LUND UNIVERSITY

Electrical Characterisation of III-V Nanowire MOSFETs

Hellenbrand, Markus

2020

Document Version:
Publisher's PDF, also known as Version of record

Link to publication

Citation for published version (APA):
Hellenbrand, M. (2020). *Electrical Characterisation of III-V Nanowire MOSFETs*. [Doctoral Thesis (compilation), Lund University]. Department of Electrical and Information Technology, Lund University.

Total number of authors:
1

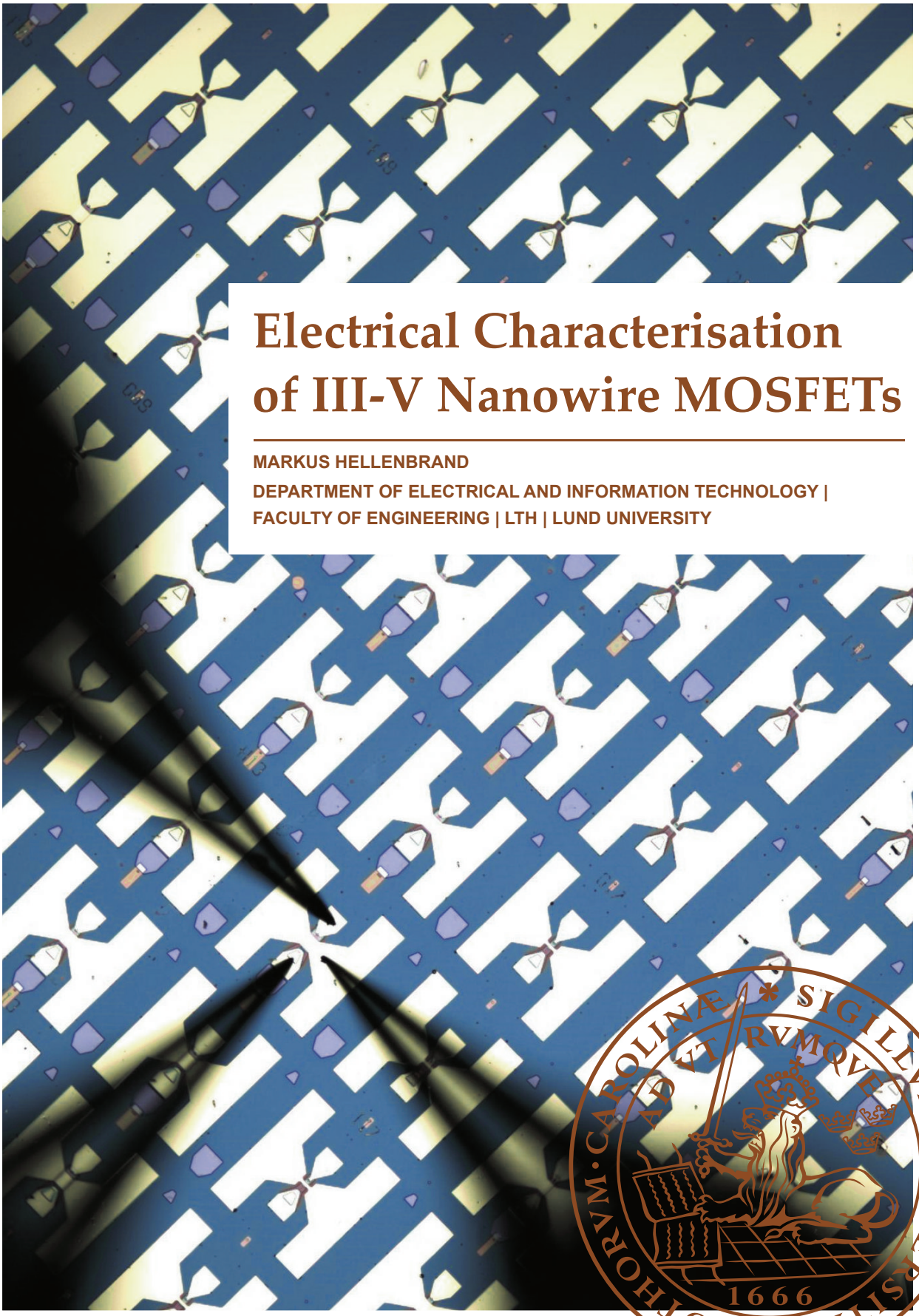
General rights
Unless other specific re-use rights are stated the following general rights apply:
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.
• Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
• You may not further distribute the material or use it for any profit-making activity or commercial gain
• You may freely distribute the URL identifying the publication in the public portal

Read more about Creative commons licenses: https://creativecommons.org/licenses/

Take down policy
If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

LUND UNIVERSITY

PO Box 117
221 00 Lund
+46 46-222 00 00



Electrical Characterisation of III-V Nanowire MOSFETs

MARKUS HELLENBRAND

DEPARTMENT OF ELECTRICAL AND INFORMATION TECHNOLOGY |
FACULTY OF ENGINEERING | LTH | LUND UNIVERSITY

Electrical Characterisation of III-V Nanowire MOSFETs

Doctoral Thesis

Markus Hellenbrand

LUND
UNIVERSITY

Department of Electrical and
Information Technology
Lund, May 2020

Academic thesis for the degree of Doctor of Philosophy, which, by due permission of the Faculty of Engineering at Lund University, will be publicly defended on Friday, 12 June, 2020, at 9:15 a.m. in lecture hall E:1406, Department of Electrical and Information Technology, Ole Römers Väg 3, 223 63 Lund, Sweden. The thesis will be defended in English.

The Faculty Opponent will be Professor Tibor Grasser, TU Wien, Austria.

Organisation:
LUND UNIVERSITY
Department of Electrical and Information Technology
Ole Römers Väg 3
223 63 Lund
Sweden

Author:
Markus Hellenbrand

Document Type:
DOCTORAL THESIS

Date of Issue:
May 2020

Sponsoring Organisation(s):
Swedish Foundation for Strategic Research (SSF)
Swedish Research Council (VR)
European Union (EU)

Title:
Electrical Characterisation of III-V Nanowire MOSFETs

Abstract:
The ever increasing demand for faster and more energy-efficient electrical computation and communication presents severe challenges for the semiconductor industry and particularly for the metal-oxide-semiconductor field-effect transistor (MOSFET), which is the workhorse of modern electronics. III-V materials exhibit higher carrier mobilities than the most commonly used MOSFET material Si so that the realisation of III-V MOSFETs can enable higher operation speeds and lower drive voltages than that which is possible in Si electronics. A lowering of the transistor drive voltage can be further facilitated by employing gate-all-around nanowire geometries or novel operation principles. However, III-V materials bring about their own challenges related to material quality and to the quality of the gate oxide on top of a III-V MOSFET channel.

This thesis presents detailed electrical characterisations of two types of (vertical) III-V nanowire transistors: MOSFETs based on conventional thermionic emission; and Tunnel FETs, which utilise quantum-mechanical tunnelling instead to control the device current and reach inverse subthreshold slopes below the thermal limit of 60 mV/decade. Transistor characterisations span over fourteen orders of magnitude in frequency/time constants and temperatures from 11 K to 370 K.

The first part of the thesis focusses on the characterisation of electrically active material defects ('traps') related to the gate stack. Low-frequency noise measurements yielded border trap densities of 10^{18} to 10^{20} cm^{-3} eV^{-1} and hysteresis measurements yielded effective trap densities – projected to the oxide/semiconductor interface – of 2×10^{12} to 3×10^{13} cm^{-2} eV^{-1}. Random telegraph noise measurements revealed that individual oxide traps can locally shift the channel energy bands by a few meV and that such defects can be located at energies from inside the semiconductor band gap all the way into the conduction band.

Small-signal radio frequency (RF) measurements revealed that parts of the wide oxide trap distribution can still interact with carriers in the MOSFET channel at gigahertz frequencies. This causes frequency hystereses in the small-signal transconductance and capacitances and can decrease the RF gains by a few decibels. A comprehensive small-signal model was developed, which takes into account these dispersions, and the model was applied to guide improvements of the physical structure of vertical RF MOSFETs. This resulted in values for the cutoff frequency f_T and the maximum oscillation frequency f_{max} of about 150 GHz in vertical III-V nanowire MOSFETs.

Bias temperature instability measurements and the integration of (lateral) III-V nanowire MOSFETs in a back end of line process were carried out as complements to the main focus of this thesis. The results of this thesis provide a broad perspective of the properties of gate oxide traps and of the RF performance of III-V nanowire transistors and can act as guidelines for further improvement and finally the integration of III-V nanowire MOSFETs in circuits.

Keywords:
MOSFET, TFET, III-V, Nanowire, Hysteresis, Low-Frequency Noise, Random Telegraph Noise, cryogenic, Reliability, Radio Frequency, Small-Signal Model.

Classification System and/or Index Terms
Electronic Engineering, Nano Technology

Language:
English

Supplementary Bibliographical Information:
–

ISBN (printed):
978-91-7895-519-0

Key title and ISSN:
Series of Licentiate and Doctoral Theses; 1654-790X, No. 130

ISBN (digital):
978-91-7895-520-6

Recipient's Notes:

Number of Pages:
156

Price:

Security Classification:
Unclassified

General Permissions:
I, the undersigned, being the copyright owner and author of the above-mentioned thesis and its abstract, hereby grant to all reference sources permission to publish and disseminate said abstract.

Signature:  *Date:* 11 May 2020

Electrical Characterisation of III-V Nanowire MOSFETs

Doctoral Thesis

Markus Hellenbrand

LUND
UNIVERSITY

Department of Electrical and
Information Technology

Lund, May 2020

Markus Hellenbrand
Department of Electrical and Information Technology
Lund University
Ole Römers Väg 3, 223 63 Lund, Sweden

Series of Licentiate and Doctoral Theses
ISSN 1654-790X, No. 130
ISBN 978-91-7895-519-0 (printed)
ISBN 978-91-7895-520-6 (digital)

© 2020 Markus Hellenbrand
This thesis is typeset using LaTeX 2ε with the body text in Palatino and Goudy Initials, headings in Helvetica, text in figures in Arial, and the body text in Appendix B in Computer Modern.

Frontispiece: Micrograph of a transistor sample with probe tips in contact.

Printed by Tryckeriet i E-huset, Lund University, Lund, Sweden.

No part of this thesis may be reproduced or transmitted in any form or by any means without written permission from the author. Distribution of the original thesis in full, however, is permitted without restriction.

"The Road goes ever on and on
Down from the door where it began.
Now far ahead the Road has gone,
And I must follow, if I can,
Pursuing it with weary feet,
Until it joins some larger way,
Where many paths and errands meet.
And whither then? I cannot say."

The Fellowship of the Ring, J. R. R. Tolkien

Abstract

The ever increasing demand for faster and more energy-efficient electrical computation and communication presents severe challenges for the semiconductor industry and particularly for the metal-oxide-semiconductor field-effect transistor (MOSFET), which is the workhorse of modern electronics. III-V materials exhibit higher carrier mobilities than the most commonly used MOSFET material Si so that the realisation of III-V MOSFETs can enable higher operation speeds and lower drive voltages than that which is possible in Si electronics. A lowering of the transistor drive voltage can be further facilitated by employing gate-all-around nanowire geometries or novel operation principles. However, III-V materials bring about their own challenges related to material quality and to the quality of the gate oxide on top of a III-V MOSFET channel.

This thesis presents detailed electrical characterisations of two types of (vertical) III-V nanowire transistors: MOSFETs based on conventional thermionic emission; and Tunnel FETs, which utilise quantum-mechanical tunnelling instead to control the device current and reach inverse subthreshold slopes below the thermal limit of 60 mV/decade. Transistor characterisations span over fourteen orders of magnitude in frequency/time constants and temperatures from 11 K to 370 K.

The first part of the thesis focusses on the characterisation of electrically active material defects ('traps') related to the gate stack. Low-frequency noise measurements yielded border trap densities of 10^{18} to 10^{20} $\mathrm{cm}^{-3}\,\mathrm{eV}^{-1}$ and hysteresis measurements yielded effective trap densities – projected to the oxide/semiconductor interface – of 2×10^{12} to 3×10^{13} $\mathrm{cm}^{-2}\,\mathrm{eV}^{-1}$. Random

telegraph noise measurements revealed that individual oxide traps can locally shift the channel energy bands by a few millielectronvolts and that such defects can be located at energies from inside the semiconductor band gap all the way into the conduction band.

Small-signal radio frequency (RF) measurements revealed that parts of the wide oxide trap distribution can still interact with carriers in the MOSFET channel at gigahertz frequencies. This causes frequency hystereses in the small-signal transconductance and capacitances and can decrease the RF gains by a few decibels. A comprehensive small-signal model was developed, which takes into account these dispersions, and the model was applied to guide improvements of the physical structure of vertical RF MOSFETs. This resulted in values for the cutoff frequency f_{T} and the maximum oscillation frequency f_{max} of about 150 GHz in vertical III-V nanowire MOSFETs.

Bias temperature instability measurements and the integration of (lateral) III-V nanowire MOSFETs in a back end of line process were carried out as complements to the main focus of this thesis. The results of this thesis provide a broad perspective of the properties of gate oxide traps and of the RF performance of III-V nanowire transistors and can act as guidelines for further improvement and finally the integration of III-V nanowire MOSFETs in circuits.

Popular Science Summary

NIMAGINABLE quantities of electrical data are produced, handled and stored all around the globe. It has been estimated that by the year 2030, with the current energy efficiency of computers and networks, the energy demand related to information and communications technology (ICT) could amount to up to 20 % of the world's total energy consumption. This is in part due to the still increasing spread of ICT to more parts of the world. It is further accelerated by emerging technologies such as the Internet of Things (IoT), Neuromorphic Computing, or Quantum Computing. In view of this, it is clear that a drastic improvement of energy efficiency in computing technology is highly necessary to cover the endless demand for ICT and maybe even to enable its more sustainable employment.

One approach to tackling parts of this enormous problem is the use of *III-V Nanowire Transistors* for advanced electronic circuits. A transistor is an electronic device with three contacts, source, gate, and drain. The gate contact does not conduct any current itself, but controls how much current can flow between source and drain. This can be thought of as similar to a dimmable light switch, which regulates the brightness of the lamps without glowing itself. In a well designed transistor, the highest possible current can be more than 100 000 times as high as the lowest current. Based on this three-contact functionality, transistors are the most fundamental components of ICT systems and act as the 'brain cells' of computers. In this capacity, several billion transistors make up the electronic 'brains' of each of our smartphones and computers and handle all the different tasks that we require of them. In a second capacity, transistors can act as amplifiers and as such, are crucial

for sending and receiving information via internet and telephone communication. An accessible example for this is audio amplifiers, which amplify the small electrical signals containing the audio information so that they can power speakers or headphones. For such communication functionalities, high operation speed is especially important.

In the vast majority of transistors in the world, the *channel*, which conducts the current between source and drain, is made out of silicon, and the *gate oxide*, by which the gate contact controls the amount of current in the channel, is made out of hafnium oxide. Silicon is widely used, because it is available in large quantities and it is relatively easy to use in fabrication. Unfortunately, current cannot flow very fast in a silicon channel, which imposes severe limitations on transistors in terms of energy efficiency and speed. The prediction, in the beginning, of rising energy demand is partly because of this limited channel conductivity. This is where III-V materials – compounds of elements from groups III and V in the periodic table – come into play. Current can flow much faster in these materials, which means that a transistor made out of III-V materials requires less energy to achieve the same performance as a silicon transistor. Alternatively, with the same power consumption, a III-V transistor can operate at a higher speed than a silicon transistor. Further improvements can be achieved by changing the geometry of a transistor, e.g. by making it into the shape of a tiny wire – a 'nanowire' – or by manipulating its physical operation principles.

Unfortunately, with III-V materials it is much more challenging than with silicon to fabricate transistors with a high material quality. It is thus not straightforward to use the superior transport properties of III-V materials to full capacity. For this thesis, I measured the performance of transistors made out of III-V materials and investigated, how their performance is affected by imperfections in the gate oxide. Measurements were carried out at frequencies from 100 µHz to 67 GHz, where the latter is 670 million million(!) times as fast as the former. For comparison, the processor of a normal home computer operates at about 2–4 GHz and typical WiFi links operate at 2.4 GHz or 5 GHz. Furthermore, transistors were measured at temperatures from −260 °C to investigate certain physical mechanisms and to explore capabilities for e.g. quantum computing or space applications, up to temperatures of 100 °C to comply with specifications for industrial applications. The thesis thus provides a broad perspective of transistors for applications from ultra-low power IoT and quantum applications to high-speed gigahertz applications for wireless communication.

Acknowledgments

"It's a dangerous business, going out of your door. You step into the Road, and if you don't keep your feet, there is no knowing where you might be swept off to."

The Fellowship of the Ring, J. R. R. Tolkien

On the long, yet in the end 'high speed' Road of my PhD journey, I have had the privilege of travelling with many wonderful people and here I express my heartfelt gratitude. First of all, thank you *Lars-Erik* and *Erik*, my supervisors, for pointing the way and guiding me through these years. *Lars-Erik*, thank you for opening the door and allowing me to pursue PhD studies in the *Nanoelectronics* group. I learned a great many things from you, not only about hard science, but also about 'academic diplomacy', the general dealings of research and academia, and especially about positivity. *Erik*, thank you for your endless patience with my unending torrents of questions, and thank you for providing a 'sounding board' for my views on 'how to do science'. I feel very lucky to have had you as my supervisors and I am certain that I will often be taking leaves out of your books in my future scientific endeavours.

Thank you *Johannes* for your enthusiasm about so many things and for all the very practical support. Among my PhD colleagues, a special thanks goes to *Elvedin*, with whom I worked the most in the beginning of my journey and who showed me a lot of 'how things are done' as a PhD student. I thank all my colleagues in our group with whom I had the privilege and the pleasure to work (in loose chronological, but mostly random order) – *Aein*, *Guntrade*, *Cezar*, *Martin*, *Kristofer*, *Jun*, *Karl-Magnus*, *Olli-Pekka*, *Sebastian*, *Lars*, *Fredrik*, *Mattias*, *Adam*, *Stefan*, *Abinaya*, *Lasse*, *Anton*, *Robin*, *Saketh*, *Navya*, *Gautham*, *Heera*, *Zhongyunshen*, *Patrik* – and all colleagues at the department. Thank you all for your fellowship and friendship and for the great working atmosphere. You are all truly special people and I wish I knew you much better than I do.

During the later stage of my PhD I had the fantastic opportunity to work with *Paul* and *Enrico* from the Tyndall Institute in Cork, and *Vamsi* and *Jacopo* from IMEC in Leuven. Thank you for many enlightening discussions and the great (ongoing) collaborations to unravel the mysteries of a few nanometres of oxide.

Thank you *Daniel*, *Anne*, *Pia*, *Dorthe*, both *Elisabeth*s, *Linda*, both *Erik*s, *Bertil*, *Göran*, *Andreas*, *Sirvan*, *Martin*, *Josef* – without your administrative and technical support we could not focus on science as much as we do.

Although it is not always apparent as a PhD student, there is a life outside of university. For framing this, first and foremost I thank my oldest and best friends *Kevin* and *Sudha* in 'good old' Heidelberg and in here in Lund, respectively, and also my 'a little bit newer' closest friends *Ali* and *Kush*. Thank you for going out the door with me, for keeping me company on the Road, and for making sure I did not get swept off.

Most of my time outside of university I probably spent at *Domino*, the catholic student community in Lund. I met so many wonderful people here that it is impossible to name all of you. Still I should like to thank some of you by name, whom I have known the best and for the longest – *Magda* and *Magnus*, *Milan*, both *Simon*s, *Leonardo*, *Märta* and *Jacob*, *Daniel*, *Imran*, *Kasia*, *Imelda*, *Camila*, *Leny*, *Anna*, *Anna Caterina*, and 'in the old days' *Elin*, *Victor*, *Alexander* and *Cassandra*, *Karin*, *Constantine*, *Ellen*, and *Evelina*. Already the list is quite long, so I hope that everyone missing understands my predicament. At least I maintain that while "I don't know half of you as well as I should like",[†] the second part of this famous quote is certainly not true. Thank you as well for your fellowship and friendship and for keeping me company on the Road. I thank the Dominican brothers in Lund for a spiritual as well as a very physical home for many years; *Björn*, *Paul-Dominique*, *Pierre-André*, *Hani*, *Henrik*, and *Johan*.

Now far ahead the Road has gone indeed, and here at the end, not of all things, but of this very Road, I am glad that my family is with me. *Mama und Papa*, *Anne und Stefan*, *Oma und Opa*, thank you for your unconditional love and kindness, for your help and support, for the very ground that my Road has ever been built upon, and thank you for providing the Grey Havens to which I can always return; because in the end, I am "only quite a little fellow in a wide world after all!"[‡]

Lund, April 2020

[†]The Fellowship of the Ring, J. R. R. Tolkien.

[‡]The Hobbit, J. R. R. Tolkien.

Contents

Preface

This thesis is the culmination of more than five years of work in the *Nanoelectronics* group at Lund University and presents detailed characterisations of vertical III-V nanowire metal-oxide-semiconductor field-effect transistors (MOSFETs) for low power and high frequency applications. The analyses span over fourteen orders of magnitude in frequency and were supported by many different researchers. The work was supervised by Professor *Lars-Erik Wernersson* and Professor *Erik Lind*.

STRUCTURE OF THE THESIS

The thesis is divided into three main parts. First, a survey of the part of the field of nanoelectronics related to this thesis will present some fundamental principles underlying the operation of III-V nanowire MOSFETs. This will put the studied transistors into perspective relative to the current state of academic research and industry by comparing some of the key metrics of transistor operation. The second part will present the conducted work pertaining to effects of gate oxide defects in DC operation and at low frequencies, which in the context of this thesis is below about 10 kHz. The third part will summarise radio frequency characterisations of nanowire MOSFETs from 10 MHz to 67 GHz. The analysis of gate oxide defects reaches over into this chapter, but more critical properties are investigated as well.

- **INTRODUCTION**
  The main body of the thesis consists of the publications appended in the back. The Introduction provides a broader and more comprehensive view than the very focussed publications and ties their work together. The Introduction is intended to be comprehensible for aspiring researchers with a Master's degree in physics or a related subject.

- **APPENDICES**

  A Measurement Setups
  Appendix A provides a description of the measurement of low-frequency noise and scattering parameters.

  B Small-Signal Analysis in Detail
  Appendix B provides a detailed treatise of the measurement and modelling of small-signal parameters beyond the content of this thesis.

- **PAPERS**
  The papers forming the main body of the thesis are reproduced in the back and listed in the following.

It should be noted that while the characterisations in this thesis were carried out in close interaction with the design of processing schemes, processing itself is not a part of this thesis. Samples were provided by E. Memišević, O.-P. Kilpi, A. Jönsson, A. Krishnaraja, C. Zota, and F. Lindelöw.

INCLUDED PAPERS

The following papers form the main body of this thesis and the respective published or draft versions are appended in the back.

Paper I: S. Netsu, M. Hellenbrand, C. B. Zota, Y. Miyamoto, and E. Lind, "A Method for Determining Trap Distributions of Specific Channel Surfaces in InGaAs Tri-Gate MOSFETs", *IEEE Journal of the Electron Devices Society*, vol. 6, pp. 408–412, Feb. 2018, doi: 10.1109/JEDS.2018.2806487.

► *I supervised and helped with the measurements and helped to write the paper.*

Paper II: M. Hellenbrand, O.-P. Kilpi, J. Svensson, E. Lind, and L.-E. Wernersson, "Low-frequency noise in nanowire and planar III-V MOSFETs", *Microelectronic Engineering*, vol. 215, pp. 110986, May 2019, doi: 10.1016/j.mee.2019.110986.

► *I carried out all measurements and analyses and wrote the paper.*

Paper III: M. Hellenbrand, E. Memišević, M. Berg, O.-P. Kilpi, J. Svensson, and L.-E. Wernersson, "Low-Frequency Noise in III–V Nanowire TFETs and MOSFETs", *IEEE Electron Device Letters*, vol. 38, no. 11, pp. 1520–1523, Sep. 2017, doi: 10.1109/LED.2017.2757538.

► *I carried out all measurements and analyses and wrote the paper.*

Paper IV: E. Memisevic, M. Hellenbrand, E. Lind, A. R. Persson, S. Sant, A. Schenk, J. Svensson, R. Wallenberg, and L.-E. Wernersson, "Individual Defects in InAs/InGaAsSb/GaSb Nanowire Tunnel Field-Effect Transistors Operating below 60 mV/decade", *Nanoletters*, vol. 17, no. 7, pp. 4373–4380, Jun. 2017, doi: 10.1021/acs.nanolett.7b01455.

▶ *I carried out the measurements and analyses pertaining to the effects of gate oxide defects and wrote the corresponding part of the paper.*

Paper V: M. Hellenbrand, E. Memišević, J. Svensson, E. Lind, and L.-E. Wernersson, "Random Telegraph Signal Noise in Tunneling Field-Effect Transistors with S below 60 mV/decade", *2017 47th European Solid-State Device Research Conference (ESSDERC)*, pp. 38–41, Sep. 2017, doi: 10.1109/ESSDERC.2017.8066586.

▶ *I carried out all measurements and analyses, wrote the paper, and presented it at the conference.*

Paper VI: M. Hellenbrand, E. Lind, O.-P. Kilpi, and L.-E. Wernersson, "Effects of traps in the gate stack on the small-signal RF response of III-V nanowire MOSFETs", *Solid-State Electronics*, under review with minor revisions pending, May 2020.

▶ *I carried out the majority of the measurements, all analyses, developed and implemented the model, and wrote the paper.*

Paper VII: O.-P. Kilpi, M. Hellenbrand, J. Svensson, E. Lind, and L.-E. Wernersson, "Vertical nanowire III–V MOSFETs with improved high-frequency gain", *Electronics Letters*, Article in Press, Apr. 2020, doi: 10.1049/el.2020.0266.

▶ *I carried out part of the high frequency measurements, analysed the small-signal model, and wrote part of the paper.*

Paper VIII: M. Hellenbrand, E. Memisevic, J. Svensson, A. Krishnaraja, E. Lind, and L.-E. Wernersson, "Capacitance Measurements in Vertical III–V Nanowire TFETs", *IEEE Electron Device Letters*, vol. 39, no. 7, pp. 943–946, May 2018, doi: 10.1109/LED.2018.2833168.

▶ *I carried out all measurements and analyses and wrote the paper.*

RELATED WORK

The following publications are not included in the thesis, but summarise related work that I was involved in. The work is divided into peer-reviewed journal papers and conference contributions and is listed according to the thematic order of the thesis.

JOURNAL PAPERS

Paper ix: E. Memišević, J. Svensson, M. Hellenbrand, E. Lind, and L.-E. Wernersson, "Scaling of Vertical InAs–GaSb Nanowire Tunneling Field-Effect Transistors on Si", *IEEE Electron Device Letters*, vol. 37, no. 5, pp. 549–552, Mar. 2016, doi: 10.1109/LED.2016.2545861.

Paper x: E. Memisevic, E. Lind, M. Hellenbrand, J. Svensson, and L.-E. Wernersson, "Impact of Band-Tails on the Subthreshold Swing of III–V Tunnel Field-Effect Transistor", *IEEE Electron Device Letters*, vol. 38, no. 12, pp. 1661–1664, Oct. 2017, doi: 10.1109/LED.2017.2764873.

Paper xi: C. Möhle, C. B. Zota, M. Hellenbrand, and E. Lind, "$1/f$ and RTS noise in InGaAs nanowire MOSFETs", *Microelectronic Engineering*, vol. 178, pp. 52–55, May 2017, doi: 10.1016/j.mee.2017.04.038.

Paper xii: M. Berg, O.-P. Kilpi, K.-M. Persson, J. Svensson, M. Hellenbrand, E. Lind, and L.-E. Wernersson, "Electrical Characterization and Modeling of Gate-Last Vertical InAs Nanowire MOSFETs on Si", *IEEE Electron Device Letters*, vol. 37, no. 8, pp. 966–969, Jun. 2016, doi: 10.1109/LED.2016.2581918.

CONFERENCE CONTRIBUTIONS

Paper xiii: E. Memisevic, J. Svensson, M. Hellenbrand, E. Lind, and L.-E. Wernersson, "Vertical InAs/GaAsSb/GaSb Tunneling Field-Effect Transistor on Si with S = 48 mV/decade and I_{on} = 10 µA/µm for I_{off} = 1 nA/µm at V_{DS} = 0.3 V", *2016 IEEE International Electron Devices Meeting*, pp. 19.1.1–19.1.4, Dec. 2016, doi: 10.1109/IEDM.2016.7838450.

Paper xiv: M. Hellenbrand, O.-P. Kilpi, J. Svensson, E. Lind, and L.-E. Wernersson, "Comparison of Low-Frequency Noise in Nanowire and Planar III-V MOSFETs", *21st Conference on Insulating Films on Semiconductors (INFOS)*, Jun. 2019.

Paper xv: C. Möhle, C. B. Zota, M. Hellenbrand, and E. Lind, "$1/f$ and RTS noise in InGaAs nanowire MOSFETs", *20th Conference on Insulating Films on Semiconductors (INFOS)*, Jun. 2017.

Paper xvi: M. Hellenbrand, E. Memisevic, J. Svensson, A. Krishnaraja, E. Lind, and L.-E. Wernersson, "Effect of Gate Oxide Defects on Tunnel Transistor RF Performance", *76th Device Research Conference (DRC)*, pp. 137–138, Jun. 2018, doi: 10.1109/DRC.2018.8442145.

Acronyms and Symbols

Here, important acronyms, abbreviations, and symbols are listed, which are recurring throughout the thesis. Some parameters, which only occur in a narrow context, are intentionally omitted; some parameters are used in more than one way, but the context is always explicitly clarified in the corresponding text. Some (compound) units are provided with prefixes to reflect the most commonly encountered notations in the literature.

