte Carlo analysis of the dynamic behavior of III–V MOSFETs for low-noise RF applications," *Solid-State Electronics*, vol. 87, pp. 51-57, September 2013.

[56] B. Razavi, RF Microelectronics, 2nd ed., B. Goodwin, J. Fuller and G. Breeze, Eds., New York: Pearson Education, 2012.

[57] D.-H. Kim, J. A. d. Alamo, D. A. Antoniadis and B. Brar, "Extraction of virtual-source injection velocity in sub-100 nm III-V HFETs," in *IEEE International Electron Devices Meeting (IEDM)*, Baltimore, MD, US, 2009.

[58] N. Lindblad, "Simulations of III-V NWFET Double-Balanced Gilbert Cells with an Improved Noise Model," Lund, 2013.

[59] D.-H. Kim and J.-S. Rieh, "CMOS 138 GHz low-power active mixer with branch-line coupler," *Electron. Lett.*, vol. 48, no. 10, pp. 554-555, May 2012.

[60] D.-H. Kim and J.-S. Rieh, "A 135 GHz Differential Active Star Mixer in SiGe BiCMOS Technology," *IEEE Microwave and Wireless Components Lett.*, vol. 22, no. 8, pp. 409-411, August 2012.

[61] O. nac, B. Cetinoneri, M. Uzunkol, Y. Atesal and G. Rebeiz, "Millimeter-Wave and THz Circuits in 45-nm SOI CMOS," in *Compound Semiconductor Integrated Circuit Symposium (CSICS)*, Waikoloa, HI, USA, 2011.

[62] N. Ghazali, H. Soetedjo, N. A. Ngah, A. Yusof, A. Dolah and M. Yahya, "DOE study on etching rate of silicon nitride (Si3N4) layer via RIE nitride etching process," in *IEEE International Conference on Semiconductor Electronics (ICSE)*, Johor Bahru, Malaysia, 2008.

[63] R. Knizikevičius, "Influence of activated polymer on the etching rate of SiO2 in CF4 + H2 plasma," *Microelectronic Engineering*, vol. 86, no. 1, pp. 55-58, January 2009.

[64] Q.-y. Tong, Q. Gan, G. Fountain, G. Hudson and P. Enquist, "Low-temperature bonding of silicon-oxide-covered wafers using diluted HF etching," *Appl. Phys. Lett.*, vol. 85, no. 14, pp. 2762-2764, October 2004.

[65] G. Gautier, L. Ventura, T. Pordié, R. Rogel and R. Jérisian, "Fabrication of deep single trenches from N-type macroporous silicon," *Thin Solid Films*, vol. 487, no. 1-2, pp. 283-287, September 2005.

[66] K. Vanheusden and A. Stesmans, "Chemical etch rates in HF solutions as a function of thickness of thermal SiO2 and buried SiO2 formed by oxygen implantation," *Appl. Phys*, vol. 69, no. 9, pp. 6656-6664, May 1991.

[67] H. Nakanishi, T. Nishimoto, R. Nakamura, A. Yotsumoto and S. Shoji, "Studies on SiO2-SiO2 bonding with hydrofluoric acid-room temperature and low stress bonding technique for MEMS," in *The Eleventh Annual International Workshop on Micro Electro Mechanical Systems (MEMS)*, Heidelberg, Germany, 1998.

[68] H. Jansen, M. d. Boer, R. Wiegerink, N. Tas, E. Smulders, C. Neagu and a. M. Elwenspoek, “RIE lag in high aspect ratio trench etching of silicon,” *Microelectronic Engineering,* vol. 35, pp. 45-50, 1197.

[69] C. M. L. Fernando Patolsky, “Nanowire nanosensors,” *Materials Today,* vol. 8, no. 4, pp. 20-28, April 2005.

[70] K. Adolfsson, H. Persson, J. Wallentin, S. Oredsson, L. Samuelson, J. O. Tegenfeldt, M. T. Borgström and C. N. Prinz, “Fluorescent Nanowire Heterostructures as a Versatile Tool for Biology Applications,” *Nano Lett.,* vol. 13, no. 10, pp. 4728-4732, August 2013.