ACRONYMS AND ABBREVIATIONS

AC	Alternating Current
ALD	Atomic Layer Deposition
BEOL	Back End of Line
BTBT	Bant-to-Band Tunnelling
BTI	Bias Temperature Instability
CMF	Correlated Mobility Fluctuations
CV	Capacitance-Voltage
DC	Direct Current
DFT	Density Functional Theory
DOS	Density of States
GAA	Gate-All-Around
ICT	Information and Communications Technology
II	Impact Ionisation

IoT	Internet of Things
IV	Current-Voltage
LFN	Low-Frequency Noise
MOSFET	Metal-Oxide-Semiconductor Field-Effect Transistor
NDR	Negative Differential Resistance
NQS	Non-Quasi-Static
RC	Resistance-Capacitance
RF	Radio Frequency
RTN	Random Telegraph Noise
***s*-parameter**	Scattering Parameter
SEM	Scanning Electron Microscope
SOI	Silicon on Insulator
TFET	Tunnel Field-Effect Transistor
WKB	Wentzel-Kramers-Brillouin
***y*-parameter**	Admittance Parameter
***z*-parameter**	Impedance Parameter

LATIN SYMBOLS

$\mathbf{Al_2O_3}$		Aluminium Oxide
$C_{\mathbf{gx}}$	F	Small-Signal Gate-to-Drain (x=d) or Gate-to-Source (x=s) Capacitance
$C_{\mathbf{it}}$	F, $\mathrm{F\,cm^{-2}}$	Interface Trap Capacitance, typically normalised by the gated area
$C_{\mathbf{ox}}$	F, $\mathrm{F\,m^{-2}}$, $\mathrm{F\,m^{-1}}$	Gate Oxide Capacitance, often normalised by area or length
$C_{\mathbf{q}}$	F, $\mathrm{F\,m^{-2}}$, $\mathrm{F\,m^{-1}}$	Gate Oxide Capacitance, often normalised by area or length
$C_{\mathbf{sd}}$	F	Small-Signal Drain-to-Source Capacitance
$D_{\mathbf{it}}$	$\mathrm{cm^{-2}\,eV^{-1}}$	Interface Trap Density
$E_{\mathbf{A}}$	eV	Activation Energy

E_C	eV	Conduction Band Energy
E_F	eV	Fermi Level Energy
E_g	eV	Band Gap
E_V	eV	Valence Band Energy
f	Hz	Frequency
f_{max}	Hz	Maximum Oscillation Frequency
f_T	Hz	Cutoff Frequency
g_{ik}	S	Small-Signal Conductance of Band-to-Band Tunnelling and Impact Ionisation
g_{ds}	S	Output Conductance
g_m	S, $\mathrm{mS\,\mu m^{-1}}$	Transconductance, often normalised by the gate width
Ge		Germanium
h		$\approx 6.626 \times 10^{-34}$ Js, Planck Constant
$\hbar$		$\approx 1.055 \times 10^{-34}$ Js, Reduced Planck Constant
h_{21}	dB	Forward Current Gain
HfO_2		Hafnium Dioxide
I_D	A, $\mathrm{mA\,\mu m^{-1}}$	Drain Current, often normalised by the gate width
I_{DS}	A, $\mathrm{mA\,\mu m^{-1}}$	Source-to-Drain Current, often normalised by the gate width
I_G	A, $\mathrm{mA\,\mu m^{-1}}$	Gate Current, often normalised by the gate width
InAs		Indium Arsenide
InGaAs		Indium Gallium Arsenide
I_S	A, $\mathrm{mA\,\mu m^{-1}}$	Source Current, often normalised by the gate width
k		Stability Factor
k_B		$\approx 1.381 \times 10^{-23}\ \mathrm{kg\,m^2\,K^{-1}\,s^{-1}}$, Boltzmann Constant
L_G	m	Gate Length
m_0		$\approx 9.109 \times 10^{-31}$ kg, Electron Rest Mass
m^*	m_0	Effective Mass
MAG	dB	Maximum Available Gain

MSG	dB	Maximum Stable Gain
n	cm^{-3}, cm^{-2}, cm^{-1}	Carrier Concentration
N_{bt}	$\mathrm{cm}^{-3}\,\mathrm{eV}^{-1}$	Border Trap Density
N_{eff}	$\mathrm{cm}^{-2}\,\mathrm{eV}^{-1}$	Effective (Defect) Density
q		$\approx 1.602 \times 10^{-19}$ C, Elemental Charge
R_{i}, R_{j}	Ω	Intrinsic Non-Quasi-Static Channel Resistances
S	mV/decade	Inverse Subthreshold Slope
Si		Silicon
$S_{\mathrm{I_D}}$	$\mathrm{A}^2\,\mathrm{Hz}^{-1}$	Drain Current Noise Power Spectral Density
$\mathbf{SiO_2}$		Silicon Dioxide
$S_{\mathrm{V_G}}$	$\mathrm{V}^2\,\mathrm{Hz}^{-1}$	Equivalent Gate Voltage Noise Power Spectral Density
t_{ox}	m	Oxide Thickness
T_{BTB}		Band-to-Band Tunnelling Probability
T	K, °C	Temperature
U	dB	Unilateral Power Gain
v_{inj}	$\mathrm{m\,s}^{-1}$	Injection Velocity
v_{th}	$\mathrm{m\,s}^{-1}$	Thermal Velocity
V_{DS}	V	Drain-to-Source Voltage
V_{fb}	V	Flatband Voltage
V_{GD}	V	Gate-to-Drain Voltage
V_{GS}	V	Gate-to-Source Voltage
V_{SD}	V	Source-to-Drain Voltage
V_{T}	V	Threshold Voltage
W_{G}	m	Gate Width
x	m	Distance in x-direction
y_{gx}, $y_{\mathrm{gx,p}}$	S	Intrinsic and Parasitic (p) Small-Signal Conductances between Gate (g) and either Drain (x=d) or Source (x=s)
y_{kl}	S	y-parameter with indices k and l

GREEK SYMBOLS

α	eV^{-1}	Non-Parabolicity Factor
α	$\mathrm{Vs\,C}^{-1}$	Scattering Factor
α		Stress-Time Exponent
$\alpha_{\mathbf{H}}$		Hooge Parameter
γ		Frequency Exponent, Voltage Acceleration Factor
κ		Dennard Scaling Factor, Relative Permittivity
λ	m	Tunnelling Attenuation Length
$\mu_{\mathbf{e/h}}$	$\mathrm{cm}^2\,\mathrm{V}^{-1}\,\mathrm{s}^{-1}$	Electron/Hole Mobility
$\mu_{\mathbf{eff}}$	$\mathrm{cm}^2\,\mathrm{V}^{-1}\,\mathrm{s}^{-1}$	Effective Carrier Mobility
$\psi_{\mathbf{s}}$	V	Surface Potential
τ	s	Characteristic Time Constant
τ_0	s	Time Constant Effective Pre-Factor
$\tau_{\mathbf{c/e}}$	s	Capture/Emission Time Constant
ω	$\mathrm{rad\,s}^{-1}$	Angular Frequency
ω_0	$\mathrm{rad\,s}^{-1}$	Angular Trap Cutoff Frequency
ξ		Universal Relaxation Time
ζ	$\mathrm{V\,m}^{-1}$	Maximum Electric Field

INTRODUCTION

1

Survey of the Field

All our lives as well as the organisation of most human societies are interfused with myriads of electronic applications. Smartphones, computers, networks and databases, advanced medical care and research have all been enabled by the invention and continuous development of a now minute device called *Metal-Oxide-Semiconductor Field-Effect Transistor* (MOSFET); and transistors have become the single most abundantly mass-produced human artefact of all time. Their continuous improvement since their invention in the 1950s has been so rapid that today's smartphones possess a computing power that only a few decades ago would have been impossible to achieve at all.

This chapter provides a brief history of the development of the MOSFET as well as benchmarks of current research in order to provide a context and a motivation for the investigations undertaken in the thesis.

1.1 MOTIVATION OF THE CONDUCTED RESEARCH

The basic functionality of a transistor consists in the regulation of a current over several orders of magnitude by a controlling voltage. Based on this principle, different types of transistors are amongst the most essential components of almost all electrical appliances. As the most prominent example, billions of transistors make up every single processor in our computers, laptops, and smartphones, but transistors are also used in smaller numbers in e.g. audio and wireless communication amplifiers, in power supply controllers or as sensors.

The rapid development of the MOSFET since its invention in the 1950s has been driven mainly by the continuous miniaturisation of device dimensions and by the occasional introduction of a new material or a change in the device geometry. Because feature sizes have now decreased to the thickness of only a few atomic layers, geometric scaling is facing fundamental physical limits. Since about the beginning of the 21st century, for further improvement of the device performance in terms of energy efficiency and speed, new materials, new architectures, and new physical operation principles have thus moved into the focus of research efforts [1].

This thesis summarises detailed characterisation work on III-V nanowire MOSFETs, mostly in vertical device architectures, which employ both new materials with respect to the industry standard Si, and a new device architecture with respect to the conventional lateral device orientation. The appeal of III-V materials will be explained in Section 1.4.1 and the appeal of the vertical nanowire architecture in Section 1.4.2. As a special kind of MOSFET, so-called *Tunnel FETs* (TFETs) have been investigated, which are based on new physical operation principles. They will be introduced in Section 1.6. The different types of measurements in this thesis are divided into those related to defects in the gate oxide and those related to high frequency operation.

1.2 A BRIEF HISTORY OF THE FIELD-EFFECT TRANSISTOR

The theoretical principle of the transistor was conceived in the 1920s by the Polish/American physicist Julius Lilienfeld and patented in 1930 [2]. It would take until the late 1940s, however, before the first practical realisation of a transistor could be achieved. In late 1947, the physicists John Bardeen, Walter Brattain, and William Shockley, working at Bell Labs, New Jersey, demonstrated the field-effect and the point-contact transistor [3, 4]. Independently, in 1948, the German physicists Herbert Mataré and Heinrich Welker made the same discovery while working at the Compagnie des Freins & Signaux Westinghouse near Paris [5]. A further important step in the history of transistors was the invention of the bipolar junction transistor by William Shockley [6], but this type of transistor was difficult to mass-produce. So the unstoppable digital revolution only took its course when the Egyptian and the Korean engineers Mohamed Atalla and Dawon Kahng, respectively, both working at Bell Labs, experimentally realised the first MOSFET. Again, the concept of the MOSFET had been known for several years, but the presence of electronic states at the surface of the semiconductor, which was to form the conductive channel of the MOSFET, prevented proper control of the current in the devices. The critical step to set the digital revolution into motion was the passivation of these interface states in the Si/SiO_2 material system [7].

Together with the invention of the Integrated Circuit and the Complementary MOS logic [8], this stable Si material system finally heralded the commercial mass production of electronic applications that we enjoy today.

1.3 BASIC MOSFET OPERATION

The basic operating principle of a MOSFET is the control of a current between two of its terminals, *source* and *drain*, by a voltage at the third terminal, the *gate*, without a current flowing into or out of the gate itself. The path between source and drain is called *channel* and the control of the gate over the channel as well as their isolation from each other is achieved by the *gate oxide*, which is a central topic of this thesis. Schematics of the device structure and the described principle of operation are depicted in Fig. 1.1. In such a so-called n-type bulk MOSFET, the source and drain contacts are realised by highly n-doped regions implanted into a p-doped substrate. In a heavily n-doped semiconductor, the Fermi level resides close to the conduction band edge so that a large concentration of electrons is available in the source. In the p-doped channel region on the other hand, the Fermi level lies close to the valance band edge so that without an externally applied voltage V_{GS} between gate and source, the resulting band structure along a cut from source to drain (Fig. 1.1(b), top) constitutes an energy barrier for the electrons in the source, which the vast majority cannot overcome. The flow of a current is thus not possible between source and drain and the transistor is in its off-state. Upon application of a positive V_{GS}, an electric field is created via the

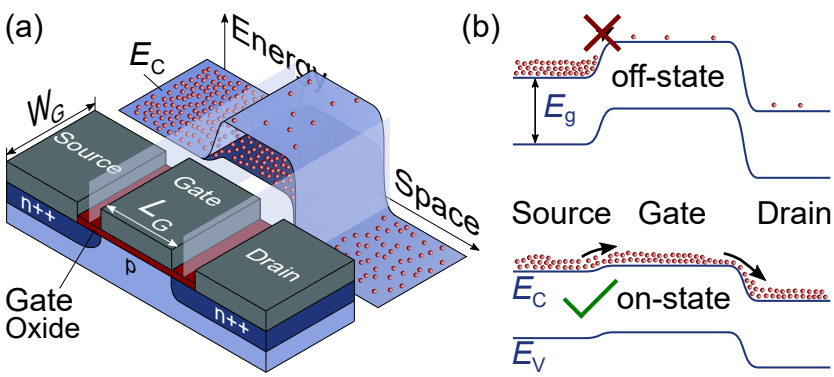


Figure 1.1: (a) Schematics of a planar MOSFET structure and the corresponding conduction band diagrams for the off- and the on-state (light blue and dark blue, respectively). (b) Corresponding energy band diagrams for the off- and the on-state in 2D along a cut from source to drain in (a).

gate oxide, which modifies the energy barrier in the channel. With increasing V_{GS}, the barrier decreases, more and more electrons can surmount it, and with a sufficiently large voltage V_{DS} between source and drain, the source-to-drain current I_{DS} increases so that the transistor gradually reaches its on-state (Fig. 1.1(b), bottom). Ideally, a MOSFET exhibits high on-currents and low off-currents. The characteristic voltage which separates the on- from the off-state, is called the *threshold voltage* V_T. A fourth contact to the bottom of the device in Fig. 1.1(a) can be used to adjust V_T, but this was not used for any devices in this thesis and is thus disregarded from here on. As mentioned before, in an ideal MOSFET, the gate oxide prevents any current I_G from flowing into the gate contact. In reality, a small leakage current is present, which is however several orders of magnitude lower than I_{DS}.

1.4 DENNARD SCALING, MOORE'S LAW AND BEYOND

A large part of the route from the first clunky, hand-sized, point-contact transistor to today's high-speed transistors of just a few nanometres was guided by scaling principles known as *Dennard scaling*. In a paper in 1974 [9], the American electrical engineer Robert H. Dennard and colleagues observed that if the gate length L_G and width W_G as well as the gate oxide thickness t_{ox} (see Fig. 1.1(a)) and the supply voltage are all scaled down by a factor $1/\kappa$, while the doping density is scaled up by a factor κ, the switching delay time of the device decreases by $1/\kappa$, while its power density (voltage times current per area) remains constant. Widespread adherence to these scaling rules led to Gordon Moore's (co-founder of Intel) famous observation that the number of transistors in an integrated circuit doubles about every two years [10], which ultimately brought us today's state of the art in electronics. In recent years, however, this constant-power scaling has slowed down, because it is facing fundamental physical limits. Other means of improving transistor performance have thus been introduced and the following sections provide an overview of some important developments during the last years.

1.4.1 MATERIALS AND MOSFET PERFORMANCE

The continuous geometric scaling of MOSFETs eventually led to gate oxide thicknesses of just a few atomic layers so that electrons in the transistor channel started tunnelling through these thin oxides. This constituted the first fundamental physical challenge for geometric scaling, since it caused increased gate currents and thus increased power consumption [11]. The challenge could eventually be overcome by the introduction of so-called *high-κ* (not to be confused with Dennard's scaling factor) gate oxides. For a planar capacitor structure, such as the gate stack of a large-area MOSFET,

the capacitance C is calculated as $C = \epsilon_0 \epsilon_r A / t_{ox}$, where ϵ_0 is the vacuum permittivity, ϵ_r (called κ in engineering) is the relative permittivity of the capacitor dielectric (i.e. the gate oxide), A is the area, and t_{ox} the thickness of the oxide. From this it can be seen that instead of thinning down the oxide, the capacitance can be increased by using a material with a higher relative permittivity κ. Since a higher permittivity, however, comes at the cost of a lower band gap E_g, which often leads to a reduced band offset between the oxide and the channel and thus again to increased tunnelling, a compromise had to be found between high values for κ and for E_g. The best compromise proved to be HfO_2 with $\kappa = 25$ (ideal case) and $E_g \approx 6$ eV [12] so that the too thin SiO_2 ($\kappa = 3.9$) gate oxide was replaced by a thicker HfO_2 gate oxide [13]. At the same time, a metal gate contact was introduced, which suppressed surface scattering at the channel/high-κ interface [14] and eliminated the problem of depletion in the previously used polysilicon contacts, which could cause unwanted shifts in the threshold voltage V_T [11].

Besides the change of the MOSFET gate stack, ongoing research is concerned with alternatives to Si as the channel material. III-V materials, which constitute the foundation of devices in this thesis, are one of the most promising options for this due to their outstanding electron transport properties [15]. The materials are referred to as III-V materials, because they are compounds of elements of the (Mendeleev) groups III and V in the periodic table of elements (IUPAC groups 13 and 15) and their outstanding transport characteristics are related to a material property known as the *effective mass* m^* of charge carriers. Generally, for parabolic bands, the effective mass can be expressed as a tensor, which depends on the wave vector $\boldsymbol{k}$. The elements (i, j) of this tensor are defined as [16, p. 243]

$$\left(\frac{1}{m_n^*(\boldsymbol{k})}\right)_{i,j} = \frac{1}{\hbar^2} \frac{\partial^2 E_n(\boldsymbol{k})}{\partial k_i \partial k_j}, \tag{1.1}$$

where $E_n(\boldsymbol{k})$ is the dispersion relation for charge carriers in the nth energy band, $\hbar$ is the reduced Planck constant, and k_i and k_j are the ith and jth element of the wave vector $\boldsymbol{k}$. In many practical cases, for small energies relative to the band edges, the dispersion relation $E(\boldsymbol{k})$ can be approximated as parabolic and isotropic so that the second derivative in (1.1) is a constant. This is known as the *effective mass approximation*. For some (doped) III-V materials, due to their low density of states, the Fermi level can move quite far into the energy bands so that $E(\boldsymbol{k})$ deviates from this parabolic approximation. From $\boldsymbol{k} \cdot \boldsymbol{p}$ perturbation theory, to a first order, this can be taken into account by a modified dispersion relation as $E(\boldsymbol{k})(1 + \alpha E(\boldsymbol{k})) = \hbar^2 k^2 / (2m^*)$ with the non-parabolicity factor α [17].

In state-of-the-art MOSFETs, the conducting channel is typically formed by a 2D electron gas and the effective gate length is so short that electrons pass the channel ballistically, i.e. without scattering. For a fully ballistic 2D MOSFET in the on-state (i.e. large V_{DS} and $V_{GS} \gg V_T$) the current can be described as [18, p. 100]†

$$I^{2D} \approx \frac{2qW_G\sqrt{2m^*}}{3\pi^2\hbar^2}\left(\frac{qC'_{ox}}{C'_{ox}+\frac{C_q^{2D}}{2}}\right)^{3/2}(V_{GS}-V_T)^{3/2}, \tag{1.2}$$

where q is the elemental charge, C'_{ox} is the series combination of the oxide capacitance and the centroid capacitance, C_q^{2D} is the 2D quantum capacitance, and all other parameters are as before. The centroid capacitance, from first-order perturbation theory, takes into account the shift in energy of the bottom of the sub bands when they are occupied by charge carriers. Due to their small channel diameters, the nanowire MOSFETs which have been studied for this thesis approach 1D transport characteristics. In this case, the current can be expressed as [18, p. 146]

$$I^{1D} = \frac{2q^2}{h}\frac{C'_{ox}}{C'_{ox}+C_q^{1D}}(V_{GS}-V_T), \tag{1.3}$$

where h is the Planck constant and C_q^{1D} the 1D quantum capacitance.

The general definition of the quantum capacitance C_q (of each occupied sub band) in (1.2) and (1.3) is $C_q = q\partial n/\partial\psi_s$ with the carrier density n (in the respective dimensionality) and the surface potential ψ_s at the top of the channel barrier [18, p. 51]. With this, C_q^{2D} and C_q^{1D} for each sub band can be expressed as

$$C_q^{2D} = \frac{q^2m^*}{\pi\hbar^2} \tag{1.4a}$$

and

$$C_q^{1D} = \frac{q^2}{\pi\hbar}\sqrt{\frac{2m^*}{E_F-\mathcal{E}(0)}}, \tag{1.4b}$$

where E_F and $\mathcal{E}(0)$ in (1.4b) denote the Fermi level and the bottom of the sub band, respectively. From (1.2)–(1.4) it can be seen that for a constant drive voltage, to a first order, higher on-currents can be achieved by using channel materials with lower effective masses. Furthermore, the quantum capacitances in (1.4) almost directly determine the density of states in a MOSFET channel so that low effective masses also entail low densities of states (DOS) [18, p. 7]. This can be beneficial for high frequency operation, as

†The expressions differ a little, because [18] uses the MOS limit, i.e. assumes that $C_q \gg C_{ox}$, which is usually not appropriate for III-V MOSFETs.

investigated in Chapter 3, since lower DOS lead to lower gate capacitances. As a another useful relation, by employing first-order perturbation theory it can be shown that in materials with parabolic bands, a low effective mass is linked to a small band gap [19, p. 69], which explains the corresponding correlation in Table 1.1.

In the drift-diffusion equation, which describes long channel MOSFETs, the current is often expressed in terms of the mobility μ rather than the effective mass. The drift-diffusion model is only valid for effective gate lengths $L_{\mathrm{eff}} \gg 2k_{\mathrm{B}}T\mu_{\mathrm{eff}}/(qv_{\mathrm{th}})$ with the effective carrier mobility μ_{eff} and the thermal velocity $v_{\mathrm{th}} = \sqrt{2k_{\mathrm{B}}T/(\pi m^*)}$ [18, p. 96]. In ballistic MOSFETs, the corresponding alternative metric to the effective mass is the injection velocity v_{inj}, which is used in a simpler formulation of the MOSFET on-current, $I = q(n^+ - n^-)v_{\mathrm{inj}}$, where n^+ and n^- are the concentrations of charge carriers with positive and negative directed moments, respectively [18]. Table 1.1 provides examples of material parameters for III-V materials used in devices in this thesis. Values for graphene are provided as a further reference, since graphene has attracted

Table 1.1: Important material parameters (at room temperature) for Si, Ge, two III-V materials used in devices in this thesis, and graphene. Values for semiconductors were obtained from [20] and values for graphene from [21–23]. All values indicate upper limits for pure and free-standing materials and are typically not achieved in actual devices. m_0 is the electron rest mass.

Material	Si	Ge	InGaAs[c]	InAs	Graphene
Band gap E_{g} (eV)	1.12[id]	0.66[id]	0.74[d]	0.35[d]	0[d]
Effective electron mass m_{e}^*/m_0	0.26	0.12	0.041	0.023	$\sim$0.033[†]
Effective light hole mass m_{h}^*/m_0	0.16	0.043	0.052	0.026	$\sim$0.045[†]
Electron injection velocity v_{inj} (m/s)	2.3×10^5	3.1×10^5	5.5×10^5	7.7×10^5	$\sim 1 \times 10^6$
Electron mobility μ_{e} (cm^2/(Vs))	1400	3900	12 000	40 000	200 000
Hole mobility μ_{h} (cm^2/(Vs))	450	1900	300	500	[‡]

[c] $In_{0.53}Ga_{0.47}As$. [id] and [d] indicate indirect and direct band gaps, respectively. [†] Note that a different definition than (1.1) has to be used for graphene, since its dispersion relation is not parabolic. [‡] Due to the almost symmetric band structure of graphene, μ_{h}^* should be similar to μ_{e}^*, but values as high as for μ_{e}^* were not reported in [22].

a lot of interest due to its potentially outstanding transport properties as well. Note that a different definition for m^* than provided by (1.1) has to be used for graphene, since it exhibits an almost linear dispersion relation.

The values in Table 1.1 clearly demonstrate the promising material properties of III-V materials with much lower effective masses and higher mobilities than the most common MOSFET material Si. It should be noted, though, that in transistors, the mobility is often greatly reduced by scattering due to ionised dopants and/or due to material defects at the channel/oxide interface and in the bulk of the oxide [24, 25]. It is much more difficult to obtain high-quality gate oxides and interfaces on III-V channels [26, 27] than on Si and effects resulting from different kinds of gate oxide defects are a central topic of this thesis. The effective mass in Si can be lowered to some degree by inducing strain in the channel [28], but for very-high-frequency and low-power transistors, it constitutes a severe limitation.

Besides III-V materials, many more candidates with promising properties are being studied for transistor applications. Examples encompass carbon nanotubes, the aforementioned graphene, or other 2D materials such as transition metal dichalcogenides, but a more thorough discussion is beyond this overview.

1.4.2 GEOMETRIES

A second major complement to the traditional geometric scaling has been the introduction of new device geometries for MOSFETs. In recent years, FinFET/tri-gate and silicon on insulator (SOI) geometries have replaced the formerly planar MOSFETs in commercial applications, and gate-all-around (GAA) structures are being investigated for the coming generations of tran-

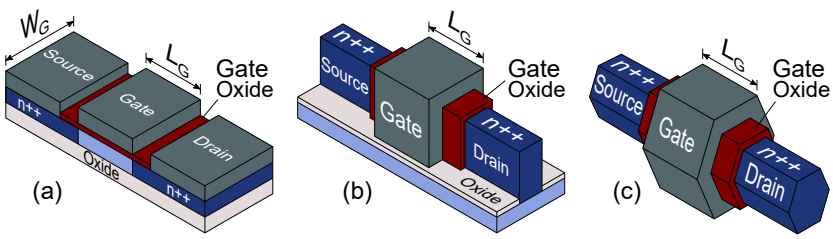


Figure 1.2: Schematic representations of advanced MOSFET structures. (a) Silicon on insulator. The buried oxide suppresses substrate leakage and reduces the intrinsic capacitances. (b) FinFET/tri-gate. Electrostatic control is strongly improved by gating the channel from three sides instead of only one. (c) Gate-all-around. Optimised electrostatic control due to gating from all sides. The nanowire structure also facilitates the co-integration of different materials.

sistors. In this thesis, especially vertical nanowire GAA MOSFETs are investigated extensively. Schematics of the different device structures are presented in Fig. 1.2 and Fig. 1.3, and a performance overview for different architectures can be found in Section 1.5.

In the SOI device structure in Fig. 1.2(a) a buried oxide underneath the channel suppresses leakage current in areas where the gate electric field decreases. It also reduces the intrinsic capacitances of the transistor, which is beneficial for high frequency applications. The FinFET/tri-gate structure in Fig. 1.2(b) takes this approach further by raising up the channel from the substrate into a fin, which can be gated from three sides. With the resulting increased electrostatic control, short-channel effects can be suppressed more efficiently so that thicker gate oxides as well as thicker and/or shorter channels are possible [29]. The ultimate electrostatic integrity can be achieved in GAA nanowire structures such as the one in Fig. 1.2(c), where the gate completely encloses the channel [30]. This structure, especially in a vertical orientation as in Fig. 1.3, also greatly facilitates the co-integration of different semiconductor materials, since the nanowire geometry allows a relatively easy strain relaxation of mismatched lattice constants [31]. This is very beneficial for engineering the band structure along the channel, which is of special

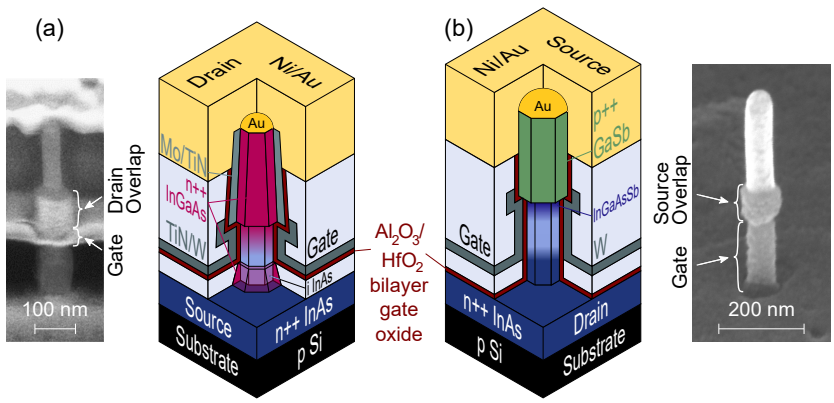


Figure 1.3: Schematic representations and scanning electron micrographs ('SEM images') of the vertical nanowire transistors studied in this thesis. (a) MOSFET; operation with the drain contact at the top provides higher on-currents and lower off-currents, but the device can also be operated with the drain in the bottom. The SEM image (courtesy O.-P. Kilpi) shows a cut (by a focused ion beam) through the sample next to a complete device. (b) TFET; the asymmetric transport principle precludes operation with the source in the bottom. The SEM image (courtesy A. Krishnaraja) was taken during processing after deposition of the gate metal.

importance for the TFETs, as explained in Section 1.6. Furthermore, a vertical geometry decouples the gate length from the area footprint of a device, which allows a dense spacing of devices on a chip even for relatively long gate lengths. Vertical nanowire MOSFETs could also prove instrumental for emerging monolithic 3D integrated chip architectures, which aim to overcome the bottleneck of data transmission between computation and memory [32]. All vertical devices fabricated in Lund and characterised in this thesis are integrated on Si substrates for compatibility with industry fabrication schemes. Details about the vertical nanowire MOSFETs can be found in [33] and about the TFETs in [34].

1.5 PERFORMANCE OVERVIEW

This section summarises state-of-the-art performance at the time of writing this thesis. Different important metrics are briefly explained and benchmarking figures provide a quantitative comparison of how the different approaches described above enhance transistor performance in practice. The section is divided into a DC and a high frequency part. Details and derivations of all metrics and the underlying transport physics can be found in textbooks such as [18] for the DC metrics and [35] for the high frequency metrics.

1.5.1 DC METRICS

The DC performance of a transistor can be characterised by different metrics, which quantify the relations between the current and the voltages. Examples of the two basic IV curves for this purpose are presented in Fig. 1.4. They are referred to as the *transfer characteristics* (Fig. 1.4(b) currents vs. gate-to-source voltage V_{GS}) and *output characteristics* (Fig. 1.4(c) drain current I_D vs. drain-to-source voltage V_{DS}). To be able to compare the quality of differently sized transistors, some of these metrics are often normalised by the gate width W_G of the transistor, see Fig. 1.1(a). For fin or GAA structures, typically, the perimeter of the gated surface is used for normalisation.

In industry, typically, an application-dependent off-current I_{off} is defined (e.g. $1\,\mathrm{nA}/\mu\mathrm{m}$) and used to determine the on-current I_{on} by moving along the voltage axis of the transfer curve in Fig. 1.4(b) by an amount V_{DS}, sometimes called the drive voltage V_{DD}, at which the transfer curve was measured. The minimum off-state current can be limited by detrimental factors such as gate leakage, direct source-to-drain tunnelling for very short channel lengths ($\lesssim$ 20 nm), band-to-band tunnelling (BTBT) on the drain side, or thermionic emission from the high-energy tail of the electron distribution. A suppression of BTBT on the drain side can be achieved by grading the channel

material composition, see Fig. 1.4(a), for which vertical III-V nanowires are very suitable [36]. To obtain a specific number, which characterises the transition between off- and on-state, the so-called threshold voltage V_T can be determined. One way to do so consistently is demonstrated in Fig. 1.4(b) as well: V_T is given by the point, where the linear extrapolation of I_D from the point of the maximum *transconductance* g_m intersects the voltage axis. The transconductance g_m in turn is the derivative of I_D with respect to V_{GS}, $g_m = \partial I_D / \partial V_{GS}$. As such, g_m is an important metric in its own right, since it describes how efficiently small changes in the gate voltage modulate the current. Typically, g_m is an important metric for the on-state performance, since it determines the gain that a transistor can provide and as such, high values for g_m are especially important for high frequency applications. The current modulation by the gate in the off-state is characterised by the *inverse subthreshold slope* $S = (\partial \log(I_D)/\partial V_{GS})^{-1}$, also referred to as *subthreshold swing*, and indicated in Fig. 1.4(b). Ideally, a MOSFET exhibits high values for g_m for high gain, and low values for S for low off-state leakage. Further on-state metrics, indicated in Fig. 1.4(c), are the on-resistance R_{on} and the output conductance g_{ds}, which are ideally as small as possible. The inverse subthreshold slope S in a conventional MOSFET is physically limited to $\sim$60 mV/decade, which will be discussed in more detail in Section 1.6. As an overview of state-of-the-art MOSFET performance, Fig. 1.5 presents reported

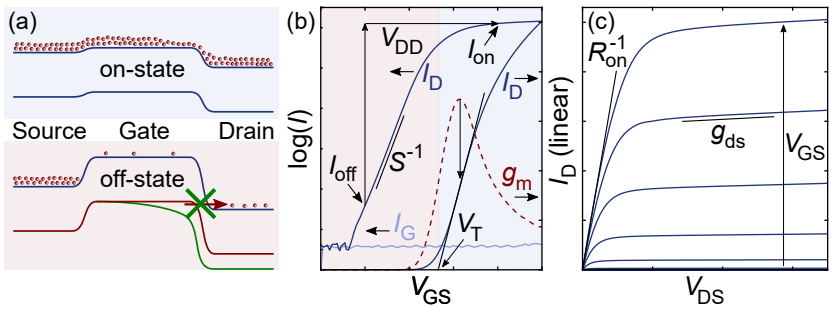


Figure 1.4: The most important IV curves to characterise the DC performance of a MOSFET alongside the band diagrams from Fig. 1.1 to illustrate the relation between currents and band alignment. (a) Band diagrams. The red valence band in the off-state illustrates band-to-band tunnelling as one of the possible sources for increased off-state current, the green valence band illustrates a compositionally graded channel with a larger band gap material on the drain side to suppress this. (b) Transfer characteristics, currents as a function of V_{GS}. Important metrics are indicated together with ways to extract them. (c) Output characteristics, I_D as a function of V_{DS}, with important metrics.

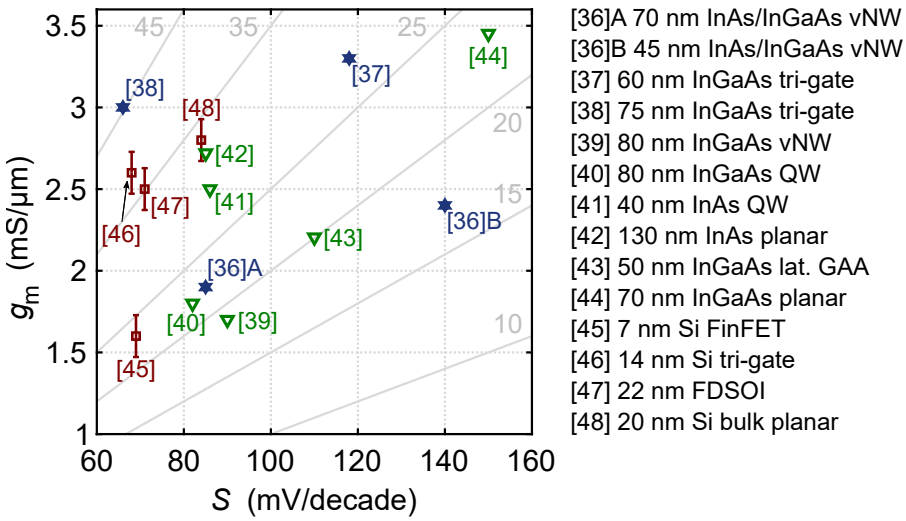


Figure 1.5: Maximum transconductance g_m vs. minimum inverse subthreshold slope S for different reported state-of-the-art MOSFETs (with different gate lengths). Blue stars and green triangles are for III-V [36–44], where blue stars indicate the types of transistors characterised in this thesis. Grey lines indicate the quality factor $Q = g_m/S$. $g_m = 3.1\text{mS}/\mu\text{m}$ was demonstrated for vertical III-V nanowire RF devices in [33], but the devices did not turn off properly and thus do not appear in the figure. Si devices (red squares, [45–48]) are provided as references and error bars are provided since the values were not reported explicitly, but calculated from reported IV curves. Values for III-V MOSFETs are for $V_{DS} = 0.5$ V and values for Si are for $V_{DS} = 0.7$–0.9 V. Acronyms in the legend: vNW – vertical nanowire; QW – quantum well; lat. GAA – lateral gate-all-around; FDSOI – fully depleted silicon on insulator.

values for g_m vs. S for different MOSFETs. III-V devices, among them the ones fabricated in Lund and characterised in this thesis, compare very well with commercial Si devices despite the largely different efforts of optimisation that have gone into the respective technologies.

1.5.2 HIGH FREQUENCY METRICS

Besides the metrics which can be obtained from DC operation, there are some metrics which describe the high frequency performance of a transistor. The most important ones in the context here are the cutoff frequency f_T and the maximum oscillation frequency f_{max}, which denote the frequencies, at which the forward current gain h_{21} and the unilateral power gain U, respectively, decrease to one so that the transistor cannot act as an amplifier anymore at

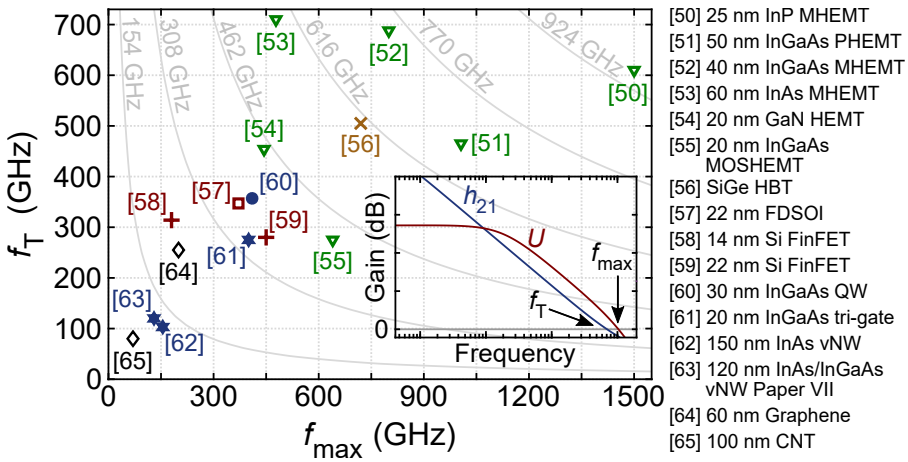


Figure 1.6: Some of the highest reported operation frequencies for different transistor technologies [50–65]. Blue stars indicate the types of transistors characterised in this thesis, commercial Si devices are marked by a red square/red plusses. The inset illustrates the extraction of f_T and f_{max}. All values except for GaN [54] are for V_{DS} between 0.3 V and ∼1 V. Although higher record values for f_T have been reported for Si (485 GHz [66]) and graphene (427 GHz [67]), values for f_{max} were not reported alongside so that these values are not included in the figure. Acronyms in the legend: MHEMT – metamorphic HEMT; PHEMT – pseudomorphic HEMT; HBT – heterojunction bipolar transistor; FDSOI – fully depleted silicon on insulator; QW – quantum well; vNW – vertical nanowire; CNT – carbon nanotube.

and beyond these frequencies [49]. A high f_T and f_{max}, to a first order, can be obtained by a high gain (high g_m) and low capacitances in the transistor. Besides h_{21} and U, the maximum stable gain MSG and the maximum available gain MAG are noteworthy, which, along with the other high frequency metrics, are treated in more detail in Chapter 3.

Because of difficulties with scalability and with long-term reliability due to gate oxide defects, III-V transistors have turned out to be more suitable for analogue high frequency applications, such as amplifiers for wireless communication or radar, rather than for digital applications. At these high frequencies, in turn, III-V transistors clearly surpass Si devices, as can be seen in Fig. 1.6, where an overview of reported state-of-the-art f_T and f_{max} is provided. Different from Fig. 1.5, which only included MOSFETs due to their superior electrostatics, Fig. 1.6 also includes other transistor technologies, most notably high electron mobility transistors (HEMTs), which achieve by far the highest operation frequencies. In HEMTs, the doping impurities to increase carrier concentrations are physically separated from the channel

in a delta doping layer and the gate isolation is realised via a Schottky barrier rather than a gate oxide, which greatly reduces the aforementioned degradation of the superior III-V transport properties. On the downside, the absence of a gate oxide in HEMTs leads to a severe increase of the off-state leakage at short ($\lesssim$ 20 nm) gate lengths and thus impedes the scalability of this technology.

1.6 NEW PHYSICS – TUNNEL FETS

The operation principle of MOSFETs inherently determines a fundamental lower limit for the inverse subthreshold slope S. This limit of $\sim$60 mV/decade can be discerned in Fig. 1.5 and can only be overcome by the introduction of new physical operation principles. Different intriguing mechanisms are under investigation to be exploited for such novel devices and a well accessible overview can be found e.g. in [68]. Here, the so-called *Tunnel FET* (TFET) will be discussed, which has long been regarded as one of the most promising device classes for operation with steep inverse subthreshold slopes.

As fermions, charge carriers in the MOSFET source and drain contacts are distributed according to the Fermi-Dirac distribution, the high-energy tail of which extends infinitely in energy. This is illustrated in Fig. 1.7(a) and as can be seen, charge carriers in this high-energy tail can overcome the channel energy barrier even in the off-state of the transistor. Due to the shape of the Fermi distribution, the inverse subthreshold slope S in a MOSFET is thus physically limited to

$$S = \left(\frac{\partial \log(I_\mathrm{D})}{\partial V_\mathrm{GS}} \right)^{-1} = \ln(10) \frac{k_\mathrm{B} T}{q} \approx 60 \text{ mV/decade}, \tag{1.5}$$

where log and ln are the logarithms with base 10 and the natural logarithm, respectively, k_B is the Boltzmann constant, and T the temperature. In order to achieve values for S, which are lower than this physical limit, the high-energy Fermi tail of the source carrier distribution has to be suppressed. In TFETs this can be achieved by a band structure as the one schematically depicted in Fig. 1.7(b), where the source band gap cuts off the Fermi tail and thus acts as a filter for the high-energy states. To turn a TFET off, the channel bands are moved upwards by the gate in the same manner as in a conventional MOSFET. This aligns the channel band gap with the source in such a way that there are no states available for source electrons to tunnel into and a current flow is not possible. This is illustrated by the red energy bands in Fig. 1.7(b). Since the Fermi tail in the source is cut off, no thermionic leakage current can flow either. In contrast to (1.5), the (ideal) inverse subthreshold slope of a TFET

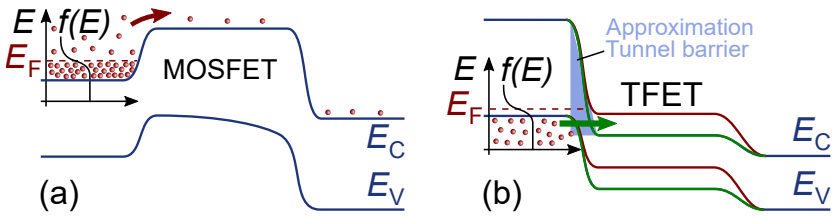


Figure 1.7: Schematic representations of the (a) MOSFET and (b) TFET band structure. (a) Even in the off-state, a small amount of charge carriers can always overcome the channel energy barrier. This limits the inverse subthreshold slope to 60 mV/decade. (b) In TFETs, the energy bands along the device are engineered in such a way that the source band gap cuts off the high-energy Fermi tail. This enables subthreshold swings below 60 mV/decade in the off-state (red bands). In the on-state (green bands), current can flow via band-to-band tunnelling.

is thus not temperature-dependent so that TFETs can be switched off with a much smaller change of V_{GS} than conventional MOSFETs.

In the on-state of a TFET, electrons in the valence band of the p-doped source can tunnel into empty states in the undoped channel and thus constitute a current. This is illustrated by the green energy bands in Fig. 1.7(b). To calculate the tunnelling probability in the on-state, the tunnel barrier, determined by the band gap, can be approximated as triangular, as indicated in Fig. 1.7(b) as well. In a one-dimensional system, the resulting band-to-band tunnelling probability T_{BTB}^{1D}, which determines the on-current, can then be calculated as [69]

$$T_{BTB}^{1D} = \exp\left(-\frac{4\sqrt{2m_e^*}E_g^{3/2}}{3q\hbar\zeta}\right), \tag{1.6}$$

where m_e^* is the effective tunnelling mass, E_g is the source band gap, and ζ is the maximum electric field at the tunnel junction. Unfortunately, this expression typically yields rather small tunnelling probabilities of only a few per cent, which reveals a challenge for ongoing TFET research: It is difficult to achieve TFET on-currents which are sufficiently large for circuit applications. III-V materials, again, can be of great benefit to tackling this challenge, since they allow flexible band engineering [31] and indeed, the III-V TFETs characterised in this thesis achieve the highest on-currents of all reported TFETs while maintaining $S < 60$ mV/decade. This is demonstrated in Fig. 1.8, where it can also be seen that irrespective of the challenge for the on-state, TFETs have demonstrated significant improvement regarding the off-state of transistors. Since the gate/channel area in TFETs is controlled in

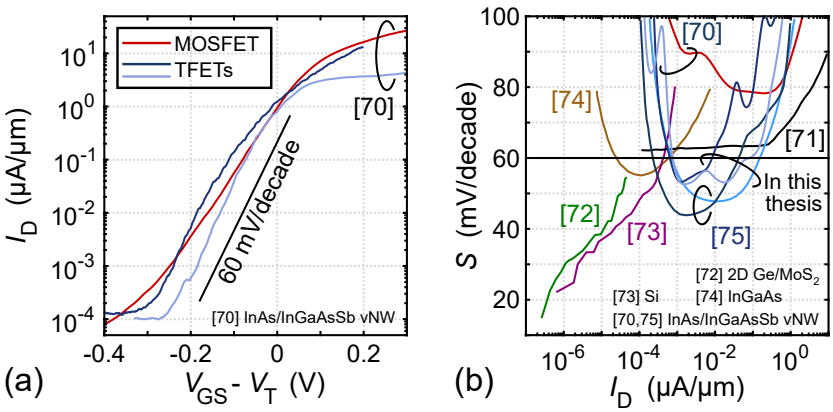


Figure 1.8: (a) Comparison of the transfer curves of a MOSFET and two vertical nanowire (vNW) TFETs with the same gate oxides [70]. It is clearly discernable that the TFETs exhibit steeper slopes. (b) Benchmark of state-of-the-art TFET subthreshold swings. Here it also becomes clear that TFETs reach steeper slopes than the fundamental 60 mV/decade. A commercial 16 nm thermionic MOSFET from [71] serves as a reference and [72–74] present some state-of-the-art TFETs published at the time of writing this thesis. The TFETs investigated in this thesis exhibit the highest currents of all TFETs, while still reaching slopes below 60 mV/decade [70,75].

the same way as in conventional MOSFETs, in some contexts, TFETs can be regarded as a special kind of MOSFET. Most of the discussions about oxide traps later in this thesis, for example, apply to TFETs in the same manner as to 'non-tunnel' MOSFETs.

Since the value of S is so important for ascertaining the successful realisation of a TFET, a dedicated comment on its measurement is appropriate. Inverse subthreshold slopes below 60 mV/decade cannot only result from a suppression of the Fermi tail, but also from the capture or emission of charge carriers by gate oxide traps at a critical time during the IV measurement. Such capture/emission processes shift the channel energy bands and thus affect the slope of the IV curve. (See Chapter 2 for more details.) It is therefore important always to measure the hysteresis of the transfer curve, ideally also to measure over different voltage intervals in the subthreshold region, and to measure at different voltage sweep rates. Such detailed measurements can uncover effects of traps and warrant the reliable characterisation of a TFET.

Since the source Fermi tail is cut off by the band gap, as mentioned, the ideal inverse subthreshold slope of a TFET is not temperature-dependent. However, the presence of gate oxide defects, band tails due to doping impurities and other material defects, and especially the presence of defects at the

tunnel junction, can all reintroduce a temperature dependence due to defect-assisted tunnelling and noise mechanisms. For this thesis, such effects were investigated amongst others in Paper III and the effect of gate oxide defects on TFET operation will be elucidated in more detail in Chapter 2. Further temperature dependences in TFETs can be caused by a degenerate doping of the source (i.e. the Fermi level is moved into the valence band) and, especially in the off-state, by Shockley-Read-Hall generation of carriers [76].

Since the treatment of the mentioned band tails can become quite tedious [77], a simple phenomenological formulation was developed in [78] (Related Work in this thesis) to describe the effect of band tails in TFETs. In this formulation, together with the effects of the channel capacitance and defects at the channel/gate oxide interface, the expression for S in a TFET becomes

$$S = \ln(10)\frac{k_B T E_0}{k_B T + E_0}\left(1 + \frac{C_q + C_{it}}{C_{ox}}\right), \tag{1.7}$$

where C_{it} and C_{ox} are the interface defect capacitance and the geometric oxide capacitance, respectively, and E_0 is an energy decay parameter, which describes the exponential decrease of the band tails into the band gap. C_{it} can be calculated as $C_{it} = qD_{it}$ with the interface defect density D_{it}. The centroid capacitance was omitted here for a simpler expression.

A quite intriguing feature of TFETs with a degenerately doped source contact is the occurrence of *negative differential conductance* (NDR) upon variation of the source voltage V_{SD} with respect to the drain bias, as illustrated in Fig. 1.9. If the biasing of the device is reversed with respect to 'normal'

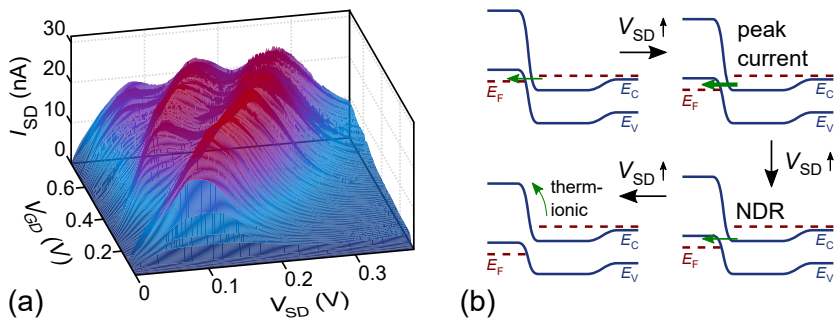


Figure 1.9: (a) Detailed measurement of the TFET NDR region at 50 K. The drain terminal was grounded and the gate-to-drain and source-to-drain voltages V_{GD} and V_{SD}, respectively, were varied. Multiple peaks were visible at all temperatures, but more pronounced at low temperatures. (b) Band diagrams corresponding to different V_{SD} for a constant V_{GD} to illustrate the different conduction regions in (a).

operation, i.e. the drain acts as the reference terminal and voltages are applied to the gate and the source instead of the gate and the drain, charge carriers can tunnel from the channel into the source instead of the other way around. At a fixed gate bias, first, as in normal operation, this causes an increase of the tunnelling current when increasing the source voltage, because the tunnelling window increases. Above a certain voltage, however, the tunnelling window starts decreasing again so that for an increase in voltage, the current decreases. Because the derivative of the current with respect to the voltage is negative in this case, the relation is called NDR. Such an NDR relation could be utilised e.g. to realise TFET-based voltage-controlled oscillators, where the NDR could compensate the positive resistance in the circuit. Fig. 1.9(a) presents a detailed measurement of the NDR characteristics as a function of both V_{SD}, the source-to-drain voltage, and V_{GD}, the gate-to-drain voltage, with a constant drain voltage of zero volts. At even higher V_{SD} than presented in Fig. 1.9(a), electrons are thermionically injected into the source so that the current after the valley between $V_{GD} \approx 0.2$ and 0.4 V increases again. This is indicated in Fig. 1.9(b). The multiple peaks beyond $V_{GD} \approx 0.2$ V in Fig. 1.9(a) go beyond the qualitative description provided here and require a more thorough analysis. Further details about the general operation principles of TFETs can be found in e.g. [69] or [79].

2

Characterisation of Gate Oxide Defects

SINCE it proves to be challenging to obtain high-quality gate oxides on top of III-V channels, it is important to understand the effects that non-idealities related to the oxide can have on the MOSFET performance. In this chapter, work on the characterisation of defects related to the gate oxide in both MOSFETs and TFETs at 'very slow' ($\gtrsim$ microseconds) time scales is summarised. The characterisations comprise hysteresis, low-frequency noise, and bias temperature instability measurements. A short overview of the chemical nature of defects related to the gate oxide at the beginning of the chapter conveys an idea of the physical origin of different defects.

2.1 PHYSICAL ORIGINS AND CHARGE EXCHANGE MECHANISMS

Typically, the high-κ gate oxide for III-V MOSFETs is grown by an atomic layer deposition (ALD) process, where alternating pulses of precursors form self-limiting atomic layers on the semiconductor surface [80]. Despite several years of research, the growth of high-κ oxides on III-V materials still gives rise to a large number of material defects which can interact with carriers in the semiconductor channel [81]. In their equilibrium state, such *traps* can either act as *donors* and emit additional electrons into the channel or as *acceptors* and capture electrons from the channel. In either case the charge state of the respective trap can change to more positive or more negative [82, 83].

Based on the distance of these defects from the high-κ/semiconductor interface and thus related to their chemical nature, the defects can be loosely divided into border traps in the 'bulk' of the oxide and interface traps [84]. To be able to interact with charge carriers in the semiconductor channel, the

energy levels of the defects have to be close to the semiconductor Fermi level. For n-type operation, as investigated in this thesis, this means that the defect energy levels have to be located in the semiconductor band gap or below to affect the off-state, and in the band gap or above to affect the on-state. According to density functional theory (DFT) calculations, there are several different material defects which lie in this range. Some typical candidates in the In(Ga)As/Al_2O_3/HfO_2 material systems in this thesis are summarised in the following and schematically illustrated in Fig. 2.1.

Border traps are typically associated with oxygen vacancies or interstitials at least a few atomic monolayers away from the oxide/semiconductor interface and their energy levels have been calculated to lie close to or above the In(Ga)As conduction band edge [82]. Border traps can interact with electrons from the semiconductor via quantum-mechanical tunnelling combined with an activation energy due to a local deformation of the trap site. If the charge state of a trap changes, the local electrostatic interactions will change and the surrounding atoms will either be drawn closer together or away from each other. Since such a structural rearrangement requires energy, a corresponding energy barrier can be associated with the capture/emission process [85]. The capture/emission time constants of such interactions can be expressed as [86–88]

$$\tau_{c/e} = \tau_0 \exp\left(\frac{x}{\lambda}\right) \exp\left(\frac{E_A}{k_B T}\right), \tag{2.1}$$

where x is the tunnelling distance, λ is the tunnelling attenuation length according to the WKB approximation, and E_A is the activation energy of the process. The effective pre-factor $\tau_0 = 1/(\sigma n v_{th})$ depends on the capture cross section σ, the carrier concentration n in the semiconductor, and the thermal velocity v_{th} of the charge carriers [86]. Due to a wide range of reported values for σ even for the same material systems, values for τ_0 can be found to vary from microseconds [89] all the way [90] to below picoseconds [91]. Especially for 'slow' (> nanoseconds) capture and emission processes, it is often assumed that the activation energy dominates the time constant.

Interface traps, as the name indicates, originate from imperfect interfaces between the III-V semiconductor and the gate oxide. Causes for interface traps with energies above the InAs or InGaAs conduction band edge can be certain faulty III-V bonds (e.g. As-As) or dangling bonds [92]. Oxygen vacancies at the interface can cause states above the conduction band edge of the semiconductor or below the valence band edge [83]. Due to the amorphous nature of the gate oxide, energy states of different defects are not localised in energy, but smeared out into distributions, see Fig. 2.2. The tails of these distributions can extend into the semiconductor band gap even from defect states with energies above or below the band gap [94]. Different III- or

V-surface oxides as well as As_{Ga} or Ga_{As} anti-sites can add additional states in the semiconductor band gap [93]. Furthermore, in the oxide outside of the semiconductor band gap, so-called metal-induced gap states (MIGs) or virtual gap states can occur due to electron wave functions from the semiconductor decaying into the oxide. There cannot be any MIGs inside the semiconductor band gap [94].

The expression to describe the interaction time constants $\tau_{c/e}$ for interface traps takes the same form as (2.1) without the tunnelling term, and E_A in this case also takes into account the distance in energy of the trap energy level from the channel Fermi level [95]. At this point it should be noted that the nomenclature for electrically active interface defects, the energies of which lie outside of the semiconductor band gap, can easily become ambiguous, since such states have to be at least a short distance away from the actual interface in order to be able to capture charge. Without this certain distance, defect sites right at the interface would be in contact with the continuous density of states of the semiconductor bands and there would be no barrier to keep charge carriers trapped.

Many studies have been carried out to reduce the magnitude of the defect distributions or to shift them out of the energy range, which is accessible during normal MOSFET operation. Different techniques of III-V surface passivation by means of forming gases have been proposed to reduce the interface state density [96] and the observation of the so-called self-cleaning effect of the ALD process [97] has lead to the introduction of the Al_2O_3/HfO_2 bilayer gate oxide used in the devices in this thesis [98]. Recent findings

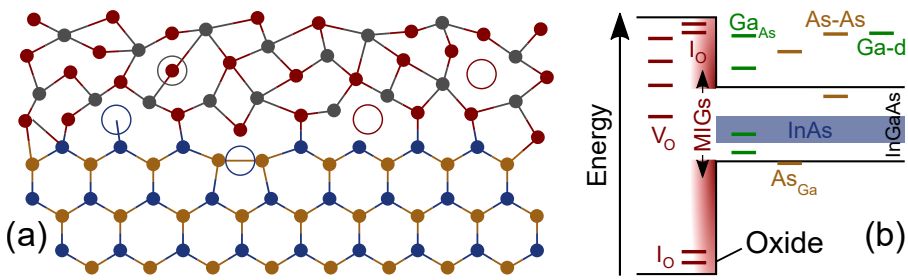


Figure 2.1: (a) 2D illustration of HfO_2 on InAs. Circles mark schematic example defects, such as oxygen vacancies (red circles), oxygen interstitials (grey circle), an As-As bond, and a dangling bond (blue circles). (b) Approximate relative positions of different defect energy levels in the HfO_2/In(Ga)As material system calculated by DFT. (Values indicated according to [82, 83, 92–94], but not to scale.) V_O – oxygen vacancies, I_O – oxygen interstitials, Ga_{As}/As_{Ga} – anti-sites, As-As – As-As bond, Ga-d – Ga dangling bond.

suggest that the Al_2O_3 might in fact not be necessary and the self-cleaning effect should work just as well with HfO_2 applied directly on top of the III-V channel [99]. Alternatively, the interface could be passivated remotely by the gate metal [100]. In order to tackle border traps, the introduction of dipole layers in the 'bulk' of the oxide has been proposed as a promising approach to shift the energy levels of the traps [101]. Since investigations in different directions constitute an active area of research, there still does not seem to exist a consensus on what would be the most promising strategy to move forward. Such investigations, however, just as the summarised results of DFT calculations, are not a part of this thesis and are only mentioned to provide an overview.

2.2 HYSTERESIS IN IV CURVES

One of the first measurements typically carried out on a transistor is the transfer sweep explained in Section 1.5 and already in this basic measurement, the effect of traps can be discerned. Fig. 2.2(a) presents a transfer curve, where V_{GS} is first 'swept' up from a certain starting point and upon reaching its highest value, immediately swept back down. Often, the two resulting IV curves do not exactly overlap each other, which is called *hysteresis*, and in the case of MOSFETs is ascribed to the charging of oxide traps. Due to the different interaction time constants (2.1) and energy levels of the traps, the width of the hysteresis depends on the sweep rate, the start and stop

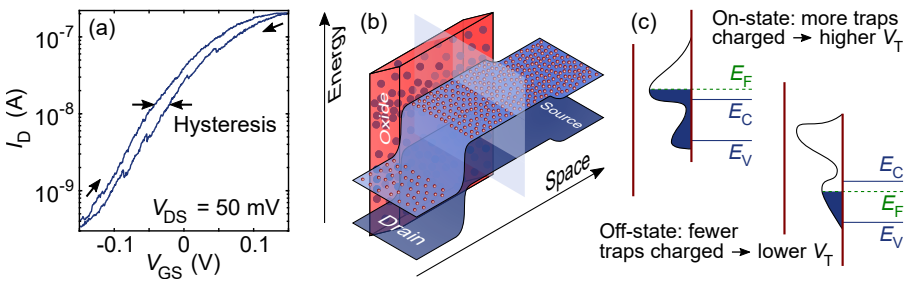


Figure 2.2: (a) Example of hysteresis in the transfer curve of a vertical nanowire TFET. In the down-sweep, individual discharging events can be discerned, which allow the quantification of the effect of individual defects and thus of single electrons. (b) 3D schematic to illustrate the relative position of oxide defects in space and energy. Red dots and blue orbs represent electrons and traps, respectively. (c) Cross section along the indicated plane in (b). The oxide defects are summarised to distributions and the different amounts of charged defects (blue areas) explain the observation of hysteresis.

voltages, and the temperature. For the most straightforward process, where traps in the oxide interact with electrons in the channel, with increasing V_{GS}, the energy levels of more and more traps end up below the Fermi level of the channel and if the electron capture time constant (2.1) is shorter than the measurement time, electrons can be captured in these traps. If the emission time constant for trapped electrons is then longer than the measurement time, the electrons remain trapped in the oxide for some time even when the Fermi level in the channel drops below the trap energy level again when V_{GS} is swept down. The trapped charge in the oxide causes a V_T shift so that during the down-sweep, at the same V_{GS}, a different I_D is measured than during the up-sweep. In the smallest devices, the capture and emission of individual electrons can be discerned from distinct steps in the IV curve such as in the example of Fig. 2.2(a). From such steps, the contribution of individual traps to the hysteresis can be determined to amount to a few millivolts. This is demonstrated in Paper IV and Paper V and will be elucidated further in Section 2.5.