[71] K. S. Novoselov, V. I. Falko, L. Colombo, P. R. Gellert, M. G. Schwab and K. Kim, “A roadmap for graphene,” *Nature,* vol. 490, no. 7419, pp. 192-200, October 2012.

[72] “Graphene Flagship,” [Online]. Available: http://graphene-flagship.eu/. [Accessed 12 February 2014].

[73] S. Mehrotra, S. Kim, T. Kubis, M. Povolotskyi, M. Lundstrom and G. Klimeck, “Engineering Nanowire n-MOSFETs at Lg < 8nm,” *IEEE Trans. Electron Devices,* vol. 60, no. 7, pp. 2171-2177, July 2013.

[74] T. D. Ladd, F. Jelezko, R. Laflamme, Y. Nakamura, C. Monroe and J. L. O'Brien, “Quantum computers,” *Nature,* vol. 464, no. 7285, pp. 45-53, March 2010.

[75] R. K. a. M. Lundstrom, “Notes on Fermi-Dirac Integrals: 3rd Edition,” September 2008. [Online]. Available: https://nanohub.org/resources/5475.

[76] J. T. Robinson, M. Jorgolli, A. K. Shalek, M.-H. Yoon, R. S. Gertner and H. Park, “Vertical nanowire electrode arrays as a scalable platform for intracellular interfacing to neuronal circuits,” *Nature Nanotech.,* vol. 7, no. 3, pp. 180-184, January 2012.

List of Acronyms

AC	Alternating Current
ALD	Atomic Layer Deposition
BHF	Buffered Hydrofluoric Acid
BJT	Bipolar-Junction-Transistor
BW	Bandwidth
CMOS	Complementary Metal-Oxide-Semiconductor Field Effect Transistor
DC	Direct Current
DHBT	Double-Hetero-Bipolar-junction Transistor
DIBL	Drain Induced Barrier Lowering
DSB	Dual Sideband
EBL	Electron Beam Lithography
FET	Field Effect Transistor
HBT	Hetero-Bipolar-junction Transistor
HEMT	High Electron Mobility Transistor
HF	Hydrofluoric Acid
IF	Intermediate Frequency
III-V	Compound semiconductor from element groups 3 and 5
IIP3	Input 3rd order Intercept Point
IMP	Intermodulation Product
ITRS	International Technology Roadmap for Semiconductors
IMIO	Impact Ionization
LFN	Low Frequency Noise
LO	Local Oscillator
MOSFET	Metal-Oxide-Semiconductor Field Effect Transistor
NF	Noise Figure
NW	Nanowire
NW-FET	Nanowire Field Effect Transistor
P1dB	1 dB compression point
PE-CVD	Plasma-Enhanced Chemical Vapor Deposition
QW-FET	Quantum Well Field Effect Transistor
RF	Radio Frequency
RTS	Random Telegraph Signal
SCE	Short-Channel-Effects
SEM	Scanning Electron Microscopy
SNR	Signal-to-Noise Ratio
SS	Sub-threshold Swing
SSB	Single Sideband
VS	Virtual Source

Appendix 1

A.1 Electron Transport

The energy-state of a particle, *E*, which is part momentum and part potential, can be described by Schrödinger's equation [10]:

$$\left[-\frac{h^2}{2m(2\pi)^2} + V(x)\right] = E(x)\Psi(x). \qquad \text{(A1)}$$

In Eq. A.1, Ψ is the particle wave function, h is the Planck's constant, m is the electron mass, and V is the potential energy. By knowing the shape of a potential, it is thus possible to calculate the possible energy states.

To calculate the electron density in a semiconductor, it can be necessary to numerically solve the Fermi-Dirac integral, $\mathcal{F}_k$, with $k = ½$ [75]:

$$\mathcal{F}_k(\eta) = \frac{1}{\Gamma(k+1)}\int_0^\eta \frac{x^k}{\exp(k-x)+1}\,dx. \qquad \text{(A2)}$$

In Eq. A2, η is the relative normalized energy state and Γ is the gamma function.