In Paper I we investigated, how different crystal surfaces of a lateral nanowire affect the hysteresis. The data points in Fig. 2.3(a) were measured by systematically varying the gate stop voltage of the transfer measurements for a fixed start voltage and a fixed sweep speed. The effective (i.e. projected to the gate oxide/semiconductor interface) concentration of traps N_{eff} can then be calculated from the measured hysteresis ΔV_T via the charged traps ΔN_{eff} corresponding to each ΔV_T by

$$\Delta N_{eff} = \Delta V_T \, C_{ox} / q = \int_{\psi_{start}}^{\psi_{stop}} N_{eff}(E) \, dE, \tag{2.2}$$

where ψ_{start} and ψ_{stop} are the surface potential at the start and the stop voltage, respectively, and C_{ox} is the geometric oxide capacitance as before. N_{eff} was modelled by two Gaussian distributions and the parameters of these distributions were used as fitting parameters to reconstruct the measure ΔV_T. The resulting N_{eff} are presented in Fig. 2.3(b).

The effect of traps can be distinguished even more clearly by increasing the sweep voltage not continually, but by fast, subsequent pulses, as indicated in the inset of Fig. 2.3(c). In between the pulses with successively increasing magnitude, the voltage always drops to a resting value for a typically much longer time than the pulse lengths themselves. With short pulses, defects with trapping time constants, which are longer than the pulse lengths, cannot trap charge carriers anymore so that there is a smaller shift in V_T as compared with a continuous sweep. This results in a smaller hysteresis and, as demonstrated in Fig. 2.3(c), in higher currents. Furthermore, during the time intervals at the resting voltage, charged traps can release carriers again, which reduces some

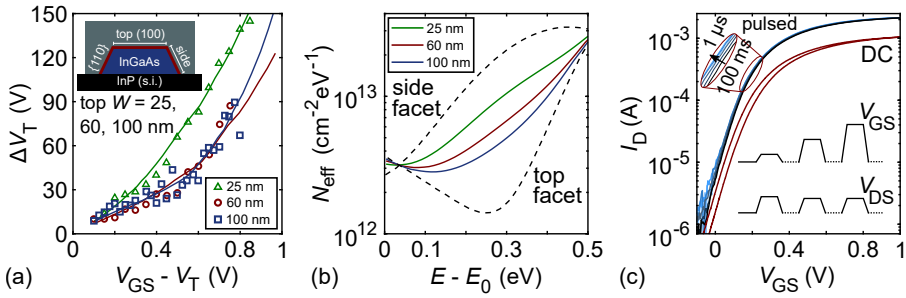


Figure 2.3: (a) Measured (symbols) and calculated (lines – fitted by N_{eff} in (b)) hysteresis ΔV_{T} as a function of increasing gate overdrive voltage $V_{\text{GS}} - V_{\text{T}}$ for three different channel geometries. The inset illustrates the channel dimensions. (b) Fitted effective defect density N_{eff} as a function of the energy E above the top of the channel barrier E_0. (c) Results from pulsed measurements on a planar InGaAs MOSFET with Al_2O_3 gate oxide. The pulsed transfer curves are clearly higher than the DC curve and the magnification resolves further differences between curves with different pulse lengths. The inset in the bottom right illustrates the pulsed measurement scheme; the wait times in between pulses are much longer than the pulses themselves.

of the accumulated V_{T} shift. The establishment of a systematic measurement scheme for pulsed measurements in the *Nanoelectronics* group was initiated during work for this thesis.

Fig. 2.3(c) presents an example of a planar InGaAs MOSFET with an Al_2O_3 gate oxide, which was used to establish the measurement scheme. In initial results on vertical nanowire MOSFETs it seems that besides electron trapping, as described before, other effects play a role as well; for example trapping of holes or the interaction of traps not only with the semiconductor channel, but also with the gate metal. Initial results from reliability measurements which corroborate these deliberations will be described in Section 2.6. All effects described in this section have also been observed in the output characteristics and can be explained by charge trapping due to V_{DS} in addition to V_{GS}.

2.3 MANY TRAPS – LOW-FREQUENCY NOISE

All transistor measurements can be subject to different forms of noise. In the IV curve in Fig. 2.2(a) for example, noise can be identified as small ripples throughout the curve. A noise source present in all electrical devices is thermal noise with a noise power spectral density $S_{\text{I}} = 4k_{\text{B}}T/R$ for a conducting resistance R. Due to the tunnelling mechanism with a low tunnelling probability $T_{\text{BTB}} \ll 1$, especially TFETs are also prone to shot

noise with $S_I = 2qI$, which depends on the device current I. Both of these mechanisms produce frequency-independent white noise. Besides these two, another important noise contribution originates from the same defects that can cause hysteresis. This contribution is often summarised under the term *low-frequency noise* (LFN) and the basis for one of the most widely used models to explain LFN as caused by oxide traps was suggested by A. L. McWhorter in 1957 [102].† The principles of this model are discussed below and a detailed treatise of LFN including new insights since 1957 can be found in [103].

Traps in the oxide, which are close to the position of the Fermi level in the channel, can continuously change their occupancy. The resulting continuous change of local electrostatics causes fluctuations in the channel energy bands and thus in the flatband voltage V_{fb}, which in turn causes fluctuations in the current. This noise origin is referred to as *generation-recombination noise* (g-r noise) and forms the basis of the McWhorter model. An important assumption in this early model is the trapping and de-trapping of electrons in defects by pure elastic tunnelling rather than by processes involving activation energies. With this assumption, (2.1) is reduced to $\tau = \tau_0 \exp(x/\lambda)$. The adequacy of this model in the light of more recent findings will be discussed in Section 2.4. Under this for now tacit assumption, and the assumption of

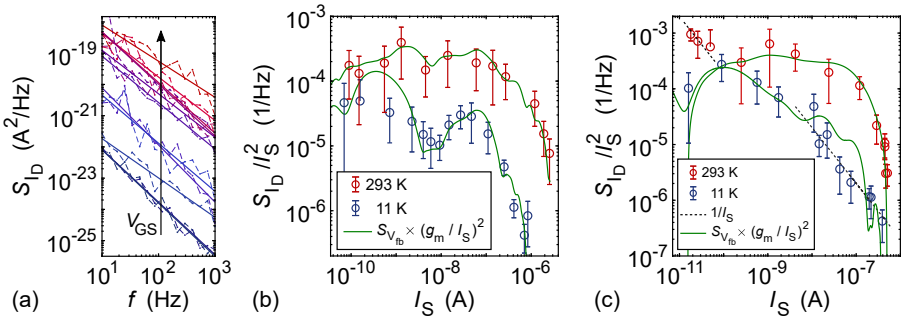


Figure 2.4: (a) Noise power spectral density S_{I_D} as a function of frequency for different V_{GS}. The slopes typically vary around –1, but are steeper in this example. This points towards a low absolute number of traps. (b) S_{I_D}/I_S^2 as a function of the source current I_S for a MOSFET at 293 K and a 1 K. At both temperatures, S_{I_D} follows the transconductance g_m^2 well, which indicates number fluctuations as the dominant noise source. (c) Same as (b) for a TFET. In some areas, S_{I_D} seems to follow $1/I_S$ instead of g_m^2, which indicates a stronger contribution of mobility fluctuations in the channel.

†Although many works concerned with LFN reference this book, it seems to be very difficult to get a hold of it. Thus, instead, [103] is recommended here for the interested reader.

a uniform trap distribution $N_{\mathrm{bt}}(x, y, z, E)$, the total power spectral density $S_{\mathrm{I_D}}$ of the noise in the device current can be calculated by integrating the power spectral density of the fluctuation in the occupancy of a single trap over both energy and space. Besides the immediate effect of fluctuations in the flatband voltage on the device currents, called *number fluctuations*, the fluctuations in local electrostatics due to charge trapping can change the amount of scattering in the channel. This can lead to fluctuations in the carrier mobility, which are called *correlated mobility fluctuations* (CMF). We investigated the importance of this CMF contribution in III-V MOSFETs in Paper II. The complete expression for $S_{\mathrm{I_D}}$, including both contributions, and normalised by the device current squared I_{D}^2, becomes [103, 104]

$$\begin{aligned} \frac{S_{\mathrm{I_D}}}{I_{\mathrm{D}}^2} &= \underbrace{\frac{q^2 k_{\mathrm{B}} T \lambda N_{\mathrm{bt}}}{f^{\gamma} L_{\mathrm{G}} W_{\mathrm{G}} C_{\mathrm{g}}^2}}_{S_{\mathrm{V_{fb}}}} \left(1 + \alpha \mu_{\mathrm{eff}} C_{\mathrm{g}} \frac{I_{\mathrm{D}}}{g_{\mathrm{m}}}\right)^2 \frac{g_{\mathrm{m}}^2}{I_{\mathrm{D}}^2} \\ &= S_{\mathrm{V_{fb}}} \left(1 + \alpha \mu_{\mathrm{eff}} C_{\mathrm{g}} \frac{I_{\mathrm{D}}}{g_{\mathrm{m}}}\right)^2 \frac{g_{\mathrm{m}}^2}{I_{\mathrm{D}}^2}, \end{aligned} \tag{2.3}$$

where C_{g} is the series combination of the oxide capacitance, the quantum capacitance, and the centroid capacitance, μ_{eff} is the effective carrier mobility, α describes the change in mobility due to charge trapping in the oxide, and the frequency exponent γ takes into account non-uniformities of N_{bt} with respect to the depth in the oxide. Since a larger number of charge carriers in the channel screens the effect of charged traps on the mobility, α depends on the device current I_{D} as $\alpha = \alpha_0 - \alpha_1 \ln(I_{\mathrm{D}})$, where α_0 and α_1 are fitting parameters [105]. α is typically in the order of 10^3–10^4 Vs/C [103].

Besides the model (2.3), which relates $S_{\mathrm{I_D}}$ to oxide traps and CMF, another widely used model explains LFN solely in terms of mobility fluctuations. This empirical model was developed by F. N. Hooge in 1969 [106] and expresses $S_{\mathrm{I_D}}$ as [103, p. 78]

$$\frac{S_{\mathrm{I_D}}}{I_{\mathrm{D}}^2} = \frac{q \alpha_{\mathrm{H}} \mu_{\mathrm{eff}} V_{\mathrm{DS}}}{f L_{\mathrm{G}}^2 I_{\mathrm{D}}}, \tag{2.4}$$

where α_{H} is the empirical and dimensionless Hooge parameter. A weak point of this model is the lack of a widely accepted physical explanation of the origins of the mobility fluctuations [103, p. 82 ff.].

In either case, (2.3) or (2.4), due to the $1/f^{\gamma}$ dependence ($\gamma = 1$ in the Hooge model), this type of noise is often referred to as $1/f$ *noise*. It can be very harmful to e.g. mixer circuits, if the $1/f$ noise of the transistor overlaps the baseband so that the noise is mixed up together with the actual signal. The frequency dependence of the power spectral density $S_{\mathrm{I_D}}$ of $1/f$ noise can be resolved with e.g. a spectrum analyser or a lock-in amplifier, as demonstrated in Fig. 2.4(a). In Paper III we demonstrated that (2.3) and

(2.4), which were originally derived for MOSFETs, can be applied to TFETs as well. So if instead of as a function of the frequency, S_{I_D} is measured at a fixed frequency – typically 10 Hz – the dominant origin of the noise in MOSFETs and TFETs can be identified by comparing the measured S_{I_D} with (2.3) and (2.4). This is exemplified in Fig. 2.4(b) and (c) for a vertical nanowire MOSFET and a vertical nanowire TFET, respectively, each at two different temperatures. In the figures, the device current is expressed as the source current I_S because of the configuration of the measurement setup, which is described in Appendix A. For both MOSFETs and TFETs, S_{I_D}/I_S^2 mostly follows the normalised transconductance g_m^2/I_S^2, which indicates oxide traps as the dominant origin of the noise rather than mobility fluctuations. Both types of LFN can occur in the same device, however, depending on e.g. channel geometry or bias range. In [107] for example, this distinction is very clear for conduction at the surface or in the core of a nanowire in different bias ranges. For TFETs, the tunnelling mechanism, which determines the current level, is unlikely to be subject to mobility fluctuations. Instead, in Fig. 2.4(c) and Paper III, the mobility fluctuations in certain bias ranges could have been induced in the channel after the tunnel junction, especially since the channels of these devices were very long ($\gtrsim$ 200 nm).

If oxide traps are identified as the dominant source of LFN, the measured S_{I_D} can be used to calculate the density N_{bt} of these traps according to (2.3). Fig. 2.5(a) presents a comparison of LFN in different III-V MOS architectures together with reported values for Si. The figure is an augmented version of the figure from Paper II, where we compared LFN for different channel geometries and their effects on the observed LFN in MOSFETs. In summary, short channel nanowire MOSFETs were affected mostly by number fluctuations, whereas long channel planar MOSFETs displayed a considerable amount of CMF as well. This may be attributed to two factors: In long channel devices, the current is determined by the mobility so that fluctuations in the mobility have a large effect, whereas the mobility is less important in short channel semi-ballistic devices; in addition, the larger surface-to-volume ratio in the short channel GAA devices can increase the ratio of number fluctuations to mobility-related effects. These different behaviours can be discerned in Fig. 2.5(a), where two different curves/values are provided for each device to separate the effects of a distributed N_{bt} and of CMF. The U-shaped curves represent distributed N_{bt} for the case of pure number fluctuations without any CMF, and the constant values in the insets represent the case of constant N_{bt}, where only the CMF are dependent on V_{GS}. The 'real' N_{bt} in most cases is very likely a combination of both, which can be deduced amongst others by examining the extreme values in Fig. 2.5(a). The highest values for the U-shaped curves exceed the amount of atoms in the

oxide, which is unphysical. The lowest values for the constant N_{bt} are lower than most reported literature values for Al_2O_3/HfO_2, even when compared with dedicated optimisation studies. This is unlikely, since the oxide in the measured devices was not explicitly optimised for low defect densities. The contributions of pure number fluctuations and CMF can however not be distinguished quantitatively in this kind of measurement.

As to the difference between III-V and Si material systems, the latter usually displays a trend to achieve lower N_{bt} values than the former and three major reasons seem probable to explain this. Fundamentally, Si channels exhibit a larger density of states so that the Fermi level does not move as far up into the conduction band in the on-state as it does in III-V channels. This means that for a given trap distribution in the gate oxide, a smaller range of this distribution and thus a smaller total number of traps is probed in Si devices. In addition, the band offset between a Si channel and the gate oxide is different from the offset for III-V channels so that different portions of the trap distributions are accessible to begin with. Thirdly, the Si fabrication process tolerates much higher annealing temperatures than III-V processes, which can

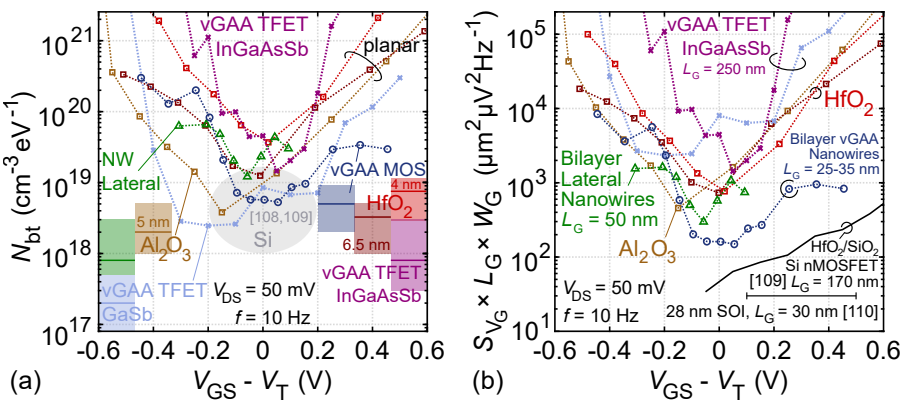


Figure 2.5: Comparison of different III-V and Si MOSFETs. (a) Border trap densities N_{bt}. U-shaped lines were calculated assuming pure number fluctuations (i.e. without CMF), straight lines with shaded standard deviations were calculated assuming constant N_{bt} with V_{GS}-dependent CMF. The constant N_{bt} are plotted in small intervals for better readability of the figure. (b) Input-referred equivalent gate voltage noise power S_{V_G}. Same colours and symbols in (a) and (b) correspond to the same devices. All III-V devices were measured during the work for this thesis and values for Si are from [108–110], where values from [110] were calculated with additional data from [111]. Squares denote planar devices, pyramids lateral nanowires, circles vertical nanowire MOSFETs, and crosses vertical nanowire TFETs.

lead to a better passivation of material defects. The most promising approach to date to remedy the effect of oxide traps in III-V material systems seems to be the introduction of a dipole layer in the oxide to shift the energies of the defects out of the operational range, as mentioned before [101].

It is somewhat striking that in Fig. 2.5(a), the values for N_{bt} for TFETs with and without CMF differ so strongly despite the current in TFETs being controlled by the tunnelling junction and not by effects limited by the carrier mobility. However, all of the TFETs in the figure featured long channels after the tunnel junction ($\gtrsim$ 200 nm) so that the CMF could have been induced after the tunnelling junction. One of the TFETs in Fig. 2.5(a) further sticks out insofar as it displays the lowest N_{bt} of all measured devices. This was confirmed in other devices on the same sample and is most likely related to processing. Of all the samples included in Fig. 2.5, this was the only one, where the gate oxide was applied as the first processing step and was not altered any further in later steps. The resulting low values for N_{bt} suggest that it would be worthwhile to adjust processing of other samples according to this scheme.

Fig. 2.5(b) presents the equivalent gate voltage noise $S_{\mathrm{V_G}} = S_{\mathrm{I_D}}/g_{\mathrm{m}}^2$, which transfers the measured current noise power $S_{\mathrm{I_D}}$ to the input, i.e. the gate voltage of the transistor. Different from N_{bt} in Fig. 2.5(a), this is a technical rather than a physical metric in the sense that it provides information about the smallest possible input voltage amplitude e.g. for amplifier design, rather than information about the physical properties of the transistor. As such, it does not depend on the assumption of a specific capture/emission mechanism. It becomes obvious that this technical metric does not simply 'translate' the physical defect properties of the oxide, when noting that the devices with the lowest N_{bt} do not necessarily exhibit the lowest $S_{\mathrm{V_G}}$. This is because devices with comparatively high N_{bt} can still exhibit high transconductances g_{m}. As a consequence, a smaller input noise $S_{\mathrm{V_G}}$ would be 'required' to cause the same current noise $S_{\mathrm{I_D}}$ that is generated by traps.

2.4 ON ELASTIC AND INELASTIC TUNNELLING IN LFN

The calculation of N_{bt} in Fig. 2.5(a) is based on the McWhorter model, which assumes elastic tunnelling as the charge trapping mechanism. Some results, especially from reliability measurements, (see Section 2.6) refute this assumption as too simplistic [88]. In this section, restrictions and prospects of the assumption of elastic tunnelling are briefly discussed.

For capture and emission processes, which are exponentially dependent on an activation energy, as in (2.1), $S_{\mathrm{I_D}}$ should vanish almost completely at temperatures as low as 11 K. Yet Fig. 2.4(b) and (c) both evidence a temper-

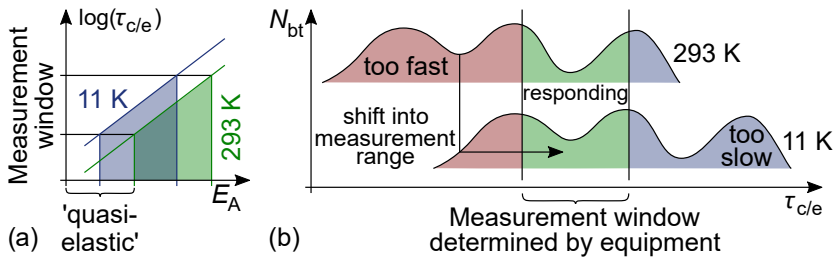


Figure 2.6: Schematic illustration how different portions of traps with a wide distribution of time constants $\tau_{c/e}$ respond at different temperatures. (a) According to an inelastic model, the trap time constants depend exponentially on the activation energy E_A. In a wide distribution of E_A, some E_A can be very small so that they trap charge 'quasi-elastically'. (b) At room temperature, these quasi-elastic traps respond too fast to be measurable, but they can shift into the accessible measurement range at low temperatures. The traps which are measurable at room temperature, on their part, shift out of the measurement range at low temperatures. If the two trap distributions with different activation energies are similar, similar LFN results are measured at the different temperatures.

ature dependence of S_{I_D}, which is linear at best, rather than exponential. In [112], an explanation for such a missing strong temperature dependence was offered for increased measurement temperatures. Different characterisation techniques suggest that trap activation energies, and thus the interaction time constants according to (2.1), are distributed over a very wide range. However, only a certain portion of traps in these wide distributions can respond in a certain measurement window at a certain temperature, while traps with time constants, which are longer than the measurement time, cannot respond. If the measurement temperature is increased, these traps with long time constants can respond faster so that at elevated temperatures, they are shifted into the measurement window. At the same time, the previously measured trap responses become so fast that they shift out of the measurement window. The measured effects due to traps can thus look very similar at different temperatures, while actually resulting from different responding trap distributions.

Under the contrasting assumption of elastic tunnelling, which leads to (2.3), a linear temperature dependence of S_{I_D} is predicted. This agrees well with the measured results in Fig. 2.4 and suggests that the line of argument for increased measurement temperatures might also hold for decreased temperatures. In a wide distribution of activation energies there can be traps with very small activation energies, which thus respond too fast to measure at e.g. room temperature. If the measurement temperature is decreased, the response

of these fast traps slows down, moves into the measurement window, and the measured results are again very similar at the different temperatures. This is illustrated in Fig. 2.6. For very low activation energies, however, charge carriers interact almost elastically with the corresponding traps so that the nomenclature in this context can be ambiguous. The adaptation of the explanation in [112] to lower temperatures is strongly supported by e.g. [90], where trap populations with low activation energies were identified at low temperatures, and it is consistent with our observations of trap responses even at gigahertz frequencies in Chapter 3/Paper VI. Furthermore, in [112], the absence of V_T shifts below 150 K was drawn upon as an argument against any elastic mechanisms as well, while we clearly observe V_T shifts even at temperatures as low as 11 K.

To be on the safe side, the calculated N_bt in Fig. 2.5(a) should be interpreted as minimum values. The involvement of activation energies can only increase the interaction time constants $\tau_\mathrm{c/e}$ so that if subject to an activation energy, the same measured $\tau_\mathrm{c/e}$ would correspond to shorter tunnelling distances and thus to higher N_bt. At temperatures as low as 11 K, however, the calculation of N_bt based on elastic tunnelling should yield reliable results. Typically, values for N_bt at these temperatures were about a factor 10–100 as low as at room temperature so that LFN measurements at several intermediate temperatures could provide information about the distribution of N_bt in energy. Whichever the 'real' interaction mechanisms may be, a thorough LFN characterisation is important for every transistor technology, since, like other noise sources, LFN can have detrimental effects on circuit performance [103].

2.5 INDIVIDUAL TRAPS – RANDOM TELEGRAPH NOISE

In strongly scaled devices, the absolute number of defects can be very small so that in some measurements, the effects of individual traps can become visible. This will happen, if these individual traps have capture and emission time constants similar to the integration time of each measurement point. An example of such individual traps was illustrated already in Fig. 2.2(a), and Fig. 2.7(a) provides an exploded view of the resulting distinct steps.

The strongly scaled III-V devices characterised in this thesis typically approach the quantum capacitance limit, where the channel potential follows the gate voltage almost one-to-one except for a possible potential drop due to interface states. From simple electrostatic considerations, the relation between V_GS and the surface potential ψ_s can be expressed as

$$\psi_\mathrm{s} = \frac{C_\mathrm{ox}}{C_\mathrm{ox} + C_\mathrm{q} + C_\mathrm{it}} V_\mathrm{GS}, \tag{2.5}$$

where C_{ox} is the gate oxide capacitance, C_q is the semiconductor quantum capacitance, and $C_{it} = qD_{it}$ is the capacitance due to an interface defect density D_{it}. For (2.5) it is assumed that C_{ox} and C_q are in series and C_{it} is in parallel with C_q. As highlighted in Fig. 2.7(a), a defect trapping a single electron increases the hysteresis by a few millivolts so that according to (2.5), the channel potential changes by a similar amount, the value somewhat reduced depending on C_{it} and C_q.

In Paper IV we demonstrated that the individual traps affecting the curve in Fig. 2.7(a) deteriorate the inverse subthreshold slope S of a transistor. Often, such a degradation is mostly attributed to interface defects D_{it} as

$$S = \ln(10)\frac{k_B T}{q}\left(1 + \frac{C_q + C_{it}}{C_{ox}}\right). \tag{2.6}$$

Since border traps can deteriorate S as well, the exclusive focus on its optimisation by improving the interface is not sufficient and border traps have to be addressed as well.

If the transistor current I_D is measured continuously and with a sufficient time resolution at a fixed bias point in close proximity to a distinct capture/emission step in the transfer curve, a phenomenon called *Random Telegraph Noise* (RTN) can be observed. In this phenomenon, the current fluctuates between two distinct levels due to the trapping and de-trapping of electrons in one specific defect. An example of such a measurement is presented in Fig. 2.7(b). Since the trapping and de-trapping follows a Poisson

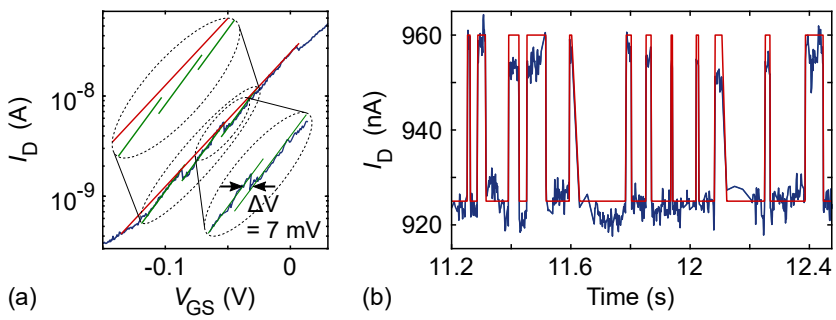


Figure 2.7: (a) Demonstration of how individual oxide traps can deteriorate the inverse subthreshold slope of a TFET. The individual slopes close to the distinct steps are mostly the same. A slope over a wider bias range, however, is clearly flatter than the individual slopes. (b) Example of RTN (different device). At a fixed bias point close to a distinct step in the transfer curve, the device current fluctuates between two distinct levels due to repeated capture and emission of an electron in the same trap.

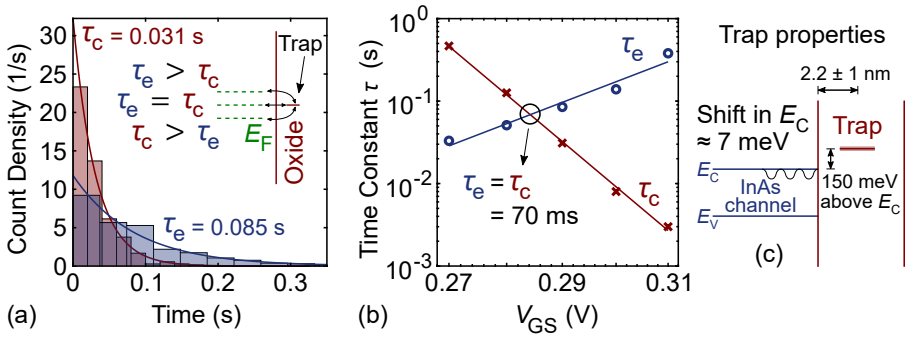


Figure 2.8: (a) Histogram of measured RTN at 11 K with fitted exponential distributions. The inset illustrates different positions of the trap energy level relative to the channel Fermi level for different V_{GS}, and the resulting relation between τ_c and τ_e. (b) Measured characteristic time constants as a function of V_{GS}; same measurement series as in (a).

point process, the distribution of the times in the high and the low current state can be described by an exponential distribution $f(t) = \tau^{-1}\exp(-t/\tau)$, where t is the time and τ the characteristic time constant of the process. An example of such distributions for the capture and emission time constants $\tau_{c/e}$ is provided in Fig. 2.8(a). From the measured distributions, the characteristic capture and emission time constants of an individual trap can be calculated. We carried out these experiments for TFETs in Paper IV and Paper V.

At cryogenic temperatures, where activation energies play a minor role, RTN measurements can be used to estimate the distance of the responsible trap from the oxide/channel interface as well as their energy level relative to the channel band edges. The energy level can be calculated from (2.5) and the transfer curve, and we observed trap responses from the subthreshold region all the way to the on-state of transistors, i.e. from inside the semiconductor band gap all the way into the conduction band. The estimation of the distance from the interface is illustrated in Fig. 2.8(b). With the assumption of elastic tunnelling, the capture and emission time constants are equal, if the trap energy level is aligned with the channel Fermi level. At this point, the tunnelling length and thus the distance of the trap from the interface can be calculated with (2.1). If the trap energy level moves below the channel Fermi level, the capture time constant will decrease and the emission time constant will increase, since now the trapped charge has to 'jump up' in energy to be released from the trap. (Vice versa, if the trap energy level moves above the channel Fermi level.) This offset between the trap energy level and the channel Fermi level, however, is equivalent to a small energy barrier. So, if in this interpretation charge carriers can surmount small energy barriers after

all, an inelastic trapping mechanism due to deformation of the defect site will have to be revisited for consistency.

In the example of Fig. 2.8(b), the time constants varied over a gate bias range of ±10–15 mV so that with (2.5) and $E = qV$ we will examine the effect of an activation energy of 15 meV in (2.1). With $E_A = 15$ meV, $\tau_c = \tau_e = 70$ ms from Fig. 2.8(b), and $\tau_0 = 1$ ps[†] in (2.1), the tunnelling distance is calculated as 1.2 nm. For pure elastic tunnelling instead, the distance would be 3.2 nm so that just from electrical measurements, the distance of the trap from the channel/oxide interface can be determined to within approximately 2 nm. The detailed understanding of individual defects is of increasing importance for ultra-scaled devices, which are affected by only a few traps. Electrical measurements as non-invasive techniques, which are applicable to complete devices rather than dedicated test structures such as capacitors, can provide a convenient means to this end.

2.6 RELIABILITY AND LIFETIME

Over time, the effects of charges trapped in the oxide can accumulate. Some defects even have interaction time constants, which are so long that their effects on the MOSFET characteristics appear as a virtually constant V_T shift. Since the integration of MOSFETs in circuits relies on fixed operating voltages, too large a V_T shift will prevent proper circuit operation. The average time until a transistor causes a failure in circuit operation is called the *lifetime* of the device and the failure criterion is often defined as an accumulated V_T shift of 30 mV over ten years.

Since measurements at such long time scales are impractical, so-called *bias temperature instability* (BTI) measurements are used to enable the extrapolation of the transistor lifetime. In such BTI measurements, electrical stress is applied to a transistor in the form of elevated V_{GS} or V_{DS} for repeated and increasing intervals of time, which leads to an increased capture of charge carriers in oxide traps, as described in Section 2.2. A schematic visualisation of this measurement scheme is presented in Fig. 2.9(a).

V_T shifts ΔV_T due to BTI stress in a wide range of MOSFETs were found to follow the general equation (e.g. (2.24) in [113, p. 47])

$$\Delta V_T \approx A_0 \exp\left(-\frac{E_A}{k_B T}\right)\left(\frac{|V_{GS} - V_{T,0}|}{t_{ox}}\right)^{\gamma} t_{stress}^{\alpha}, \tag{2.7}$$

where A_0 is a scaling pre-factor, E_A is the apparent activation energy of the ΔV_T process, $V_{T,0}$ is the threshold voltage before stress, t_{ox} is the gate oxide

[†] $\tau_0 \approx 1$ ps was chosen based on high frequency measurements, see Chapter 3.

thickness, t_{stress} is the stress time, and γ and α are the voltage acceleration factor and the stress time exponent, respectively. For the relaxation times in between stress periods, a *universal relaxation model* was found to describe the recovery of V_{T}. With the *universal relaxation time* $\xi = t_{\mathrm{relaxation}}/t_{\mathrm{stress,total}}$ and fitting parameters B and β, the V_{T} recovery $r(\xi)$ can be expressed as (e.g. (2.25) in [113, p. 49])

$$r(\xi) = \frac{A}{1 + B\xi^{\beta}}. \tag{2.8}$$

A in this expression is the V_{T} shift at the end of the previous stress period and is thus equal to (2.7), but with the total t_{stress} of all previous stress times added together. With these two relations, (2.7) and (2.8), measurements of ΔV_{T} for increasing stress times t_{stress} and the subsequent V_{T} recovery in relaxation intervals $t_{\mathrm{relaxation}}$ in between the application of stress should enable the extrapolation of the expected lifetimes of the transistors as well as conclusions about the responsible trap distributions.

BTI measurements on vertical III-V nanowire MOSFETs were carried out in a collaboration with V. Putcha and J. Franco at IMEC in Belgium as a complement to the previously described characterisations. At first glance, unfortunately, the measurements do not seem to obey (2.7) or (2.8) at all. The representative relaxation traces in Fig. 2.9(b) clearly and repeatedly change their slope and their sign so that their magnitudes with increasing stress times are not monotonous. This indicates that besides the 'normal' charge trapping, it is likely that different competing mechanisms are involved in these complex traces, such as a discharge of traps to the gate metal instead of the channel or the capture of holes from the valence band. A schematic summary of possible processes involved is provided in Fig. 2.9(c). With these considerations, (2.8)

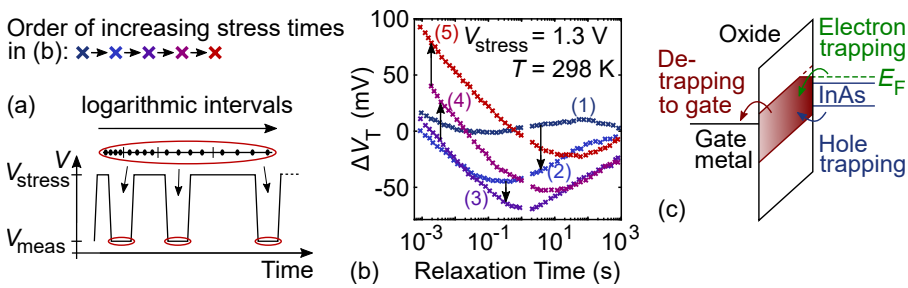


Figure 2.9: (a) BTI measurement scheme, adapted from [114]. The durations of the stress periods are typically increased logarithmically, while the duration of the relaxation periods is kept constant with logarithmically spaced measurement points. (b) V_{T} recovery measured on a vertical III-V nanowire MOSFET. (c) Possible mechanisms, which could explain the complex V_{T} recovery in (b).

can then be fitted to the measured data, if the total relaxation in Fig. 2.9(b) is separated into a positive and a negative component. From such fits, the physical origins, i.e. the involved trap distributions, can be determined, e.g. by a compact-physics framework, such as the one presented in [115] (*Comphy*).

2.7 OUTLOOK – TOWARDS A UNIFIED TRAP DISTRIBUTION

Although all of the previously analysed effects – hysteresis, LFN, RTN, BTI – manifest themselves in the operation of a MOSFET in different ways, all effects can be explained on the basis of trap distributions in the gate oxide. It seems likely then that there is a single common distribution, which causes all of the observed effects. In an ongoing collaboration with E. Caruso and P. Hurley from the Tyndall National Institute in Ireland, we are endeavouring to find a model for such a common defect distribution. If such a model was found, it would be interesting, in a more advanced attempt, to also include the high frequency trap responses studied in the chapter hereafter. Retroactively, this might lead to different trap distributions for the low frequency range again, since the very fast traps observed at gigahertz frequencies should always be able to respond at lower frequencies as well.

In this chapter, a frequency range from about 100 µHz (BTI) to about 10 kHz (LFN) was covered and hysteresis and RTN measurements are within this spectrum as well. Characterisations in the following chapter carry on from 10 MHz, which leaves a gap in the covered time constants of electrically active defects. Capacitance-voltage (CV) measurements, which typically range from below kilohertz to about a megahertz, can be a means to close this gap to some extent. They are, however, not covered in this thesis. Furthermore, the measurement range of the transconductance-frequency dispersion, which will be discussed in the following chapter, can be extended to frequencies in this gap by a lock-in amplifier as used for the LFN measurements. It is thus possible to electrically characterise oxide traps over the whole frequency range from microhertz to gigahertz.

3

Radio Frequency Characterisation

Due to their low effective mass and the resulting high injection velocities and low densities of states, III-V MOSFETs are of special interest for high frequency applications. This chapter summarises the work on radio frequency (RF) characterisation of (vertical) III-V nanowire MOSFETs carried out during work for this thesis. In the first section, the small-signal model will be presented, which was developed and used to explain different observations in the RF performance of III-V nanowire MOSFETs. In the second section, improvements to the design of vertical RF transistors, based on this model, will be reviewed, and in the last section, RF characterisations will be employed to examine some physical properties of TFETs. The measurement procedure for the small-signal parameters is described in Appendix A and a detailed treatise of the development of the model in Appendix B.

3.1 THE SMALL-SIGNAL MODEL INCLUDING GATE OXIDE DEFECTS

If for a voltage $v_{\mathrm{TOTAL}} = V_{\mathrm{DC}} + v_{\mathrm{ac}}$ the AC component v_{ac} is small with respect to the DC component V_{DC}, the effect of this AC component on the transistor currents i at the DC bias point can be approximated by the linear term of a Taylor expansion

$$i_{\mathrm{TOTAL}} \approx I_{\mathrm{DC}}(V_{\mathrm{GS}}, V_{\mathrm{DS}}) + \frac{\partial i_{\mathrm{ac}}}{\partial v_{\mathrm{gs}}} v_{\mathrm{gs}} + \frac{\partial i_{\mathrm{ac}}}{\partial v_{\mathrm{ds}}} v_{\mathrm{ds}}. \tag{3.1}$$

For this expression, the source terminal is assumed to be the common ground terminal, which will be maintained throughout this chapter. For such small AC variations, especially at high frequencies, it is sometimes sufficient to study only this linearised *small-signal* response of a transistor rather than the

complete *large-signal* response. The aim of RF small-signal modelling is then the expression of the electrical response of a device under test by an equivalent circuit consisting of linear elements.

At radio frequencies it is not practical to work with currents and voltages anymore. Instead, the linear electrical response of networks and devices is measured via complex-valued *scattering parameters* (*s*-parameters), which provide information about the amplitude and the phase of a signal. For such a characterisation in general, a device under test can be represented as a 'black box' with input and output ports, and the incident, reflected, and transmitted power is measured. For an n-port network this results in an $n \times n$ matrix of s-parameters, which describes the complete small-signal response of the network. For a detailed introduction to RF network analysis see e.g. [116].

A more intuitive and completely equivalent set of parameters is the set of *admittance parameters* (y-parameters), which will be used in the following analysis. In a common source configuration, a MOSFET has two ports, namely gate (port 1) and drain (port 2), as illustrated in Fig 3.1. The general definitions of the y-parameters for such a two-port are

$$y_{kl} = \left.\frac{i_k}{v_l}\right|_{v_{\neg l}=0} \quad \text{(3.2a)} \qquad \text{and} \qquad \begin{bmatrix} i_1 \\ i_2 \end{bmatrix} = \begin{bmatrix} y_{11} & y_{12} \\ y_{21} & y_{22} \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}, \quad \text{(3.2b)}$$

where all currents and voltages are the small-signal AC components of (3.1). In words, y_{kl} is the linearised small-signal admittance, which the current related to port k experiences as a function of the voltage at port l with the respective other voltage kept at zero volts. In the following, brief explanations of the different small-signal components will be provided, before the equivalent circuit model (Fig. 3.2) and the associated equations are summarised on a double page as a convenient overview for practical use. The model was submitted for publication as Paper VI and a more detailed discussion of its development can be found in Appendix B. Detailed derivations of the basic small-signal components can be found in e.g. [117].

At the heart of the small-signal model is the transconductance g_{m}, which was already introduced in Chapter 1 and which is essential for the gain of a transistor. In DC operation, g_{m} appears as a constant at each bias point, but in a frequency sweep, for III-V MOSFETs, g_{m} typically exhibits a dispersion of the form $\ln(\omega/\omega_0)$ due to the presence of gate oxide traps [118]. Here, ω is the angular frequency and ω_0 is the inverse of the effective pre-factor τ_0 in the expression for the capture/emission time constant $\tau_{\mathrm{c/e}}$ in (2.1). ω_0 is thus the angular frequency above which traps cannot respond to the input signal anymore. The movement of charge in the MOSFET channel is

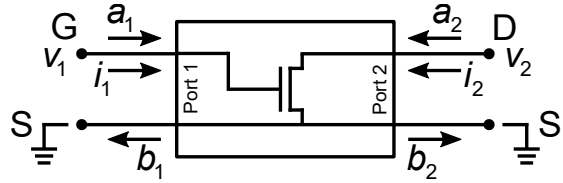


Figure 3.1: Illustration of a MOSFET as a two-port. G – gate, D – drain, S – source. For the actual *s*-parameter measurement, the different powers a_i and b_j are measured at the ports. For *y*-parameter modelling, the source is grounded and the currents i_k as functions of the voltages v_l are calculated.

expressed by capacitances, which are defined as $C_{kl} = (\text{-})\partial Q_k/\partial v_l$, where the (-) applies if $k \neq l$ [117, p. 397 ff.].† The mutual differential capacitance $C_\text{m} = C_\text{dg} - C_\text{gd}$ balances the charge between the two inversely related intrinsic capacitances C_dg and C_gd. Besides the intrinsic capacitances, so-called parasitic capacitances due to overlaps of conductive elements in the physical device structure affect the MOSFET small-signal response. In the following, intrinsic capacitances are denoted by a subscript 'i' and parasitic capacitances by a subscript 'p'. As described in more detail in Section 3.2 and similar to g_m, the intrinsic and parasitic capacitances are affected by the presence of oxide traps as well. The total resulting admittances, which take into account these effects, are denoted as y_gx and $y_\text{gx,p}$, where 'x' stands for 's' (source) or 'd' (drain). Additional constant conductances in $y_\text{gs,p}$ and $y_\text{gd,p}$ take into account DC leakage.

At low frequencies, the carrier response in the channel can be treated as instantaneous, but at operation frequencies approaching the inverse of the channel transition time, this assumption is not valid anymore. To a first degree, the resulting *non-quasi-static* (NQS) effects can be taken into account by intrinsic channel resistances R_i and R_j in series with the intrinsic capacitances [49, p. 364]. R_g, R_s, and R_d in Fig. 3.2 are the access resistances of the transistor and g_ds on the drain side is the output conductance as in Fig. 1.4. Two controlled current sources, denoted g_i1 and g_i2, take into account impact ionisation (II) and band-to-band tunnelling (BTBT), which can occur in III-V MOSFETs at high gate-to-drain electric fields [36, 119, 120].

The analytic expressions for the *y*-parameters become much easier, if first, the access resistances are subtracted from the measured data. In a well behaved transistor, R_g, R_s, and R_d (along with the parasitic capacitances)

†In the case of C_gs and C_sd this does not seem to make sense, if the source is kept at zero bias. From the beginning, the common-source (CS) configuration yields the capacitances C_gg, C_gd, C_dg, and C_dd, cf. (3.2b). C_gs and C_sd are the result of rewriting the CS model to a hybrid-π model as in Fig. 3.2.

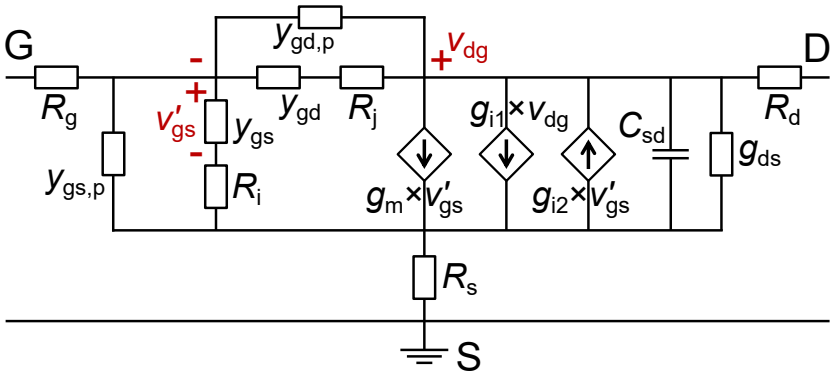


Figure 3.2: Comprehensive small-signal model as developed and used in this thesis. G, S, and D denote gate, source, and drain, respectively, and equations for all parameters are provided in the text.

can be determined from an s-parameter measurement in the off-state of the transistor, where the model in Fig. 3.2 is drastically simplified and only the access resistances and the parasitic capacitances remain. In the off-state, at high frequencies, the real parts of the impedance parameters (z-parameters), which are again equivalent to the s- and y-parameters, become

$$\boldsymbol{Z}_{\mathbf{R}} = \mathrm{Re}(\boldsymbol{Z}_{\textbf{off-state}}) = \begin{bmatrix} R_\mathrm{s}+R_\mathrm{g} & R_\mathrm{s} \\ R_\mathrm{s} & R_\mathrm{s}+R_\mathrm{d} \end{bmatrix}. \tag{3.3}$$

This matrix can simply be subtracted from the z-parameter matrix of the actual measurement, which is to be modelled. The remaining intrinsic y-parameters (i.e. without R_g, R_s, and R_d) according to (3.2a) are

$$y_{11} = \left.\frac{i_1}{v_1}\right|_{v_2=0} = y_{\mathrm{gs,p}} + \frac{y_\mathrm{gs}}{1+y_\mathrm{gs}R_\mathrm{i}} - y_{12}, \tag{3.4a}$$

$$y_{12} = \left.\frac{i_1}{v_2}\right|_{v_1=0} = -y_{\mathrm{gd,p}} - \frac{y_\mathrm{gd}}{1+y_\mathrm{gd}R_\mathrm{j}}, \tag{3.4b}$$

$$y_{21} = \left.\frac{i_2}{v_1}\right|_{v_2=0} = \frac{g_\mathrm{m}}{1+y_\mathrm{gs}R_\mathrm{i}} + y_{12} - g_{\mathrm{i1}} - \frac{g_{\mathrm{i2}}}{1+y_\mathrm{gs}R_\mathrm{i}}, \tag{3.4c}$$

$$y_{22} = \left.\frac{i_2}{v_2}\right|_{v_1=0} = g_\mathrm{ds} + \mathrm{j}\omega C_\mathrm{sd} - y_{12} + g_{\mathrm{i1}}, \tag{3.4d}$$

with equations for the different components as in (3.5)–(3.9). The y-parameters (3.4) possess an inherent universality in the sense that different aspects, such as the effect of border traps, II, or BTBT, which might be too detailed for some

applications, can be removed without further ado to obtain simpler expressions for all parameters. The detailed expression for the transconductance is

$$\begin{aligned} g_{\mathrm{m}} = & g_{\mathrm{m,i}} \left[1 + \gamma_1 \ln(\omega/\omega_0) + \mathrm{j}\alpha \left(1 + \gamma_2 \ln(\omega/\omega_0)\right)\right] \\ & - \mathrm{j}\omega \left(C_{\mathrm{m}} - C_{\mathrm{m},\omega} \ln(\omega/\omega_0)\right), \end{aligned} \tag{3.5}$$

where $g_{\mathrm{m,i}}$ is the intrinsic transconductance without the effect of traps, γ_1, γ_2, and $C_{\mathrm{m},\omega}$ are the parameters to describe the frequency dispersions due to traps, and α is a scaling factor for the dispersion of the imaginary part of the transconductance, since it is much smaller than the real part. The logarithmic expressions for the frequency dispersions can be obtained from a numerical solution of the trap model, which will be described in Section 3.2. The intrinsic admittances, when including the effects of traps, are

$$y_{\mathrm{gx}} = \omega \times g_{\mathrm{gx},\omega} + \mathrm{j}\omega \left(C_{\mathrm{gx,i}} - C_{\mathrm{gx},\omega} \ln(\omega/\omega_0)\right), \tag{3.6}$$

where 'x' stands for 's' or 'd' for source or drain, respectively, $C_{\mathrm{gx,i}}$ are the intrinsic capacitances without the effect of traps, and the dispersion due to traps is modelled by $g_{\mathrm{gx},\omega}$ and $C_{\mathrm{gx},\omega}$. Similar expressions describe the parasitic elements

$$y_{\mathrm{gx,p}} = \omega \times g_{\mathrm{gxp},\omega} + \mathrm{j}\omega \left(C_{\mathrm{gxp},0} - C_{\mathrm{gxp},\omega} \ln(\omega/\omega_0)\right) + g_{\mathrm{gx,l}}, \tag{3.7}$$

where $g_{\mathrm{gx,l}}$ additionally takes into account DC leakage and the other parameters are analogous to (3.6). Simple analytic expressions for the intrinsic NQS resistances R_{i} and R_{j} can be derived by an RC relaxation time approach by equating the time it takes carriers to traverse half the channel with the time to halfway charge $C_{\mathrm{gs,i}}(C_{\mathrm{gd,i}})$ through $R_{\mathrm{i}}(R_{\mathrm{j}})$. The results are

$$R_{\mathrm{i}} = \frac{1}{1.4\, g_{\mathrm{m,i}}} \tag{3.8a}$$

and

$$R_{\mathrm{j}} = \frac{1}{1.4\, g_{\mathrm{m,i}} \times C_{\mathrm{gd,i}}/C_{\mathrm{gs,i}}}. \tag{3.8b}$$

Finally, BTBT and II are modelled by

$$g_{\mathrm{i}k} = \frac{g_{k0}}{1 + \mathrm{j}\omega\tau_{\mathrm{i}}} \tag{3.9}$$

with the (for simplicity) common characteristic time constant τ_{i}. (Note that this time constant τ_{i} is different from the $\tau_{\mathrm{c/e}}$ which describes traps.) A sufficient number of constraints for fitting the different components of the small-signal model to the measured data is provided by the different real and imaginary parts as well as the low and high frequency ranges of the different y-parameters.

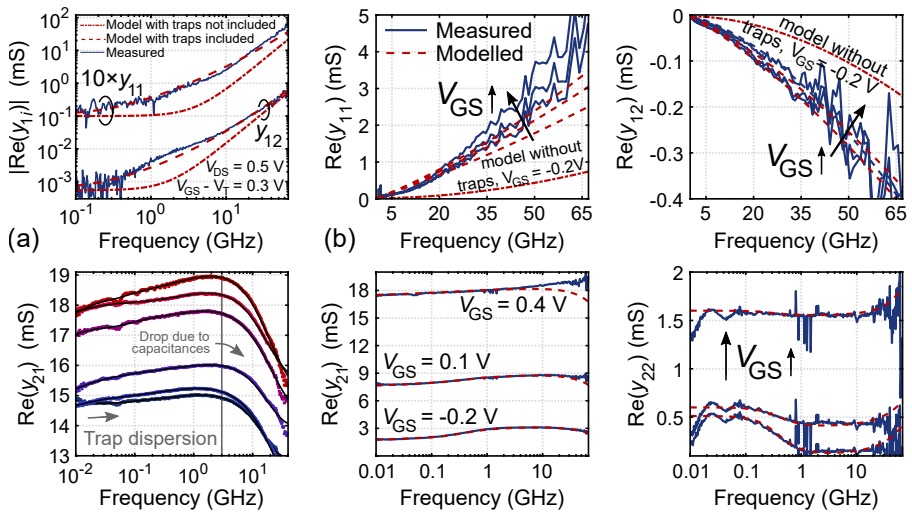


Figure 3.3: Real parts of the measured (blue, solid lines) and modelled (red, dashed lines) y-parameters for a vertical III-V nanowire MOSFET. (a) $\mathrm{Re}(y_{11})$, $\mathrm{Re}(y_{12})$ (same device), and $\mathrm{Re}(y_{21})$ (several different devices) with logarithmic frequency axes to highlight the dispersions due to traps. (b) Linear frequency axes (except $\mathrm{Re}(y_{22})$) to demonstrate how the model fits the full frequency range. The increase in the two lower curves of $\mathrm{Re}(y_{22})$ below about 1 GHz as well as the change in the slope of $\mathrm{Re}(y_{21})$ at about 0.1 GHz in the curves for $V_{\mathrm{GS}} = 0.1$ V and -0.2 V are caused by BTBT and impact ionisation.

As can be seen in Fig. 3.3 and Fig. 3.4, which are reproduced (and augmented) from Paper VI, (3.4)–(3.9) enable the accurate and comprehensive modelling of RF measurements on III-V (nanowire) MOSFETs and can thus provide detailed insights into the limiting and enabling factors of such transistors in terms of high frequency performance. The investigations of these factors during work for this thesis will be summarised in Section 3.3, but first, the origin of the frequency dispersions in the previous equations will be discussed in a dedicated section.

3.2 ON THE ORIGIN OF THE OXIDE TRAP COMPONENTS

The dispersions in the conductances in Fig. 3.3(a), in the capacitances in Fig. 3.4(b), and in the gains in Fig. 3.5 in Section 3.3 were observed in almost all measurements on vertical III-V nanowire MOSFETs and they can be explained by the presence of traps in the gate stack. For some of the parameters – the transconductance, the conductances, and the unilateral power gain – such

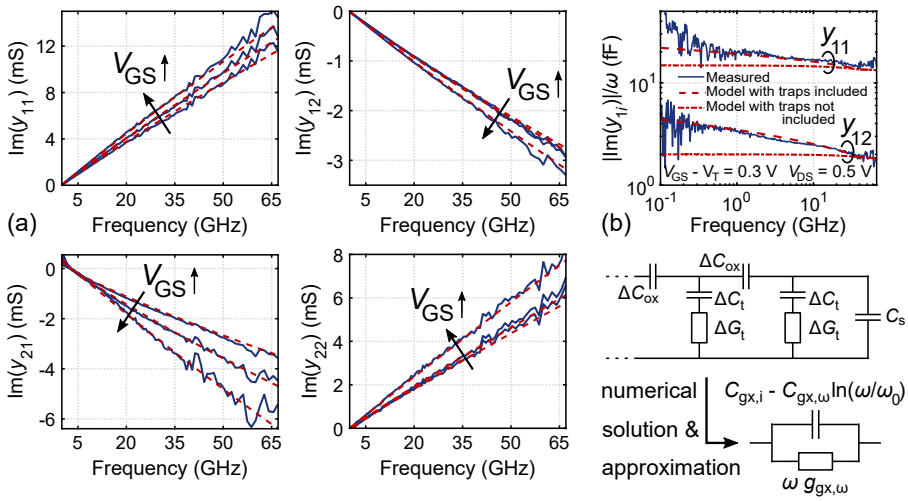


Figure 3.4: Imaginary parts of the measured (blue, solid lines) and modelled (red, dashed lines) y-parameters in Fig. 3.3. (a) Since fewer traps can respond at higher frequencies, the dispersion due to traps is not visible on the linear scale. (b) Top: $\mathrm{Im}(y_{11})$ and $\mathrm{Im}(y_{12})$ with a logarithmic frequency axis to distinguish the dispersion due to traps. Bottom: Schematic of the distributed RC model for traps, which can explain all dispersions in the y-parameters.

effects had been investigated previously [118, 121, 122], but the dispersions in the capacitances, the forward current gain, and the maximum stable gain had not been resolved. The linear and logarithmic frequency dependences in (3.5), (3.6), and (3.7), which describe the different dispersions, can be obtained from a numerical solution of the surface potential in a distributed RC network, which is also commonly used to model oxide traps in CV measurements [123]. In [123], it was demonstrated how such a distributed RC network can model traps with varying distances from the oxide/semiconductor interface and how the total response of this network can be represented by equivalent lumped elements with the aforementioned frequency dependences. Fig. 3.4(b) illustrates such a distributed RC model alongside experimental evidence for the logarithmic frequency dispersion in the capacitances from Paper VI.

We found that on the whole, the effect of traps on the RF performance is measurable, but comparatively small. This is an important conclusion, which helps to set the right priorities for the optimisation of the RF performance of MOSFETs. Further details can be found in Paper VI, a detailed derivation of the g_m-f dispersion in [118], and a holistic, detailed treatise on the evolution of the small-signal model in Appendix B. Furthermore, for consistency between measurement techniques, the measurement and modelling of the RF

dispersion lead to the previously used value of ∼1 ps for the pre-factor τ_0 in the expression for the capture/emission time constant $\tau_{c/e}$ (2.1).

3.3 IMPROVED RF PERFORMANCE

Two of the most important high frequency metrics of a transistor are the cutoff frequency f_T and the maximum oscillation frequency f_{max}, which were compared for different technologies in Section 1.5.2. f_T and f_{max} denote the frequencies, at which the forward current gain h_{21} and the unilateral power gain U, respectively, become equal to one so that at and above f_T and f_{max}, the transistor cannot act as an amplifier anymore. Besides these two frequencies, the maximum stable gain MSG, the maximum available gain MAG, and the stability factor k, which are defined below, are important metrics for small-signal amplifier design. For practical use, definitions of the different parameters are gathered here and for consistency, all definitions are given in terms of the y-parameters [49]. (Definitions based on other, equivalent, parameters, also exist.) The forward current gain h_{21} is defined straightforwardly as

$$h_{21} = \frac{|y_{21}|}{|y_{11}|}, \tag{3.10}$$

the unilateral power gain U is defined as

$$U = \frac{|y_{21} - y_{12}|^2}{4[\mathrm{Re}(y_{11})\mathrm{Re}(y_{22}) - \mathrm{Re}(y_{12})\mathrm{Re}(y_{21})]}, \tag{3.11}$$

and the stability factor k is defined as

$$k = \frac{2\mathrm{Re}(y_{11})\mathrm{Re}(y_{22}) - \mathrm{Re}(y_{12}\,y_{21})}{|y_{12}\,y_{21}|}. \tag{3.12}$$

For $k < 1$, a transistor is potentially unstable, i.e. self-oscillations can occur. The maximum stable gain in this case, i.e. if the transistor is stabilised, is

$$\mathrm{MSG} = \frac{|y_{21}|}{|y_{12}|}. \tag{3.13}$$

For $k > 1$, a transistor is unconditionally stable and can provide the gain

$$\mathrm{MAG} = \frac{|y_{21}|}{|y_{12}|}\left(k - \sqrt{k^2 - 1}\right). \tag{3.14}$$

To obtain these metrics for a measured device, the access resistances R_g, R_s, and R_d have to be added again to the intrinsic y-parameters from (3.4).

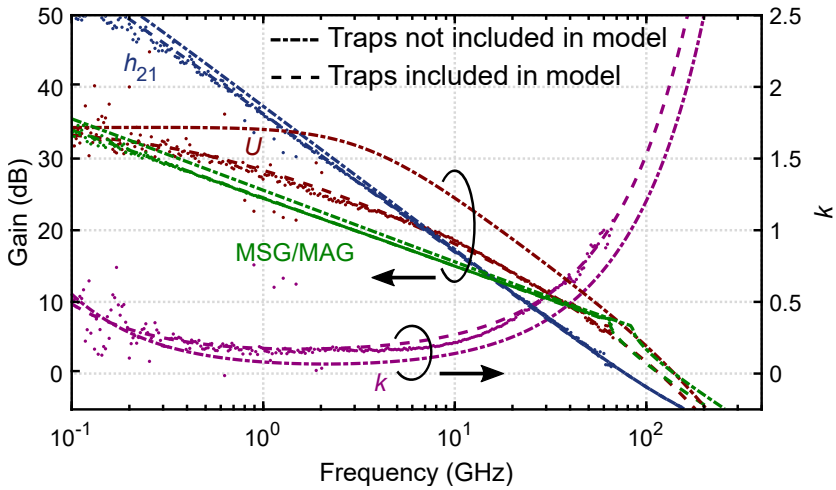


Figure 3.5: The different gains h_{21}, U, MSG/MAG, and the stability factor k for a vertical nanowire MOSFET. Dots are measured values and the dashed lines are calculated according to the model in Fig. 3.2 with (3.4)–(3.14), which fits the measurements well. The dash-dotted lines do not take into account the effects of oxide traps and the deviation of the model from the measured values is obvious. Extrapolation of the model including traps results in $f_\mathrm{T} \approx 150$ GHz and $f_\mathrm{max} \approx 160$ GHz.