A.2 Mixer Simulations

To evaluate circuit performance of theoretically enhanced InAs nanowire FET structures, a Verilog-A model was constructed using a compact virtual source model fitted to measured data. A RC-network of extrinsic parasitic elements was added using estimated values for realistic layouts [40]. The implemented active mixer that was used for circuit simulations is presented in Fig. A.1, showing the NW-FETs, M_1, M_2 and M_3, symbolized with schematic nanowires. The name and location of the different input and output voltages are indicated along with labeling of the different components.

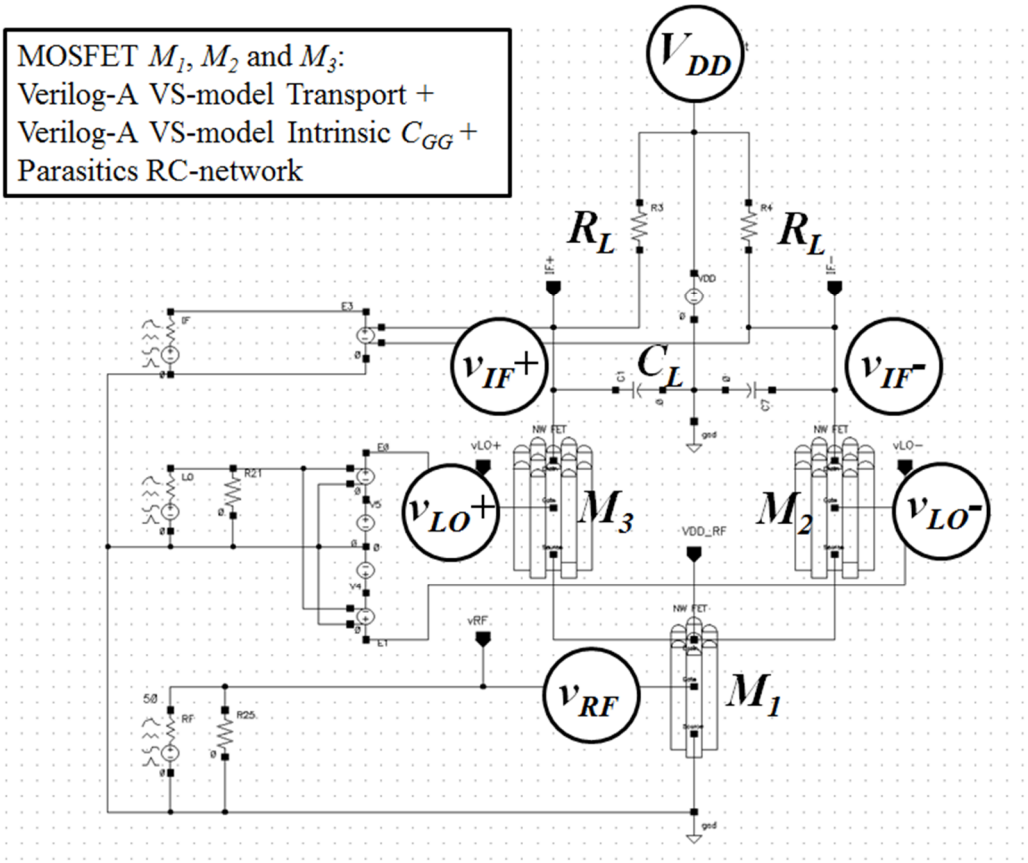

Fig. A.1 *Mixer circuit implemented in Cadence using a Verilog-A model for intrinsic transistor transport calculations as well as the intrinsic capacitances. The extrinsic capacitances are modelled with a small signal model using values determined for realistic layouts. The different load elements and input and output voltages indicated where the parameters have the same meaning as in Fig. 5.3.*