Approximative (disregarding second- and higher-order terms) analytic expressions for f_T and f_max can be obtained by setting (3.10) and (3.11) equal to one and solving for the frequency. If the access resistances R_s and R_d are not extremely large, and when disregarding the effect of traps at high frequencies, this leads to [117, pp. 437 ff.]

$$f_\mathrm{T} \approx \frac{1}{2\pi}\left[\frac{C_\mathrm{gg,t}}{g_\mathrm{m}} + (R_\mathrm{s}+R_\mathrm{d})\left(C_\mathrm{gd,t} + \frac{g_\mathrm{ds}}{g_\mathrm{m}}C_\mathrm{gg,t}\right)\right]^{-1} \tag{3.15}$$

and

$$f_\mathrm{max} \approx \sqrt{\frac{f_\mathrm{T}}{8\pi R_\mathrm{g}\left(C_\mathrm{gd,t} + 2\pi f_\mathrm{T}\Psi\right)}}, \tag{3.16}$$

where [124]

$$\Psi = \frac{g_\mathrm{ds}}{g_\mathrm{m}^2}\left[C_\mathrm{gg,t}^2 + \frac{R_\mathrm{s}}{R_\mathrm{g}}(C_\mathrm{gs,i}+C_\mathrm{gs,p})^2 + \frac{R_\mathrm{i}}{R_\mathrm{g}}C_\mathrm{gs,i}^2 + \frac{R_\mathrm{d}}{R_\mathrm{g}}C_\mathrm{gd,t}^2\right]. \tag{3.17}$$

Here, as a simplification, the common observation of $C_\mathrm{gd,p} \gg C_\mathrm{gd,i}$ was exploited, i.e. that the parasitic drain capacitance is typically much larger

than the intrinsic one. With this, $C_{gd,t}$ in (3.15)–(3.17) is the effective measured combination of $C_{gd,p}$ and $C_{gd,i}$ and $C_{gg,t} = C_{gs,p} + C_{gs,i} + C_{gd,t}$. In (3.15) it can be noted that the intrinsic R_i drops out of the expression, while in (3.16) it adds a contribution to the main limitation of f_{max} by the extrinsic R_g. If R_s and R_d can be made very small ($\lesssim 1\ \Omega$), (3.15)–(3.17) can be simplified further by disregarding the terms with R_s and R_d.

These expressions for f_T and f_{max} finally demonstrate, why the previously mentioned low gate capacitances due to low DOS in III-V MOSFETs can be an advantage. Furthermore, (3.15)–(3.17) make it possible to draw some straightforward and general conclusions: A high f_T requires a high transconductance, a low g_{ds}, low R_s and R_d, and low capacitances; and a high f_{max} requires a high f_T, low R_s and R_d, especially a low R_g, and among the capacitances, especially a low $C_{gd,t}$. While these conclusions are obvious from the analytic expressions, the experimental difficulty in verifying this for vertical III-V nanowire MOSFETs consisted in the examination whether all of the assumptions during the derivations and approximations are justified and no other shortcomings – e.g. very high contact resistances, the gate oxide, leakage – supersede the critical parameters in (3.15)–(3.17).

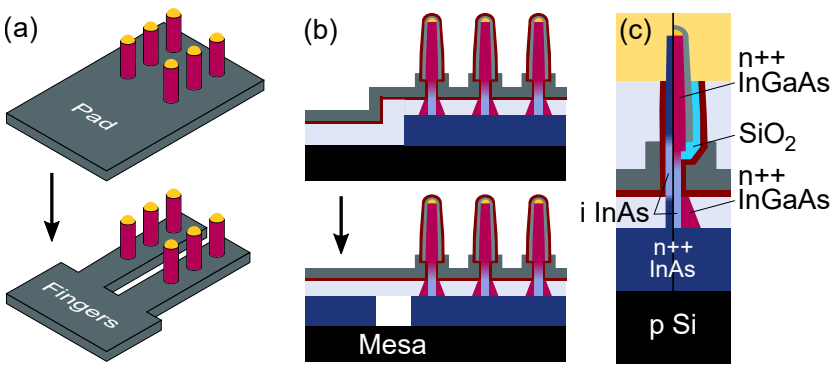


Figure 3.6: Schematic illustrations of the most important changes to the structural layout of vertical nanowire RF MOSFETs. (a) The gate metal pads were changed to finger contacts around the nanowires to reduce the parasitic $C_{gs,p}$. (b) For a further reduction of $C_{gs,p}$, the pad for the gate via was raised from the substrate. Isolation between the raised pad and the bottom contact of the transistor was realised by air bridges. (c) Left: Previous layout. Right: A highly doped foot was introduced to be able to increase the distance between gate metal and bottom contact without increasing the access resistance. Furthermore, a SiO_2 side wall spacer was introduced at the top of the transistor to reduce the parasitic $C_{gd,p}$ due to an overlap between gate metal and top metal.

In Fig. 3.5, the dispersions in the gains are clearly discernible when comparing the measured data with values as calculated by a small-signal model which does not take into account the effect of traps. With (3.4)–(3.14), these dispersions can be readily understood: The frequency dependences in the transconductance g_{m} (3.5) and in the admittances y_{gx} (3.6) and $y_{\mathrm{gx,p}}$ (3.7), which affect the y-parameters (3.4), also change the frequency dependences of h_{21}, U, MSG, and k, all of which are calculated from the y-parameters. A more detailed discussion can be found in Paper VI and in Appendix B.

As part of the work for this thesis, it was investigated, which parameters are crucial for scaling high frequency vertical nanowire MOSFETs. We found that the major challenge from a processing point of view, besides the obvious improvement from increasing the transconductance, is the reduction of the parasitic capacitances. A finger gate layout to reduce the parasitic capacitances on the bottom side of the MOSFETs had been developed previously for a gate-first process [62], i.e. where during processing, the gate stack is applied as one of the first steps. This gate-first scheme can entail large access resistances due to difficulties in aligning the gate, source, and drain contacts [125]. With the help of the modelling carried out during work for this thesis, the finger gate process, schematically depicted in Fig. 3.6(a) was successfully transferred to a gate-last process, as demonstrated by the SEM image in Fig. 3.7(a). The gate-last process enables a much better alignment of the gate and drain contacts [33]. Besides the necessity for gate fingers, we found that it is crucial to ensure a proper mesa separation between the gate and the bottom contact, as depicted in the schematic in Fig. 3.6(b)

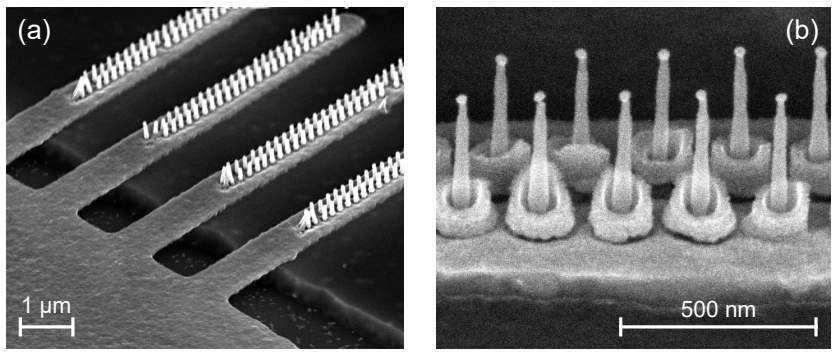


Figure 3.7: SEM images (courtesy O.-P. Kilpi) taken during the processing of vertical nanowire RF MOSFETs to demonstrate the improvements illustrated in Fig. 3.6. (a) Gate fingers suspended over a trench, which separates the gate pad for the via contact from the source mesa. (b) Openings between the gate metal and the nanowires for the SiO_2 side wall spacers, cf. Fig. 3.6(c).

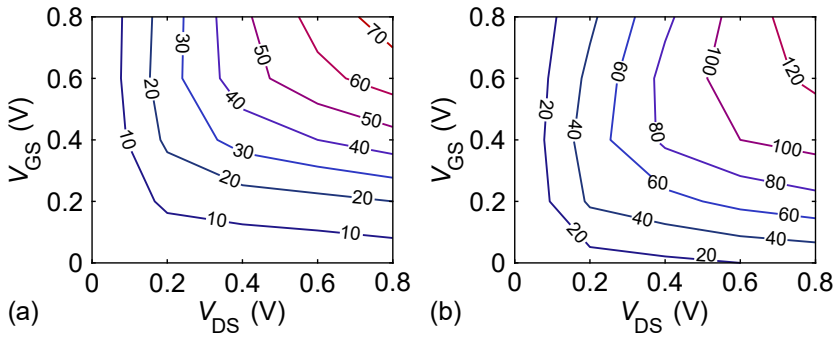


Figure 3.8: (a) f_T and (b) f_{max} as a function of V_{GS} and V_{DS} at 85 °C. Although f_T and f_{max} continuously increase with increasing V_{GS} and V_{DS} here, they typically decrease again at even higher voltages, when g_m decreases.

and discernible in the SEM image in Fig. 3.7(a). Even with gate fingers, there is some geometric parasitic capacitance left between the fingers and the MOSFET bottom contact. To be able to reduce this remaining parasitic by increasing the distance of the gate from the bottom contact without increasing the bottom access resistance, the growth of a highly doped foot around the bottom of the nanowires was introduced, see Fig. 3.6(c). In order to reduce the parasitic capacitance at the top of the gate, which originates from an overlap between the self-aligned gate and the top metal contact, a SiO_2 side wall spacer was developed, as illustrated in Fig. 3.6(c) and demonstrated in Fig. 3.7(b). This side wall spacer can at the same time be used to realise a field plate on the top side of the gate so that if the top contact is connected as the drain, the drain breakdown voltage is increased [33]. Fig. 3.9(c) in Section 3.4 demonstrates the measured capacitances of a MOSFET in an intermediate state of the described structural development. The devices on this sample featured finger gates, but no side wall spacers, and the distance between the gate pad and the bottom contact was still too small, which is evident from the large values for C_{gs}. Still, these devices reached values of $f_T \gtrsim 120$ GHz and $f_{max} \gtrsim 130$ GHz, as demonstrated in Paper VII and in Fig. 3.5.

While the characterisation of individual devices can provide many insights into their transport physics, ultimately, transistors are intended to work in concert in some form of circuit. As a step in this direction, a III-V-compatible low temperature back end of line (BEOL) process with four metal layers was developed in the *Nanoelectronics* group by S. Andrić [126]. We measured the RF performance of lateral nanowire MOSFETs embedded in this BEOL as a function of V_{GS} and V_{GD} and, to investigate the stability of the process at temperatures specified for industrial applications (up to 85 °C), as a function of temperature. Results for f_T and f_{max} at 85 °C are presented in Fig. 3.8. The comparison of both lateral and vertical nanowire MOSFETs before and

after the application of the BEOL structure in [127] reveals that the process does not significantly deteriorate the DC metrics of the transistors. However, the BEOL structure contained additional capacitances of a few femtofarads, which reduced f_{T} and f_{max} by about 50 GHz when compared with the free-standing transistors. Based on the subsequent analysis of the BEOL structure, it should be possible to reduce these additional capacitances so that the expected reduction in RF performance should become smaller.

3.4 TFET CAPACITANCES

For TFETs, small-signal measurements and modelling were applied to analyse the coupling of the channel charge to the source and the drain terminal and again to analyse the effect of traps. The latter are similar to the effects in MOSFETs and the findings were presented at the 76th Device Research Conference (Related Work). The investigation of the charge coupling resulted in Paper VIII and is summarised in the following.

As in conventional MOSFETs, the coupling of the charge in the channel to the source and drain terminals can be described by capacitances. For the drain side, the intrinsic gate-to-drain capacitance $C_{\mathrm{gd,i}} = \text{-}\partial Q_{\mathrm{D}}/\partial v_{\mathrm{d}}$ can be calculated via the charge Q_{D} injected into the channel from the drain, and the small-signal variation in the drain voltage ∂v_{d};

$$Q_{\mathrm{D}} = q\, L_{\mathrm{G}} \sum_n \int_{E_0}^{\infty} \frac{f_D(E)}{1 + T_{BTB}(E)} D_{1D}(E - E_n)\, dE, \tag{3.18}$$

where E_0 is the energy at the top of the barrier in the channel, f_{D} is the Fermi-Dirac distribution in the drain contact, $D_{1\mathrm{D}}$ is the 1D density of states in the channel, E_n is the bottom of the nth sub band, and the tunnelling probability T_{BTB} takes into account the small amount of charge that can tunnel from the channel into the source. For (3.18) it was assumed that the nanowires in the devices confine charge transport to one dimension, which brings about the use of the 1D density of states.

While the coupling between the channel and the drain in a TFET is basically the same as in a conventional MOSFET, due to the similar band structures on the drain side, it is very different on the source side due to the small tunnelling probability T_{BTB}. Our measurements of the on-current and its calculation in Paper VIII confirmed that T_{BTB} is in the order of only a few per cent so that the charge in the channel is effectively decoupled from small-signal variations with respect to the source. This results in very small intrinsic gate-to-source capacitances $C_{\mathrm{gs,i}}$ at all (reasonable) bias conditions, as can be seen in Fig. 3.9(a) and (b). Furthermore, it effectively turns the TFET into a two-terminal device in terms of small-signal voltages so that $C_{\mathrm{gd,i}}$

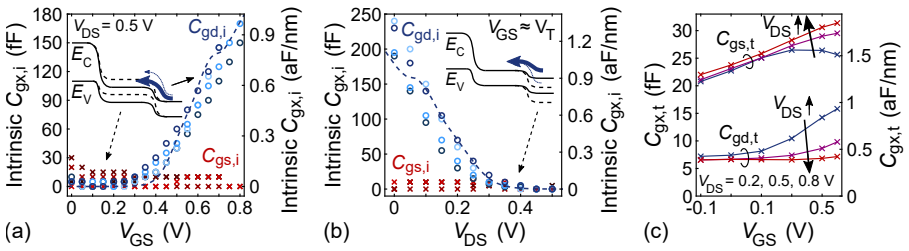


Figure 3.9: Intrinsic capacitances measured in TFETs as a function of V_{GS} (a) and of V_{DS} (b). The insets depict schematics of the band diagrams at the different bias conditions to illustrate the variation of the capacitances. (c) Total capacitances measured in a MOSFET as comparison, reproduced from [63]. Note that the TFETs consisted of arrays of ∼2000 nanowires to provide sufficient gain for the RF measurements despite the small T_{BTB}, while the MOSFETs consisted of only 180 nanowires. The absolute capacitances are thus very different, but when normalised to the total gate width, the C_{gd} are very similar in MOSFETs and TFETs, as expected.

becomes symmetric with respect to AC variations in the gate and the drain voltage. This is confirmed by the good agreement of the measured $C_{gd,i}$ and its calculation via (3.18) as a function of the gate and the drain voltages V_{GS} and V_{DS} in 3.9(a) and (b), respectively.

Explanations for the observed variations of $C_{gd,i}$ are illustrated by the schematic band diagrams in the insets of Fig. 3.9. For a variation of either V_{GS} or V_{DS}, the gate-to-drain electric field changes, which changes the amount of electrons injected into the channel from the drain. At the same time, the tunnelling probability between source and channel remains small so that no change can be observed in $C_{gs,i}$. This is in contrast to the capacitances in conventional MOSFETs, presented in Fig. 3.9(c), where both $C_{gd,t}$ and $C_{gs,t}$ vary with varying V_{GS} and V_{DS}. Note that in Fig. 3.9(a) and (b) the (large) parasitic capacitances are subtracted for clarity, while in Fig. 3.9(c), the total capacitances are plotted, which include parasitics.

In terms of performance, the TFETs achieved f_T and f_{max} of about 3 GHz, which was mostly limited by the low transconductance due to the low tunnelling probability, and by the large parasitic capacitances, which were not optimised for RF operation. With an improved RF layout similar to the improvements investigated for MOSFETs before, it should be possible to considerably increase f_T and f_{max}. At the same time, TFETs are not a contestant for high frequency operation. With their steep switching and the resulting low operating voltages, they are rather suited for use in e.g. low power applications for the Internet of Things, such as sensing applications, where moderate operating frequencies are sufficient.

4

Outlook and Final Words

Over the course of this thesis, many different aspects of III-V nanowire MOSFETs have been investigated and the breadth of these investigations provided a comprehensive understanding of these kinds of devices. However, while many aspects of III-V nanowire MOSFETs have been studied, the intention to write a coherent thesis around one or two central topics sometimes required disregarding possible avenues for further detailed investigation. Even with only two main topics, there were things which could have been studied in more detail. As a proposal for future research, examples of some of these areas for further investigation are given below.

It would be interesting to study a possible dependence of the low-frequency noise (LFN) on the channel material. Different band alignments between the channel conduction band and the gate oxide as well as different densities of states (DOS) in the channel could place different parts of the trap distribution in the electrically accessible range. As discussed in Section 2.3, the band alignment and the DOS in the channel are different in Si material systems and as such, are two of the likely reasons for typically smaller LFN in Si devices. Furthermore, it would be interesting to study in more detail the apparent difference in LFN in long channel and short channel devices that was observed in Paper II. This could lead to an expression for LFN which takes into account the higher ballisticity in short channel III-V MOSFETs. A more detailed measurement of the LFN temperature dependence would also be of interest so as to resolve the partitioning between elastic and inelastic contributions. In the IV curves of TFETs, we clearly and repeatedly observed multiple peaks both in the 'normal' on-state and in the negative differential resistance regions. It would be very interesting to resolve the origins of these peaks.

Many questions remain unresolved after five years of intensive research and in fact, many new questions have been unearthed. But this is and will always be the way of science; a balance has to be struck between the amount and the depth of investigation. The choice to focus this thesis on oxide defects and high frequency performance was made halfway into the five years of this PhD project. This provided some guidelines as to which observations to leave aside for the time being, with the hope, perhaps, to pick them up again at a later time. An experiment planned, in fact, for the months after writing this thesis is the measurement of the stability of small-signal RF parameters. This combines RF measurements and reliability measurements and would be very relevant in the design of RF circuits.

So, to summarise this thesis; III-V nanowire MOSFETs have been subjected to a wide range of measurements over a wide range of frequencies. Oxide traps are a major concern for the stability and the reliability of these devices, and somewhat informally, "anything can happen" to electrical devices in the presence of traps. In fact, this variability goes so far that it is now being explored to make use of in resistive memory applications [128]. Despite the difficulties due to traps, excellent performance has been demonstrated in III-V nanowire MOSFETs and, often, the limiting factors of specific samples, and especially of the yield, can be traced back to irregularities in the processing tools of the university cleanroom. Such irregularities occur due to drift in the tool parameters or the very different use of these tools by many different users, with different projects, using different materials, and different processing conditions. With this in mind I believe that the question whether or not III-V MOSFETs should or will be adopted in industry on a large scale is not so much one of performance, but purely one of processing yield. If industry thus decided to integrate III-V MOSFETs into their designs, I am fairly certain that ways would be found to make it happen. The International Roadmap for Devices and Systems [129] predicts that vertical nanowire MOSFETs will play a central role in monolithic 3D integration by the year 2030, so chances are good that in a few years, the results of this academic thesis will contribute to the industrial development of new user applications. Specifically for III-V nanowire MOSFETs, this might be in the form of co-integration with Si logic; for TFETs it might be in the sensing level of monolithically integrated circuits, and especially in IoT applications.

As the final words on a personal note, I aim to continue on an academic Road with postdoctoral research as the next milestone. And whither then? I cannot say ...

Bibliography

[1] S. Salahuddin, K. Ni, and S. Datta, "The era of hyper-scaling in electronics," *Nature Electronics*, vol. 1, no. 8, pp. 442–450, Aug. 2018. doi: 10.1038/s41928-018-0117-x

[2] J. E. Lilienfeld, "Method and apparatus for controlling electric currents," Patent US 1 745 175, Jan., 1930.

[3] J. Bardeen and W. H. Brattain, "The transistor, a semi-conductor triode," *Phys. Rev.*, vol. 74, pp. 230–231, Jul. 1948. doi: 10.1103/PhysRev.74.230

[4] W. Shockley and G. L. Pearson, "Modulation of conductance of thin films of semi-conductors by surface charges," *Phys. Rev.*, vol. 74, pp. 232–233, Jul. 1948. doi: 10.1103/PhysRev.74.232

[5] H. F. Mataré and H. Welker, "Crystal device for controlling electric currents by means of a solid semiconductor," Patent US 2 673 948, Mar., 1954.

[6] W. Shockley, "Circuit element utilizing semiconductive material," Patent US 2 569 347, Sep., 1951.

[7] D. Kahng and M. M. Atalla, "Silicon-silicon dioxide field induced surface devices," in *IRE-AIEE Solid-state Device Res. Conf*, Jun. 1960.

[8] F. M. Wanlass and C. Sah, "Nanowatt logic using field-effect metal-oxide semiconductor triodes," in *International Solid State Circuits Conference*, Feb. 1963. doi: 10.1142/9789814503464_0081

[9] R. H. Dennard, F. H. Gaensslen, H. Yu, V. L. Rideout, E. Bassous, and A. R. LeBlanc, "Design of ion-implanted MOSFET's with very small physical dimensions," *IEEE Journal of Solid-State Circuits*, vol. 9, no. 5, pp. 256–268, Oct. 1974. doi: 10.1109/JSSC.1974.1050511

[10] G. E. Moore, "Progress in digital integrated electronics," in *1975 IEEE International Electron Devices Meeting*, Dec. 1975. doi: 10.1109/N-SSC.2006.4804410 pp. 11–13.

[11] S.-H. Lo, D. A. Buchanan, and Y. Taur, "Modeling and characterization of quantization, polysilicon depletion, and direct tunneling effects in MOSFETs with ultrathin oxides," *IBM Journal of Research and Development*, vol. 43, no. 3, pp. 327–337, May 1999. doi: 10.1147/rd.433.0327

[12] P. W. Peacock and J. Robertson, "Band offsets and Schottky barrier heights of high dielectric constant oxides," *Journal of Applied Physics*, vol. 92, no. 8, pp. 4712–4721, Oct. 2002. doi: 10.1063/1.1506388

[13] K. Mistry, C. Allen, C. Auth, B. Beattie, D. Bergstrom, M. Bost, M. Brazier, M. Buehler, A. Cappellani, R. Chau *et al.*, "A 45nm logic technology with high-k+metal gate transistors, strained silicon, 9 Cu interconnect layers, 193nm dry patterning, and 100% Pb-free packaging," in *2007 IEEE International Electron Devices Meeting*, Dec. 2007. doi: 10.1109/IEDM.2007.4418914 pp. 247–250.

[14] R. Chau, S. Datta, M. Doczy, B. Doyle, J. Kavalieros, and M. Metz, "High-κ/metal-gate stack and its MOSFET characteristics," *IEEE Electron Device Letters*, vol. 25, no. 6, pp. 408–410, Jun. 2004. doi: 10.1109/LED.2004.828570

[15] H. Riel, L.-E. Wernersson, M. Hong, and J. A. del Alamo, "III–V compound semiconductor transistors—from planar to nanowire structures," *MRS Bulletin*, vol. 39, no. 8, pp. 668—-677, Aug. 2014. doi: 10.1557/mrs.2014.137

[16] H. Ibach and H. Lüth, *Solid-State Physics*, 4th ed. Springer. ISBN 978-3-540-93803-3

[17] E. Lind, "High frequency III–V nanowire MOSFETs," *Semiconductor Science and Technology*, vol. 31, no. 9, p. 093005, Aug. 2016. doi: 10.1088/0268-1242/31/9/093005

[18] M. Lundström and J. Guo, *Nanoscale Transistors*. Springer, 2006. ISBN 978-0-387-28002-8

[19] P. Y. Yu and M. Cardona, *Fundamentals of Semiconductors*. Springer. ISBN 978-3-642-00709-5

[20] "Physical properties of semiconductors," http://www.ioffe.ru/SVA/NSM/Semicond/index.html, accessed: 2020-02-19.

[21] K. Zou, X. Hong, and J. Zhu, "Effective mass of electrons and holes in bilayer graphene: Electron-hole asymmetry and electron-electron interaction," *Phys. Rev. B*, vol. 84, p. 085408, Aug. 2011. doi: 10.1103/PhysRevB.84.085408

[22] K. I. Bolotin, K. J. Sikes, Z. Jiang, M. Klima, G. Fudenberg, J. Hone, P. Kim, and H. L. Stormer, "Ultrahigh electron mobility in suspended graphene," *Solid State Communications*, vol. 146, no. 9, pp. 351–355, Feb. 2008. doi: 10.1016/j.ssc.2008.02.024

[23] R. S. Deacon, K.-C. Chuang, R. J. Nicholas, K. S. Novoselov, and A. K. Geim, "Cyclotron resonance study of the electron and hole velocity in graphene monolayers," *Phys. Rev. B*, vol. 76, p. 081406, Aug. 2007. doi: 10.1103/PhysRevB.76.081406

[24] T. P. O'Regan, M. V. Fischetti, B. Sorée, S. Jin, W. Magnus, and M. Meuris, "Calculation of the electron mobility in III-V inversion layers with high-κ dielectrics," *Journal of Applied Physics*, vol. 108, no. 10, p. 103705, Nov. 2010. doi: 10.1063/1.3500553

[25] A. Sonnet, R. Galatage, P. Hurley, E. Pelucchi, K. Thomas, A. Gocalinska, J. Huang, N. Goel, G. Bersuker, W. Kirk *et al.*, "Remote phonon and surface roughness limited universal electron mobility of $In_{0.53}Ga_{0.47}As$ surface channel MOSFETs," *Microelectronic Engineering*, vol. 88, no. 7, pp. 1083–1086, Mar. 2011. doi: 10.1016/j.mee.2011.03.120 Proceedings of the 17th Biennial International Insulating Films on Semiconductor Conference.

[26] L. Lin and J. Robertson, "Defect states at III-V semiconductor oxide interfaces," *Applied Physics Letters*, vol. 98, no. 8, p. 082903, Feb. 2011. doi: 10.1063/1.3556619

[27] C. Hinkle, E. Vogel, P. Ye, and R. Wallace, "Interfacial chemistry of oxides on $In_xGa_{(1-x)}As$ and implications for MOSFET applications," *Current Opinion in Solid State and Materials Science*, vol. 15, no. 5, pp. 188–207, May 2011. doi: 10.1016/j.cossms.2011.04.005

[28] S. Thompson, N. Anand, M. Armstrong, C. Auth, B. Arcot, M. Alavi, P. Bai, J. Bielefeld, R. Bigwood, J. Brandenburg *et al.*, "A 90 nm logic technology featuring 50nm strained silicon channel transistors, 7 layers of Cu interconnects, low k ILD, and 1 um^2 SRAM cell," in *2002 IEEE International Electron Devices Meeting*, Dec. 2002. doi: 10.1109/IEDM.2002.1175779 pp. 61–64.

[29] K. Suzuki, T. Tanaka, Y. Tosaka, H. Horie, and Y. Arimoto, "Scaling theory for double-gate SOI MOSFET's," *IEEE Transactions on Electron Devices*, vol. 40, no. 12, pp. 2326–2329, Dec. 1993. doi: 10.1109/16.249482

[30] C. P. Auth and J. D. Plummer, "Scaling theory for cylindrical, fully-depleted, surrounding-gate MOSFET's," *IEEE Electron Device Letters*, vol. 18, no. 2, pp. 74–76, Feb. 1997. doi: 10.1109/55.553049

[31] L.-E. Wernersson, C. Thelander, E. Lind, and L. Samuelson, "III-V nanowires—extending a narrowing road," *Proceedings of the IEEE*, vol. 98, no. 12, pp. 2047–2060, Dec. 2010. doi: 10.1109/JPROC.2010.2065211

[32] M. M. Shulaker, G. Hills, R. S. Park, R. T. Howe, K. Saraswat, H.-S. P. Wong, and S. Mitra, "Three-dimensional integration of nanotechnologies for computing and data storage on a single chip," *Nature*, vol. 547, pp. 74–78, Jul. 2017. doi: 10.1038/nature22994

[33] O.-P. Kilpi, "Vertical III-V nanowire MOSFETs," Ph.D. dissertation, Lund University. ISBN 978-91-7895-293-9 Oct. 2019. [Online]. Available: https://lup.lub.lu.se/search/ws/files/69641798/Thesis_online.pdf

[34] E. Memisevic, "Vertical III-V nanowire tunnel field-effect transistor," Ph.D. dissertation, Lund University. ISBN 978-91-7753-255-2 Aug. 2017. [Online]. Available: https://lup.lub.lu.se/search/ws/files/30742687/Thesis_Vertical_III_V_Nanowire_TFET_Memisevic_1.pdf

[35] P. Roblin and H. Rohdin, *High-speed heterostructure devices*. Cambridge University Press, 2002. ISBN 978-0-511-06903-1

[36] O.-P. Kilpi, J. Svensson, E. Lind, and L.-E. Wernersson, "Electrical properties of vertical InAs/InGaAs heterostructure MOSFETs," *IEEE Journal of the Electron Devices Society*, vol. 7, pp. 70–75, Oct. 2018. doi: 10.1109/JEDS.2018.2878659

[37] C. B. Zota, L.-E. Wernersson, and E. Lind, "Single suspended InGaAs nanowire mosfets," in *2015 IEEE International Electron Devices Meeting*, Dec. 2015. doi: 10.1109/IEDM.2015.7409808 pp. 31.4.1–31.4.4.

[38] C. B. Zota, F. Lindelow, L.-E. Wernersson, and E. Lind, "InGaAs tri-gate MOSFETs with record on-current," in *2016 IEEE International Electron Devices Meeting*, Dec. 2016. doi: 10.1109/IEDM.2016.7838336 pp. 3.2.1–3.2.4.

[39] X. Zhao, C. Heidelberger, E. A. Fitzgerald, W. Lu, A. Vardi, and J. A. del Alamo, "Sub-10 nm diameter InGaAs vertical nanowire MOSFETs," in *2017 IEEE International Electron Devices Meeting*, Dec. 2017. doi: 10.1109/IEDM.2017.8268407 pp. 17.2.1–17.2.4.

[40] T.-W. Kim, D.-H. Koh, C.-S. Shin, W.-K. Park, T. Orzali, C. Hobbs, W. P. Maszara, and D.-H. Kim, "L_g = 80-nm trigate quantum-well $In_{0.53}Ga_{0.47}As$ metal–oxide–semiconductor field-effect transistors with Al_2O_3/HfO_2 gate-stack," *IEEE Electron Device Letters*, vol. 36, no. 3, pp. 223–225, Mar. 2015. doi: 10.1109/LED.2015.2393554

[41] S. Lee, C. Huang, D. Cohen-Elias, B. J. Thibeault, W. Mitchell, V. Chobpattana, S. Stemmer, A. C. Gossard, and M. J. W. Rodwell, "Highly scalable raised source/drain InAs quantum well MOSFETs exhibiting $I_{ON} = 482\ \mu A/\mu m$ at $I_{OFF} = 100$ nA/μm and $V_{DD} = 0.5$ V," *IEEE Electron Device Letters*, vol. 35, no. 6, pp. 621–623, Jun. 2014. doi: 10.1109/LED.2014.2317146

[42] S. Chang, X. Li, R. Oxland, S. Wang, C. Wang, R. Contreras-Guerrero, K. Bhuwalka, G. Doornbos, T. Vasen, M. Holland *et al.*, "InAs N-MOSFETs with record performance of I_{on} = 600 µA/µm at I_{off} = 100 nA/µm (V_d = 0.5 V)," in *2013 IEEE International Electron Devices Meeting*, Dec. 2013. doi: 10.1109/IEDM.2013.6724639 pp. 16.1.1–16.1.4.

[43] N. Waldron, S. Sioncke, J. Franco, L. Nyns, A. Vais, X. Zhou, H. C. Lin, G. Boccardi, J. W. Maes, Q. Xie *et al.*, "Gate-all-around InGaAs nanowire FETS with peak transconductance of 2200 µS/µm at 50nm Lg using a replacement fin RMG flow," in *2015 IEEE International Electron Devices Meeting*, Dec. 2015. doi: 10.1109/IEDM.2015.7409805 pp. 31.1.1–31.1.4.

[44] J. Lin, X. Cai, Y. Wu, D. A. Antoniadis, and J. A. del Alamo, "Record maximum transconductance of 3.45 ms/μm for III-V FETs," *IEEE Electron Device Letters*, vol. 37, no. 4, pp. 381–384, Apr. 2016. doi: 10.1109/LED.2016.2529653

[45] S. Narasimha, B. Jagannathan, A. Ogino, D. Jaeger, B. Greene, C. Sheraw, K. Zhao, B. Haran, U. Kwon, A. Mahalingam *et al.*, "A 7nm CMOS technology platform for mobile and high performance compute application," in *2017 IEEE International Electron Devices Meeting*, Dec. 2017. doi: 10.1109/IEDM.2017.8268476 pp. 29.5.1–29.5.4.

[46] C.-H. Jan, F. Al-Amoody, H.-Y. Chang, T. Chang, Y.-W. Chen, N. Dias, W. Hafez, D. Ingerly, M. Jang, E. Karl *et al.*, "A 14 nm SoC platform technology featuring 2nd generation tri-gate transistors, 70 nm gate pitch, 52 nm metal pitch, and 0.0499 um^2 SRAM cells, optimized for low power, high performance and high density SoC products," in *2015 Symposium on VLSI Technology*, Jun. 2015. doi: 10.1109/VLSIT.2015.7223683 pp. T12–T13.

[47] R. Carter, J. Mazurier, L. Pirro, J. Sachse, P. Baars, J. Faul, C. Grass, G. Grasshoff, P. Javorka, T. Kammler *et al.*, "22nm FDSOI technology for emerging mobile, internet-of-things, and RF applications," in *2016 IEEE International Electron Devices Meeting*, Dec. 2016. doi: 10.1109/IEDM.2016.7838029 pp. 2.2.1–2.2.4.

[48] H.-J. Cho, K.-I. Seo, W. C. Jeong, Y.-H. Kim, Y. D. Lim, W. W. Jang, J. G. Hong, S. D. Suk, M. Li, C. Ryou *et al.*, "Bulk planar 20nm high-k/metal gate CMOS technology platform for low power and high performance applications," in *2011 IEEE International Electron Devices Meeting*, Dec. 2011. doi: 10.1109/IEDM.2011.6131556 pp. 15.1.1–15.1.4.

[49] P. Roblin and H. Rohdin, *Small- and large-signal AC models for the long-channel MODFET*. Cambridge University Press, ch. 11. ISBN 978-0-511-06903-1

[50] X. Mei, W. Yoshida, M. Lange, J. Lee, J. Zhou, P. Liu, K. Leong, A. Zamora, J. Padilla, S. Sarkozy *et al.*, "First demonstration of amplification at 1 THz using 25-nm InP high electron mobility transistor process," *IEEE Electron Device Letters*, vol. 36, no. 4, pp. 327–329, Apr. 2015. doi: 10.1109/LED.2015.2407193

[51] D. Kim, J. A. del Alamo, P. Chen, Wonill Ha, M. Urteaga, and B. Brar, "50-nm E-mode $In_{0.7}Ga_{0.3}As$ PHEMTs on 100-mm InP substrate with $f_{max} > 1$ THz," in *2010 IEEE International Electron Devices Meeting*, Dec. 2010. doi: 10.1109/IEDM.2010.5703453 pp. 30.6.1–30.6.4.

[52] D. Kim, B. Brar, and J. A. del Alamo, "$f_T = 688$ GHz and $f_{max} = 800$ GHz in $L_g = 40$ nm $In_{0.7}Ga_{0.3}As$ MHEMTs with $g_{m\_max} > 2.7\,\mu S/\mu m$," in *2011 IEEE International Electron Devices Meeting*, Dec. 2011. doi: 10.1109/IEDM.2011.6131548 pp. 13.6.1–13.6.4.

[53] E.-Y. Chang, C.-I. Kuo, H.-T. Hsu, C.-Y. Chiang, and Y. Miyamoto, "InAs thin-channel high-electron-mobility transistors with very high current-gain cutoff frequency for emerging submillimeter-wave applications," *Applied Physics Express*, vol. 6, no. 3, p. 034001, Mar. 2013. doi: 10.7567/apex.6.034001

[54] Y. Tang, K. Shinohara, D. Regan, A. Corrion, D. Brown, J. Wong, A. Schmitz, H. Fung, S. Kim, and M. Micovic, "Ultrahigh-speed GaN high-electron-mobility transistors with f_T/f_{max} of 454/444 GHz," *IEEE Electron Device Letters*, vol. 36, no. 6, pp. 549–551, Jun. 2015. doi: 10.1109/LED.2015.2421311

[55] A. Tessmann, A. Leuther, F. Heinz, F. Bernhardt, L. John, H. Massler, L. Czornomaz, and T. Merkle, "20-nm $In_{0.8}Ga_{0.2}As$ MOSHEMT MMIC technology on silicon," *IEEE Journal of Solid-State Circuits*, vol. 54, no. 9, pp. 2411–2418, Sep. 2019. doi: 10.1109/JSSC.2019.2915161

[56] B. Heinemann, H. Rücker, R. Barth, F. Bärwolf, J. Drews, G. G. Fischer, A. Fox, O. Fursenko, T. Grabolla, F. Herzel *et al.*, "SiGe HBT with f_T/f_{max} of 505 GHz/720 GHz," in *2016 IEEE International Electron Devices Meeting*, Dec. 2016. doi: 10.1109/IEDM.2016.7838335 pp. 3.1.1–3.1.4.

[57] S. N. Ong, S. Lehmann, W. H. Chow, C. Zhang, C. Schippel, L. H. K. Chan, Y. Andee, M. Hauschildt, K. K. S. Tan, J. Watts *et al.*, "A 22nm FDSOI technology optimized for RF/mmWave applications," in *2018 IEEE Radio Frequency Integrated Circuits Symposium (RFIC)*, Jun. 2018. doi: 10.1109/RFIC.2018.8429035 pp. 72–75.

[58] J. Singh, J. Ciavatti, K. Sundaram, J. S. Wong, A. Bandyopadhyay, X. Zhang, S. Li, A. Bellaouar, J. Watts, J. G. Lee *et al.*, "14-nm FinFET technology for analog and RF applications," *IEEE Transactions on Electron Devices*, vol. 65, no. 1, pp. 31–37, Jan. 2018. doi: 10.1109/TED.2017.2776838

[59] H.-J. Lee, S. Rami, S. Ravikumar, V. Neeli, K. Phoa, B. Sell, and Y. Zhang, "Intel 22nm FinFET (22FFL) process technology for RF and mmWave applications and circuit design optimization for FinFET technology," in *2018 IEEE International Electron Devices Meeting*, Dec. 2018. doi: 10.1109/IEDM.2018.8614490 pp. 14.1.1–14.1.4.

[60] J. Wu, Y. Fang, B. Markman, H. Y. Tseng, and M. J. W. Rodwell, "L_g = 30 nm InAs channel MOSFETs exhibiting f_{max} = 410 GHz and f_t = 357 GHz," *IEEE Electron Device Letters*, vol. 39, no. 4, pp. 472–475, Apr. 2018. doi: 10.1109/LED.2018.2803786

[61] C. B. Zota, F. Lindelöw, L.-E. Wernersson, and E. Lind, "High-frequency InGaAs tri-gate MOSFETs with f_{max} of 400 GHz," *Electronics Letters*, vol. 52, no. 22, pp. 1869–1871, Oct. 2016. doi: 10.1049/el.2016.3108

[62] S. Johansson, E. Memisevic, L.-E. Wernersson, and E. Lind, "High-frequency gate-all-around vertical InAs nanowire MOSFETs on Si substrates," *IEEE Electron Device Letters*, vol. 35, no. 5, pp. 518–520, May 2014. doi: 10.1109/LED.2014.2310119

[63] O.-P. Kilpi, M. Hellenbrand, J. Svensson, E. Lind, and L.-E. Wernersson, "Vertical nanowire III–V MOSFETs with improved high-frequency gain," *Electronics Letters*, vol. –, no. –, pp. –, Apr. 2020. doi: 10.1049/el.2020.0266 Article in Press.

[64] Y. Wu, X. Zou, M. Sun, Z. Cao, X. Wang, S. Huo, J. Zhou, Y. Yang, X. Yu, Y. Kong *et al.*, "200 GHz maximum oscillation frequency in CVD graphene radio frequency transistors," *ACS Applied Materials & Interfaces*, vol. 8, no. 39, pp. 25 645–25 649, Sep. 2016. doi: 10.1021/acsami.6b05791

[65] Y. Cao, G. J. Brady, H. Gui, C. Rutherglen, M. S. Arnold, and C. Zhou, "Radio frequency transistors using aligned semiconducting carbon nanotubes with current-gain cutoff frequency and maximum oscillation frequency simultaneously greater than 70 GHz," *ACS Nano*, vol. 10, no. 7, pp. 6782–6790, Jun. 2016. doi: 10.1021/acsnano.6b02395

[66] S. Lee, B. Jagannathan, S. Narasimha, A. Chou, N. Zamdmer, J. Johnson, R. Williams, L. Wagner, J. Kim, J. Plouchart *et al.*, "Record RF performance of 45-nm SOI CMOS technology," in *2007 IEEE International Electron Devices Meeting*, Dec. 2007. doi: 10.1109/IEDM.2007.4418916 pp. 255–258.

[67] R. Cheng, J. Bai, L. Liao, H. Zhou, Y. Chen, L. Liu, Y.-C. Lin, S. Jiang, Y. Huang, and X. Duan, "High-frequency self-aligned graphene transistors with transferred gate stacks," *Proceedings of the National Academy of Sciences*, vol. 109, no. 29, pp. 11 588–11 592, Jul. 2012. doi: 10.1073/pnas.1205696109

[68] S. Cristoloveanu, J. Wan, and A. Zaslavsky, "A review of sharp-switching devices for ultra-low power applications," *IEEE Journal of the Electron Devices Society*, vol. 4, no. 5, pp. 215–226, May 2016. doi: 10.1109/JEDS.2016.2545978

[69] A. C. Seabaugh and Q. Zhang, "Low-voltage tunnel transistors for beyond CMOS logic," *Proceedings of the IEEE*, vol. 98, no. 12, pp. 2095–2110, Dec. 2010. doi: 10.1109/JPROC.2010.2070470

[70] M. Hellenbrand, E. Memišević, M. Berg, O.-P. Kilpi, J. Svensson, and L.-E. Wernersson, "Low-frequency noise in III-V nanowire TFETs and MOSFETs," *IEEE Electron Device Letters*, vol. 38, no. 11, pp. 1520–1523, Nov. 2017. doi: 10.1109/LED.2017.2757538

[71] S. Wu, C. Y. Lin, M. C. Chiang, J. J. Liaw, J. Y. Cheng, S. H. Yang, M. Liang, T. Miyashita, C. H. Tsai, B. C. Hsu *et al.*, "A 16nm FinFET CMOS technology for mobile SoC and computing applications," in *2013 IEEE International Electron Devices Meeting*, Dec. 2013. doi: 10.1109/IEDM.2013.6724591 pp. 9.1.1–9.1.4.

[72] D. Sarkar, X. Xie, W. Liu, W. Cao, J. Kang, Y. Gong, S. Kraemer, P. M. Ajayan, and K. Banerjee, "A subthermionic tunnel field-effect transistor with an atomically thin channel," *Nature*, vol. 526, no. 7571, pp. 91–95, Sep. 2015. doi: 10.1038/nature15387

[73] R. Gandhi, Z. Chen, N. Singh, K. Banerjee, and S. Lee, "CMOS-compatible vertical-silicon-nanowire gate-all-around p-type tunneling FETs with $\leq$ 50-mV/decade subthreshold swing," *IEEE Electron Device Letters*, vol. 32, no. 11, pp. 1504–1506, Nov. 2011. doi: 10.1109/LED.2011.2165331

[74] A. Alian, Y. Mols, C. C. M. Bordallo, D. Verreck, A. Verhulst, A. Vandooren, R. Rooyackers, P. G. D. Agopian, J. A. Martino, A. Thean *et al.*, "InGaAs tunnel FET with sub-nanometer EOT and sub-60 mV/dec subthreshold swing at room temperature," *Applied Physics Letters*, vol. 109, no. 24, pp. 243 502–1–243 502–4, Dec. 2016. doi: 10.1063/1.4971830

[75] E. Memisevic, J. Svensson, M. Hellenbrand, E. Lind, and L.-E. Wernersson, "Vertical InAs/GaAsSb/GaSb tunneling field-effect transistor on Si with S = 48 mV/decade and I_{on}= 10 µA/µm for I_{off}= 1 nA/µm at V_{DS}= 0.3 V," in *2016 IEEE International Electron Devices Meeting*, Dec. 2016. doi: 10.1109/IEDM.2016.7838450 pp. 19.1.1–19.1.4.

[76] S. Mookerjea, D. Mohata, T. Mayer, V. Narayanan, and S. Datta, "Temperature-dependent I-V characteristics of a vertical $In_{0.53}Ga_{0.47}As$ tunnel FET," *IEEE Electron Device Letters*, vol. 31, no. 6, pp. 564–566, Jun. 2010. doi: 10.1109/LED.2010.2045631

[77] S. Sant and A. Schenk, "The effect of density-of-state tails on band-to-band tunneling: Theory and application to tunnel field effect transistors," *Journal of Applied Physics*, vol. 122, no. 13, p. 135702, Oct. 2017. doi: 10.1063/1.4994112

[78] E. Memisevic, E. Lind, M. Hellenbrand, J. Svensson, and L. E. Wernersson, "Impact of band-tails on the subthreshold swing of III-V tunnel field-effect transistor," *IEEE Electron Device Letters*, vol. 38, no. 12, pp. 1661–1664, Dec. 2017. doi: 10.1109/LED.2017.2764873

[79] A. M. Ionescu and H. Riel, "Tunnel field-effect transistors as energy-efficient electronic switches," *Nature*, vol. 479, no. 7373, pp. 329–337, Nov. 2011. doi: 10.1038/nature10679

[80] R. W. Johnson, A. Hultqvist, and S. F. Bent, "A brief review of atomic layer deposition: from fundamentals to applications," *Materials Today*, vol. 17, no. 5, pp. 236–246, Jun. 2014. doi: 10.1016/j.mattod.2014.04.026

[81] J. Franco, V. Putcha, A. Vais, S. Sioncke, N. Waldron, D. Zhou, G. Rzepa, P. J. Roussel, G. Groeseneken, M. Heyns *et al.*, "Characterization of oxide defects in InGaAs MOS gate stacks for high-mobility n-channel MOSFETs (invited)," in *2017 IEEE International Electron Devices Meeting*, Dec. 2017. doi: 10.1109/IEDM.2017.8268347 pp. 751–754.

[82] K. Xiong, J. Robertson, M. C. Gibson, and S. J. Clark, "Defect energy levels in HfO_2 high-dielectric-constant gate oxide," *Applied Physics Letters*, vol. 87, no. 18, p. 183505, Oct. 2005. doi: 10.1063/1.2119425

[83] D. Liu and J. Robertson, "Oxygen vacancy levels and interfaces of Al_2O_3," *Microelectronic Engineering*, vol. 86, no. 7, pp. 1668–1671, Mar. 2009. doi: 10.1016/j.mee.2009.03.011

[84] D. M. Fleetwood, P. S. Winokur, R. A. Reber, T. L. Meisenheimer, J. R. Schwank, M. R. Shaneyfelt, and L. C. Riewe, "Effects of oxide traps, interface traps, and "border traps" on metal-oxide-semiconductor devices," *Journal of Applied Physics*, vol. 73, no. 10, pp. 5058–5074, May 1993. doi: 10.1063/1.353777

[85] F. Schanovsky, W. Gös, and T. Grasser, "Multiphonon hole trapping from first principles," *Journal of Vacuum Science & Technology B*, vol. 29, no. 1, p. 01A201, Jan. 2011. doi: 10.1116/1.3533269

[86] F. P. Heiman and G. Warfield, "The effects of oxide traps on the MOS capacitance," *IEEE Transactions on Electron Devices*, vol. 12, no. 4, pp. 167–178, Apr. 1965. doi: 10.1109/T-ED.1965.15475

[87] M. J. Kirton and M. J. Uren, "Capture and emission kinetics of individual Si:SiO_2 interface states," *Applied Physics Letters*, vol. 48, no. 19, pp. 1270–1272, May 1986. doi: 10.1063/1.97000

[88] T. Grasser, "Stochastic charge trapping in oxides: From random telegraph noise to bias temperature instabilities," *Microelectronics Reliability*, vol. 52, no. 1, pp. 39–70, Oct. 2011. doi: 10.1016/j.microrel.2011.09.002 2011 Reliability of Compound Semiconductors (ROCS) Workshop.

[89] H. Chen, Y. Yuan, B. Yu, J. Ahn, P. C. McIntyre, P. M. Asbeck, M. J. W. Rodwell, and Y. Taur, "Interface-state modeling of Al_2O_3–InGaAs MOS from depletion to inversion," *IEEE Transactions on Electron Devices*, vol. 59, no. 9, pp. 2383–2389, Sep. 2012. doi: 10.1109/TED.2012.2205255

[90] V. Putcha, J. Franco, A. Vais, S. Sioncke, B. Kaczer, D. Linten, and G. Groeseneken, "On the apparent non-Arrhenius temperature dependence of charge trapping in IIIV/high-*k* MOS stack," *IEEE Transactions on Electron Devices*, vol. 65, no. 9, pp. 3689–3696, Aug. 2018. doi: 10.1109/TED.2018.2851189

[91] I. Lundström and C. Svensson, "Tunneling to traps in insulators," *Journal of Applied Physics*, vol. 43, no. 12, pp. 5045–5047, Dec. 1972. doi: 10.1063/1.1661067

[92] J. Robertson, Y. Guo, and L. Lin, "Defect state passivation at III-V oxide interfaces for complementary metal–oxide–semiconductor devices," *Journal of Applied Physics*, vol. 117, no. 11, p. 112806, Mar. 2015. doi: 10.1063/1.4913832

[93] G. Greene-Diniz, K. J. Kuhn, P. K. Hurley, and J. C. Greer, "First principles modeling of defects in the $Al_2O_3/In_{0.53}Ga_{0.47}As$ system," *Journal of Applied Physics*, vol. 121, no. 7, p. 075703, Feb. 2017. doi: 10.1063/1.4975033

[94] J. Robertson, "Model of interface states at III-V oxide interfaces," *Applied Physics Letters*, vol. 94, no. 15, p. 152104, Apr. 2009. doi: 10.1063/1.3120554

[95] W. Shockley and W. T. Read, "Statistics of the recombinations of holes and electrons," *Phys. Rev.*, vol. 87, pp. 835–842, Sep. 1952. doi: 10.1103/PhysRev.87.835

[96] Y.-C. Fu, U. Peralagu, D. A. J. Millar, J. Lin, I. Povey, X. Li, S. Monaghan, R. Droopad, P. K. Hurley, and I. G. Thayne, "The impact of forming gas annealing on the electrical characteristics of sulfur passivated $Al_2O_3/In_{0.53}Ga_{0.47}As$ (110) metal-oxide-semiconductor capacitors," *Applied Physics Letters*, vol. 110, no. 14, p. 142905, Apr. 2017. doi: 10.1063/1.4980012

[97] M. Milojevic, F. S. Aguirre-Tostado, C. L. Hinkle, H. C. Kim, E. M. Vogel, J. Kim, and R. M. Wallace, "Half-cycle atomic layer deposition reaction studies of Al_2O_3 on $In_{0.2}Ga_{0.8}As$ (100) surfaces," *Applied Physics Letters*, vol. 93, no. 20, p. 202902, Nov. 2008. doi: 10.1063/1.3033404

[98] J. Wu, A. S. Babadi, D. Jacobsson, J. Colvin, S. Yngman, R. Timm, E. Lind, and L.-E. Wernersson, "Low trap density in InAs/high-*k* nanowire gate stacks with optimized growth and doping conditions," vol. 16, no. 4, pp. 2418–2425, Mar. 2016. doi: 10.1021/acs.nanolett.5b05253

[99] R. Timm, A. R. Head, S. Yngman, J. V. Knutsson, M. Hjort, S. R. McKibbin, A. Troian, O. Persson, S. Urpelainen, J. Knudsen *et al.*, "Self-cleaning and surface chemical reactions during hafnium dioxide atomic layer deposition on indium arsenide," *Nature communications*, vol. 9, no. 1, p. 1412, Apr. 2018. doi: 10.1038/s41467-018-03855-z

[100] S. Yoshida, D. H. L. Lin, R. Suzuki, Y. Miyanami, N. Collaert, T. Hosoi, T. Shimura, and H. Watanabe, "Analysis of III–V oxides at high-k/InGaAs interfaces induced by metal electrodes," *Japanese Journal of Applied Physics*, vol. 58, no. 5, p. 051010, Apr. 2019. doi: 10.7567/1347-4065/ab0256

[101] J. Franco, A. Vais, S. Sioncke, V. Putcha, B. Kaczer, B. . Shie, X. Shi, R. Mahlouji, L. Nyns, D. Zhou *et al.*, "Demonstration of an InGaAs gate stack with sufficient PBTI reliability by thermal budget optimization, nitridation, high-k material choice, and interface dipole," in *2016 IEEE Symposium on VLSI Technology*, Jun. 2016. doi: 10.1109/VLSIT.2016.7573371 pp. 1–2.

[102] E. Burstein, R. H. Kingston, and A. L. McWhorter, *Semiconductor surface physics*. University of Pennsylvania Press, 1957.

[103] M. von Haartman and M. Östling, *Low-Frequency Noise in Advanced MOS Devices*. Springer, 2007. ISBN 978-1-4020-5909-4

[104] G. Ghibaudo, O. Roux, C. Nguyen-Duc, F. Balestra, and J. Brini, "Improved analysis of low frequency noise in field-effect MOS transistors," *physica status solidi (a)*, vol. 124, no. 2, pp. 571–581, Feb. 1991. doi: 10.1002/pssa.2211240225

[105] A. Pacelli, S. Villa, A. L. Lacaita, and L. M. Perron, "Quantum effects on the extraction of MOS oxide traps by $1/f$ noise measurements," *IEEE Transactions on Electron Devices*, vol. 46, no. 5, pp. 1029–1035, May 1999. doi: 10.1109/16.760413

[106] F. Hooge, "$1/f$ noise is no surface effect," *Physics Letters A*, vol. 29, no. 3, pp. 139–140, Apr. 1969. doi: 10.1016/0375-9601(69)90076-0

[107] K.-M. Persson, B. G. Malm, and L.-E. Wernersson, "Surface and core contribution to $1/f$-noise in InAs nanowire metal-oxide-semiconductor field-effect transistors," *Applied Physics Letters*, vol. 103, no. 3, pp. 033 508–1–033 508–4, Jul. 2013. doi: 10.1063/1.4813850

[108] D. Lopez, S. Haendler, C. Leyris, G. Bidal, and G. Ghibaudo, "Low-frequency noise investigation and noise variability analysis in high-k/metal gate 32-nm CMOS transistors," *IEEE Transactions on Electron Devices*, vol. 58, no. 8, pp. 2310–2316, Aug. 2011. doi: 10.1109/TED.2011.2141139

[109] E. Simoen, A. Veloso, Y. Higuchi, N. Horiguchi, and C. Claeys, "On the oxide trap density and profiles of 1-nm EOT metal-gate last CMOS transistors assessed by low-frequency noise," *IEEE Transactions on Electron Devices*, vol. 60, no. 11, pp. 3849–3855, Nov. 2013. doi: 10.1109/TED.2013.2279892

[110] E. G. Ioannidis, S. Haendler, A. Bajolet, T. Pahron, N. Planes, F. Arnaud, R. A. Bianchi, M. Haond, D. Golanski, J. Rosa *et al.*, "Low frequency noise variability in high-k/metal gate stack 28nm bulk and FD-SOI CMOS transistors," in *2011 IEEE International Electron Devices Meeting*, Dec. 2011. doi: 10.1109/IEDM.2011.6131581 pp. 18.6.1–18.6.4.

[111] B. K. Esfeh, V. Kilchytska, V. Barral, N. Planes, M. Haond, D. Flandre, and J.-P. Raskin, "Assessment of 28nm UTBB FD-SOI technology platform for RF applications: Figures of merit and effect of parasitic elements," *Solid-State Electronics*, vol. 117, pp. 130–137, Dec. 2016. doi: 10.1016/j.sse.2015.11.020

[112] H. Reisinger, T. Grasser, W. Gustin, and C. Schlünder, "The statistical analysis of individual defects constituting nbti and its implications for modeling DC- and AC-stress," in *2010 IEEE International Reliability Physics Symposium*, May 2010. doi: 10.1109/IRPS.2010.5488858 pp. 7–15.

[113] V. Putcha, "Reliability characterization of gate-stacks for III-V channel MOSFETs," Ph.D. dissertation, Feb. 2019. [Online]. Available: https://lirias.kuleuven.be/retrieve/532454

[114] B. Kaczer, T. Grasser, J. Roussel, J. Martin-Martinez, R. O'Connor, B. J. O'Sullivan, and G. Groeseneken, "Ubiquitous relaxation in BTI stressing—new evaluation and insights," in *2008 IEEE International Reliability Physics Symposium*, Apr. 2008. doi: 10.1109/RELPHY.2008.4558858 pp. 20–27.

[115] G. Rzepa, J. Franco, B. O'Sullivan, A. Subirats, M. Simicic, G. Hellings, P. Weckx, M. Jech, T. Knobloch, M. Waltl *et al.*, "Comphy — a compact-physics framework for unified modeling of BTI," *Microelectronics Reliability*, vol. 85, pp. 49–65, Apr. 2018. doi: 10.1016/j.microrel.2018.04.002

[116] D. M. Pozar, *Microwave Network Analysis*, 4th ed. John Wiley & Sons, Inc., ch. 4. ISBN 978-0-470-63155-3

[117] W. Liu, *Fundamentals of III-V Devices*. John Wiley & Sons, Inc., 1999. ISBN 978-0-471-29700-0

[118] S. Johansson, M. Berg, K. M. Persson, and E. Lind, "A high-frequency transconductance method for characterization of high-κ border traps in III-V MOSFETs," vol. 60, no. 2, pp. 776–781, Feb. 2013. doi: 10.1109/TED.2012.2231867

[119] M. Isler and K. Schünemann, "Impact-ionization effects on the high-frequency behavior of HFETs," vol. 52, no. 3, pp. 858–863, Mar. 2004. doi: 10.1109/TMTT.2004.823553

[120] S. Johansson, M. Egard, S. G. Ghalamestani, B. M. Borg, M. Berg, L.-E. Wernersson, and E. Lind, "RF characterization of vertical InAs nanowire wrap-gate transistors integrated on Si substrates," *IEEE Transactions on Microwave Theory and Techniques*, vol. 59, no. 10, pp. 2733–2738, Oct. doi: 10.1109/TMTT.2011.2163076

[121] M. Egard, L. Ohlsson, M. Arlelid, K. Persson, B. M. Borg, F. Lenrick, R. Wallenberg, E. Lind, and L.-E. Wernersson, "High-frequency performance of self-aligned gate-last surface channel $In_{0.53}Ga_{0.47}As$ MOSFET," *IEEE Electron Device Letters*, vol. 33, no. 3, pp. 369–371, Mar. 2012. doi: 10.1109/LED.2011.2181323

[122] C. B. Zota, G. Roll, L.-E. Wernersson, and E. Lind, "Radio-frequency characterization of selectively regrown InGaAs lateral nanowire MOSFETs," *IEEE Transactions on Electron Devices*, vol. 61, no. 12, pp. 4078–4083, Dec. 2014. doi: 10.1109/TED.2014.2363732

[123] Y. Yuan, L. Wang, B. Yu, B. Shin, J. Ahn, P. C. McIntyre, P. M. Asbeck, M. J. W. Rodwell, and Y. Taur, "A distributed model for border traps in Al_2O_3-InGaAs MOS devices," *IEEE Electron Device Letters*, vol. 32, no. 4, pp. 485–487, Apr. 2011. doi: 10.1109/LED.2011.2105241

[124] E. Lind, M. Hellenbrand, Lindelöw, and S. Andric, "Improved small signal model n-type, demonstrated p-type small-signal models," Lund University, Tech. Rep., 2018.

[125] M. Berg, O. Kilpi, K. Persson, J. Svensson, M. Hellenbrand, E. Lind, and L. Wernersson, "Electrical characterization and modeling of gate-last vertical InAs nanowire MOSFETs on Si," *IEEE Electron Device Letters*, vol. 37, no. 8, pp. 966–969, Jun. 2016. doi: 10.1109/LED.2016.2581918

[126] S. Andric, L. Ohlsson, and L.-E. Wenrersson, "Low-temperature front-side BEOL technology with circuit level multiline thru-reflect-line kit for III-V MOSFETs on silicon," in *2019 92nd ARFTG Microwave Measurement Conference*, Jan. 2019. doi: 10.1109/ARFTG.2019.8637222 pp. 1–4.

[127] S. Andric, L. Ohlsson Fhager, F. Lindelöw, O.-P. Kilpi, and L.-E. Wernersson, "Low-temperature back-end-of-line technology compatible with III-V nanowire MOSFETs," *Journal of Vacuum Science & Technology B*, vol. 37, no. 6, p. 061204, Oct. 2019. doi: 10.1116/1.5121017

[128] T.-C. Chang, K.-C. Chang, T.-M. Tsai, T.-J. Chu, and S. M. Sze, "Resistance random access memory," *Materials Today*, vol. 19, no. 5, pp. 254–264, Jun. 2016. doi: 10.1016/j.mattod.2015.11.009

[129] A. A. Vyas, A. Kumar, B. Sahu, C. K. P. Cheung, C.-P. Chang, C. Henderson, E. S. G. Bersuker, G. Klimeck, H. Bu, J. Stathis *et al.*, "International roadmap for devices and systems 2018 update," IEEE, Tech. Rep. [Online]. Available: https://irds.ieee.org/editions/2018/more-moore

APPENDICES

A

Measurement Setups

In this appendix, the measurement setups for low-frequency noise and small-signal parameters used for this thesis are summarised, so that future researchers may reassemble the setups.

A.1 LOW-FREQUENCY NOISE

The setup consists of an SR570 low-noise current preamplifier (transconductance amplifier), an SR830 lock-in amplifier, a Keysight B2912A source/measure unit, and a probe station to contact the devices. A lock-in amplifier is actually not the ideal tool, since measurements take a long time ($\sim$2 hours per device). A spectrum analyser would enable much faster measurements, but the intrinsic low-frequency noise of all spectrum analysers available in the measurement lab at the time of this thesis was too high for the analysers to be suitable for transistor characterisation. The preamplifier is ideally powered by its internal battery. A schematic of the connexions is provided in Fig. A.1(a). The MOFSET drain is connected to the preamplifier's 'Virtual Null Input' via a coax/triax cable and the preamplifier supplies a DC drain bias of 50 mV via its built-in bias source. 50 mV is a typical value to measure noise to get a measurable current while not bending the bands in the channel too much, which would smear out the energy resolution of the probed noise. Gate and source voltage (with the source as the reference terminal with $V_S = 0$ V) are supplied by the SMU and the source current as the device current for later calculations is recorded continually for each bias point.

The preamplifier 'Output' is connected to the lock-in amplifier 'Input A' via a coax cable and the lock-in amplifier is controlled by a LabView script,

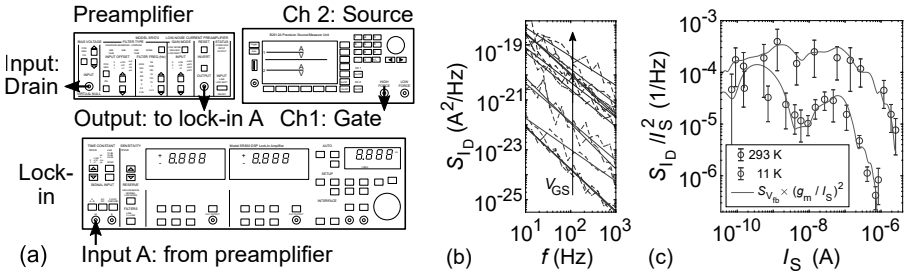


Figure A.1: (a) Measurement setup for LFN. (b) Frequency sweep. (c) Measurement at 10 Hz.

Table A.1: Settings for LFN measurement. P indicates preamplifier, L indicates lock-in amplifier.

	Setting	Value
P	Bias Voltage	on, pos, 50 mV
P	Filter Type	none
P	Gain Mode	Low Noise or High BW, depending on frequency range
P	Input Offset	off
P	Filter Freq	–
P	Sensitivity	depends on device current, see manual and 'Status' signal lamp
P	Invert	–
L	Signal Input	
	Input	A
	Couple	AC
	Ground	depends on setup
L	Filters	Line $+ 2 \times$ Line
L	Time Constant	higher: better S/N ration, longer measurement time, lower: worse S/N, faster measurement
L	Sensitivity	depends on device current
L	Reserve	Low Noise

which varies the measurement frequency. (A re-implementation of the script in MATLAB is highly recommended.) A complete noise characterisation typically consists of a frequency sweep for each gate bias, as presented in Fig. A.1(b), and a measurement at a fixed frequency of 10 Hz, which results in Fig. A.1(c). The different bias points should be distributed evenly over the current in the transfer curve of the transistor to provide evenly spaced points for the current axis in Fig. A.1(c). A transfer curve should be recorded before and after the noise measurement to document, if there is any significant drift during the measurement, and to provide the transconductance, which is required for the analysis. Possible drifts in the IV curve would have to be taken into account in the analysis. For a frequency sweep as in Fig. A.1(b), 50 measurements were taken per frequency, for the measurement at 10 Hz as in Fig. A.1(c), 12×50 measurements were taken for increased averaging.

Each gate bias point has to be set manually and the sensitivity of the amplifiers has to be adjusted to not exceed the compliance settings and the bandwidth. An excess of the current compliance is indicated by 'overload' signal lamps, but an excess of the bandwidth of the preamplifier is not indicated. Furthermore, at each bias point, the measured device current should be left to stabilise itself before the noise measurement is started, so that a constant device current can be recorded for each bias point. Ideally, the amplification of the preamplifier is set as high as possible, but the amplification is limited by the bandwidth of the amplifier, which is documented in the manual. For each bias point, the device current and the preamplifier amplification have to be recorded so that the current power spectral density S_{I_D} can be calculated from the measured lock-in voltage. Typical values for all settings are listed in Table A.1.

A.2 RADIO FREQUENCY MEASUREMENT

The setup for RF measurements consists of a vector-network-analyser (VNA) for the actual *s*-parameter measurement, a source/measure unit (SMU) to bias the transistor, and a probe station to contact the transistor. The SMU is connected, typically via coax cables, to the respective channel inputs of the VNA, which are used for the *s*-parameter measurement, and the VNA ports are connected to the probe station via cables specified for the respective measurement range. Special ground-signal-ground probes have to be used to contact the transistor pads, which have to be designed in such a way that they comply with RF specifications as well. See Fig. A.2 for schematics of the pad layout and the probes.

Before the actual measurement, the setup has to be calibrated to move the measurement reference plane to the tip of the probes. There are different

Figure A.2: (a) Schematics of two different kinds of RF probes. (b) Calibration structures for LRRM calibration: 50 Ω, short, 'thru'. (c) De-embedding pads, 'open' (left) and 'short' (right).

ways to calibrate the setup; all measurements for this thesis were calibrated by a load-reflect-reflect-match (LRRM) scheme. (Together with the back end of line process, for future use, a thru-reflect-line (TRL) calibration kit was developed by S. Andrić.) For the LRRM calibration, a substrate with dedicated calibration structures is used, see Fig. A.2(b). The calibration is performed by a dedicated software and then stored on the VNA, so that the VNA can automatically apply the calibration to the measurements. For ideal calibration, the exact substrate as well as the exact probes are specified in the calibration software. For this thesis, WinCal was used for calibration. A second step before the actual *s*-parameter measurement is required to be able to subtract the effect of the metal contact pads from the measurements later on. For this, dedicated 'de-embedding' structures have to be available on the measured sample; see Fig. A.2(c) for a schematic depiction of such structures. The de-embedding structures are basically the same as the actual contact pads, but without a transistor underneath. The 'open' structure presents an open circuit between source, gate, and drain, and the 'short' structure presents a short circuit between all terminals. *s*-parameters are measured on the de-embedding pads and stored for later use in data analysis.

Table A.2: Reference VNA settings for RF measurements used for this thesis.

Setting	Value
Measured Traces	s_{11}, s_{12}, s_{21}, s_{22}
Frequency Range	10 MHz–67 GHz
IF Bandwith	50 Hz
Sweep Type	logarithmic
Number of Points	401
RF Power	-27 dB

The actual measurement of the *s*-parameters of the transistors is ideally controlled with a script. Most of the time during the work for this thesis, this was achieved by a LabView script, but towards the end, an improved and completely automated MATLAB script was developed together with S. Andrić. *s*-parameters are typically measured at different bias points and over a specified frequency range, typically the full range of the VNA. The VNA settings, which were used for measurements in this thesis, are listed in Table A.2. The IF (intermediate frequency) bandwidth specifies, which bandwidth should be used at the intermediate down-conversion stages of the VNA. A smaller IF bandwidth results in more accurate measurements, but a longer measurement time.

B

Small-Signal Analysis in Detail

This document began as a personal sorting of notes, but was soon used as the basis for teaching material in a lab exercise of a Master's course, and finally grew into a comprehensive treatise on all the work conducted on RF measurements and modelling during my time as a PhD student. As such, there are many overlaps with Chapter 3 in this thesis and with Paper VI, but this document provides many more details. It remains with the *Nanoelectronics* group as e.g. an introduction for new PhD students and as possible teaching material for courses. The document can of course still be updated to improve it, but it works well as a self-contained and comprehensive overview in its current form.

The MOSFET Small-Signal Model

Markus Hellenbrand
markus.hellenbrand@gmail.com

13 May 2020

1 INTRODUCTION

This document is meant to capture (almost) all aspects of the small-signal model(s) that can be used to describe the high-frequency performance of III-V (MOS) transistors. While very detailed derivations of the general small-signal response of transistors can be found in e.g. [1] or [2], the abundance of information is often not suitable for practical work and important equations are often scattered to places asunder. The intention of this document is thus not to provide detailed derivations of all properties, but rather to provide a basic understanding and a (at least someewhat) consolidated summary of the small-signal model as well as a document for practical use to e.g. implement a script for the analysis of measured data. As such, this is sort of a collection of the most important equations from mostly [1] and [2] interspersed with accompanying explanations.

The first basic model, which is introduced, with small variations, can often be found in literature as the core of any small-signal transistor analysis. Example measurements on a lateral nanowire MOSFET will demonstrate success and limitations of this model. Based on these observations, more detailed additional effects such as the effect of oxide traps, band-to-band tunnelling, and impact ionisation are added to the core model later on. The effects of the different model components are illustrated by example measurements as well. Practical use of this document is intended to start with de-embedded measurement data. While there is certainly room for improvement, the document is self-contained and provides a comprehensive overview.

2 CALIBRATION, MEASUREMENT, DE-EMBEDDING

At high frequencies, it is not practical anymore to measure voltages and currents as in a DC measurement. Since at high frequencies, the cables used in a measurement setup are in the order of – or much larger than – the wavelength of the measured signals, they have to be treated as waves. This means that dedicated RF cables and probes have to be used in the measurement in order to avoid power losses in the signal. The response of the transistor is measured in the form of scattering parameters (s-parameters), which describe the transmission and reflection of the incident power waves at the terminals of the transistor. The theoretical analysis of the s-parameters treated in Chapter 4.

The measurement of these s-parameters is carried out with a Vector Network Analyser (VNA). All stages of the measurement setup, from the port of the VNA through the cables and the probes down to the sample, have to be impedance-matched in order to avoid reflections the intermediate interfaces. Different frequency ranges require different cable specifications and different measurement probes. Typically, common to all probes is the ground-signal-ground (GSG) configuration as illustrated in Fig. 1(a); the frequency range of cables is linked to their diameter, see Table 1 for an overview.

Table 1: Diameters, frequencies, and mating of RF cables.

Outer dielectric diameter	Frequency up to	Mates with
SMA	24 GHz	2.92 mm & 3.5 mm
3.5 mm	34 GHz	2.92 mm & SMA
2.92 mm	40 GHz	3.5 mm & SMA
2.4 mm	50 GHz	1.85 mm
1.85 mm	70 GHz	2.4 mm
1.0 mm	110 GHz	1.0 mm

Figure 1: Schematics of (a) two different RF probes, (b) the calibration structures used for an LRRM calibration (from left to right: load, short, through), and (c) open (left) and short (right) de-embedding structures to remove the effect of the transistor contact pads.

In order to be able to measure the intrinsic response of a transistor, first, the reference plane of the measurement has to be moved from the port of the VNA to the tip of the measurement probes. This is achieved with a calibration of the measurement setup with the help of dedicated calibration structures. Fig. 1(b) provides schematics of the structures, which are used for a load-reflect-reflect-match (LRRM) calibration. The reflection and transmission coefficients of these structures are well known and by measuring the structures and comparing the results with the known coefficients, the measurement reference plane can be moved accordingly. (The calculations are done by a software, equations not included here.)

Since the physical dimensions of transistors are much smaller than the probe tips, large metal pads have to be added on the transistor samples in order to be able to establish electrical contact between the transistor and the measurement setup. These metal pads add large capacitances and inductances to the RF response of the total structure so that the reference plane has to be move one more time to remove the effect of the pads. This process is called *de-embedding* and in the case of an *open-short* de-embedding it is carried out with the help of structures as the e.g. ones in Fig. 1(c). These de-embedding structures have to be exactly the same as the metal contact pads, except that the transistor is not present in the structure. As can be seen in Fig. 1(c), one structure represents an open circuit in the centre, and the other a short circuit. The measured response of the pads can be subtracted from the measured response of the complete structure (with transistor) so that the resulting response contains only the effect of the intrinsic transistor.

3 SMALL-SIGNAL PARAMETERS

As mentioned previously, the measured quantity in a high frequency characterisation are s-parameters, which describe the reflection and transmission of power waves incident to the transistor. For this characterisation method, the transistor is represented as a *two-port network*, where the gate corresponds to the input (port 1), the drain to the output (port 2), and the source is the common ground terminal. This representation is visualised in Fig. 2, where the a_j and b_j are the power flowing into and out of the ports, respectively, as

$$a_j = \frac{v_j + Z_0 i_j}{2\sqrt{Z_0}} \qquad (1a) \qquad\qquad b_j = \frac{v_j - Z_0 i_j}{2\sqrt{Z_0}}. \qquad (1b)$$

Here, v_j and i_j are the voltage and the current, respectively, at port j, and Z_0 is the characteristic impedance of the system. With (1a) and (1b), the complete two-port response can be described with the so-called $\boldsymbol{S}$-matrix as

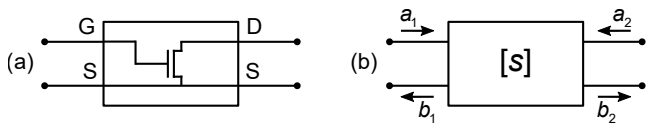


Figure 2: Two-port representation of a transistor. (a) Illustration of the port assignment. G, D, and S stand for gate, drain, and source, respectively. (b) Power input a and output b.

$$\begin{bmatrix} b_1 \\ b_2 \end{bmatrix} = \begin{bmatrix} s_{11} & s_{12} \\ s_{21} & s_{22} \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \end{bmatrix} \quad (2) \qquad \text{with} \qquad s_{kl} = \left.\frac{b_k}{a_l}\right|_{a_{\neg l}=0}. \quad (3)$$

For modelling, instead of s-parameters, admittance parameters (y-parameters) are much more intuitive, since they can be derived for a transistor small-signal model in a straightforward manner. Formally, the set of four y-parameters is completely equivalent to the set of s-parameters described by (2) and (3). The formal relation between the s-parameters and the y-parameters is [1, p. 249]

$$y'_{11} = \frac{(1-s_{11})(1+s_{22})+s_{12}s_{21}}{(1+s_{11})(1+s_{22})-s_{12}s_{21}}, \quad (4a) \qquad y'_{12} = \frac{-2s_{12}}{(1+s_{11})(1+s_{22})-s_{12}s_{21}}, \quad (4b)$$

$$y'_{21} = \frac{-2s_{21}}{(1+s_{11})(1+s_{22})-s_{12}s_{21}}, \quad (4c) \qquad y'_{22} = \frac{(1+s_{11})(1-s_{22})+s_{12}s_{21}}{(1+s_{11})(1+s_{22})-s_{12}s_{21}}. \quad (4d)$$

Note that the y'_{kl} in (4) are normalised with the characteristic impedance Z_0 of the system (usually 50 Ω) so that the unnormalised $y = y'/Z_0$.

While (4) may not be intuitive, the definitions of the y-parameters will be very intuitive, since they relate the y-parameters to the physical properties of the transistor. The y-parameters are completely equivalent to the s-parameters. Generally, the admittance y is the reciprocal of the impedance z, so that $y = 1/z = i/v$, where i is a current and v a voltage. In the small-signal model, i and v correspond to small AC values at the ports of the transistor around the DC large-signal bias point. The underlying and important assumption here is that the response of a transistor to a small AC signal can be approximated as linear around the DC bias point:

$$i_{\text{total}}(V_{\text{DC}} + \delta v_{\text{ac}}) \approx i_{\text{total}}(V_{\text{DC}}) + \frac{\partial i_{\text{total}}}{\partial v_{\text{ac}}}\delta v_{\text{ac}} \approx I_{\text{DC}} + i_{\text{ac}}. \quad (5)$$

Here, i_{total} is the total device current, uppercase letters with index DC denote (large-signal) DC values, and lowercase letters with index ac denote small AC variations. An example for the total source-to-drain current i_{DS} of a transistor as function of the gate and drain bias v_{GS} and v_{DS}, respectively, would thus become

$$i_{\text{DS}}(v_{\text{GS}}, v_{\text{DS}}) \approx I_{\text{DS}}(V_{\text{GS}}, V_{\text{DS}}) + \frac{\partial i_{\text{DS}}}{\partial v_{\text{gs}}}\delta v_{\text{gs}} + \frac{\partial i_{\text{DS}}}{\partial v_{\text{ds}}}\delta v_{\text{ds}} = I_{\text{DS}}(V_{\text{GS}}, V_{\text{DS}}) + g_{\text{m}}\delta v_{\text{gs}} + g_{\text{ds}}\delta v_{\text{ds}}. \quad (6)$$

Here, lowercase letters with uppercase indices denote quantities consisting of a (large-signal) DC and a (small-signal) AC part, g_{m} is the transconductance, and g_{ds} the output conductance. The other notations are as before.

In (6) we see two specific examples of admittances: the transconductance g_{m} and the output conductance g_{ds}. More generally, the admittances y_{kl} are complex values and their general definition, analogous to (3, is

$$y_{kl} = \left.\frac{i_k}{v_l}\right|_{v_{\neg l}=0} \quad (7)$$

with the currents i_k and voltages v_l at the respective ports k and l.

4 THE BASIC SMALL-SIGNAL MODEL

With the general definition of (7), we can now approach the small-signal model for a MOSFET and derive the concrete y-parameters. A basic small-signal model, which captures the essential RF response of a MOSFET in the on-state, is provided in Fig. 3. The physical meaning of the different components will be explained below. First, for practicality, the equations for the y-parameters are provided close to the figure.

4.1 THE INTRINSIC MODEL PARAMETERS

The expressions for the y-parameters become much easier, if the gate, source, and drain resistances R_{g}, R_{s}, and R_{d}, are removed first. The procedure for this, as well as a way to determine their values, will

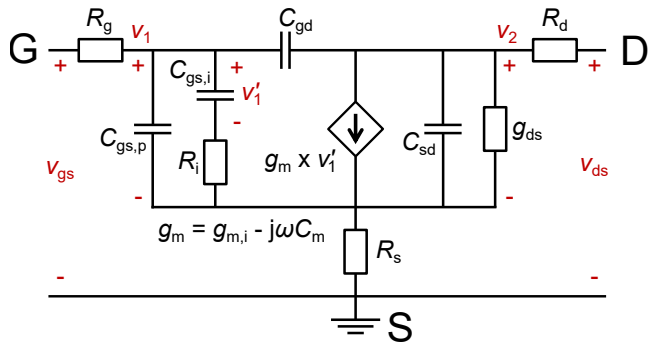


Figure 3: The most basic small-signal model that captures the essential intrinsic (and parasitic) high-frequency properties of a MOSFET.

be explained in 4.2. After removing the resistances, with (7), and with the angular frequency ω and the imaginary unit j, the *intrinsic* y-parameters for Fig. 3 (with v_1 and v_2 marked) become

$$y_{11} = \left.\frac{i_1}{v_1}\right|_{v_2=0} = \mathrm{j}\omega\left(C_{\mathrm{gs,p}} + C_{\mathrm{gd}}\right) + \frac{\mathrm{j}\omega C_{\mathrm{gs,i}}}{1 + \mathrm{j}\omega C_{\mathrm{gs,i}} R_{\mathrm{i}}} \tag{8a}$$

$$y_{12} = \left.\frac{i_1}{v_2}\right|_{v_1=0} = -\mathrm{j}\omega C_{\mathrm{gd}} \tag{8b}$$

$$y_{21} = \left.\frac{i_2}{v_1}\right|_{v_2=0} = \frac{g_{\mathrm{m,i}} - \mathrm{j}\omega C_{\mathrm{m}}}{1 + \mathrm{j}\omega R_{\mathrm{i}} C_{\mathrm{gs,i}}} - \mathrm{j}\omega C_{\mathrm{gd}} \tag{8c}$$

$$y_{22} = \left.\frac{i_2}{v_2}\right|_{v_1=0} = g_{\mathrm{ds}} + \mathrm{j}\omega\left(C_{\mathrm{sd}} + C_{\mathrm{gd}}\right). \tag{8d}$$

If the denominators in the fractions of y_{11} and y_{21} are made real, they become $1 + \omega^2 R_{\mathrm{i}}^2 C_{\mathrm{gs,i}}^2$. For large parts of the typical measurement range (up to 67 GHz), $\omega^2 R^2 C^2 \ll 1$, which simplifies these new denominators to approximately one. The physical meaning of the different components is as follows:

- R_{s}, R_{d}, and R_{g} are the source, drain, and gate resistance respectively. They contain the contact and access resistance of the different terminals.
- $C_{\mathrm{gs,p}}$ is the parasitic gate-to-source capacitance. It originates in the physical extension of the gate metal, which forms parasitic capacitances with the source of the MOSFET.
- $C_{\mathrm{gs,i}} = -\partial Q_{\mathrm{g}}/\partial v_{\mathrm{s}}$ is the intrinsic gate-to-source capacitance, which models the charge on the gate – and thus in the channel – as a function of the source small-signal voltage. (Note that $C_{\mathrm{gs,i}}$ actually comes from rewriting the the y-parameters of the initial common source model, so that the definition makes sense despite $v_{\mathrm{s}} = 0$ in Fig. 3.)
- R_{i} is the intrinsic channel resistance. It is not a resistance in the ohmic sense and it is not the same as the channel resistance as observed in DC measurements, either. Instead, its origin is the finite time that it takes for charge carriers to move through the channel. A short derivation is provided below this list.
- C_{gd} is the gate-to-drain capacitance. In this simplified model it contains both the intrinsic $C_{\mathrm{gd,i}} = -\partial Q_{\mathrm{g}}/\partial v_{\mathrm{d}}$ and the parasitic $C_{\mathrm{gd,p}}$. Both $C_{\mathrm{gd,i}}$ and $C_{\mathrm{gd,p}}$ have origins analogous to the gate-to-source capacitances. In the case of C_{gd}, the parasitic contribution usually dominates, so that the intrinsic $C_{\mathrm{gd,i}}$ can be disregarded. For the same reason, an intrinsic resistance R_{j}, analogous to the intrinsic R_{i}, is disregarded. In the more detailed model later on, R_{j} will be included.
- For the controlled current source $g_{\mathrm{m}} = g_{\mathrm{m,i}} - \mathrm{j}\omega C_{\mathrm{m}}$, $g_{\mathrm{m,i}}$ is the intrinsic transconductance and C_{m} is the mutual differential capacitance, which balances the charge in the channel. It is defined as $C_{\mathrm{dg}} - C_{\mathrm{gd}}$.
- C_{sd} is the source-to-drain capacitance. Physically, just as C_{gd} and $C_{\mathrm{gd,i}}$, C_{sd} should appear in series with a corresponding resistance to model the delay of charge carriers in the channel. In the case of C_{sd}, however, such a resistance would always be masked by the relatively large g_{ds} in

parallel with C_{sd}, so that the resistance is omitted.
- g_{ds} is the output conductance of the transistor.

A Short Derivation of R_{i} [3]
In DC operation and at low frequencies, the movement of charge carriers in the transistor is much faster than the change in the bias conditions so that it can be assumed that the charge carriers inside the transistor adapt to a change in the bias conditions instantaneously. If the frequencies of operation approach the inverse of the time it takes for charge carriers to traverse the channel, this cannot be assumed any longer. The delay due to the finite velocity of the charge carriers can be modelled with R_{i} and the intrinsic capacitance $C_{\mathrm{gs,i}}$.

The time it takes charge carriers to traverse half the channel of a MOSFET is approximately $t_{1/2} = L_{\mathrm{G}}/(2v^+)$, where L_{G} is the gate length and v^+ is the average velocity of the charge carriers. This can be equated with the time $t_{\mathrm{RC}} = \ln(2)\tau = 0.7R_{\mathrm{i}}C_{\mathrm{gs,i}}L_{\mathrm{G}}W_{\mathrm{G}}$ that it takes to halfway charge the RC element consisting of R_{i} and $C_{\mathrm{gs,i}}$. W_{G} is the width of the transistor. Equating $t_{1/2}$ and t_{RC} yields $R_{\mathrm{i}} = 1/(1.4W_{\mathrm{G}}C_{\mathrm{gs,i}}v^+)$. This can be further related to the transconductance $g_{\mathrm{m,i}}$ since $g_{\mathrm{m,i}} = (\partial Q/\partial v_{\mathrm{gs}})v^+ + Q(\partial v^+/\partial v_{\mathrm{gs}}) \approx (\partial Q/\partial v_{\mathrm{gs}})v^+ \approx W_{\mathrm{G}}C_{\mathrm{gs,i}}v^+$, so that $R_{\mathrm{i}} \approx 1/(1.4g_{\mathrm{m,i}})$. This approach corresponds to the so-called relaxation time model. A further short explanation can be found in e.g. [4] and more details can be found in e.g. [5]. Physically, to iterate, intrinsic resistances corresponding to R_{i} should be placed in series with the intrinsic C_{gd} and C_{sd}, as mentioned above. For C_{gd}, this resistance will be added in a more detailed model later on; for C_{sd}, the effect of the intrinsic resistance is masked by the relatively large g_{ds} in parallel.

4.2 OFF-STATE, ACCESS RESISTANCES, AND PARASITIC CAPACITANCES

In order to be able to use the expressions for the intrinsic y-parameters in (8), the external resistances R_{g}, R_{s}, and R_{d} have to removed. One way to determine their values is a measurement of the s-parameters in the off-state of the transistor, a so-called *cold-FET measurement*. For such a measurement, if the gate voltage V_g is chosen far in the off-state of the transistor and the drain voltage is set to 0 V, the small-signal model of Fig. 3 reduces to that in Fig. 4. The elements, which are missing compared with Fig. 3, disappear because there is no current flowing through the device.

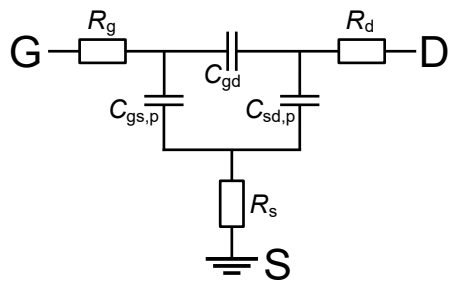


Figure 4: Small-signal model in the off-state.

At sufficiently high frequencies (above about a few GHz), the capacitances in Fig. 4 will short-circuit possibly still present intrinsic conductances so that the real part of the impedance will be dominated by the resistances R_{g}, R_{s}, and R_{d}. The real part of the matrix of impedance parameters (z-parameters) then becomes

$$\boldsymbol{Z}_{\mathbf{R}} = \mathrm{Re}(\boldsymbol{Z}_{\textbf{off-state}}) = \begin{bmatrix} R_{\mathrm{s}} + R_{\mathrm{g}} & R_{\mathrm{s}} \\ R_{\mathrm{s}} & R_{\mathrm{s}} + R_{\mathrm{d}} \end{bmatrix}, \tag{9}$$

so that the resistances can be read easily from plotting the real parts of the z-parameters. Just as for the y-parameters, the set of z-parameters is completely equivalent to the measured s-parameters, so that the z-parameters can be calculated directly from the measured s-parameters. For completeness, the corresponding equations are listed below [1, p. 249]:

$$z'_{11} = \frac{(1+s_{11})(1-s_{ss}) + s_{12}s_{21}}{(1-s_{11})(1-s_{22}) - s_{12}s_{21}} \tag{10a}$$

$$z'_{12} = \frac{2s_{12}}{(1-s_{11})(1-s_{22}) - s_{12}s_{21}} \tag{10b}$$

$$z'_{21} = \frac{2s_{21}}{(1-s_{11})(1-s_{22}) - s_{12}s_{21}} \tag{10c}$$

$$z'_{22} = \frac{(1-s_{11})(1+s_{22}+s_{12}s_{21})}{(1-s_{11})(1-s_{22}) - s_{12}s_{21}} \tag{10d}$$

Again, just as for the y-parameters, the z' are normalised to the characteristic impedance of the system so that the unnormalised $z = z'Z_0$.

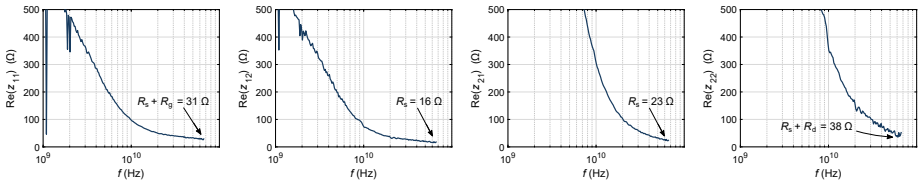


Figure 5: Real parts of the off-state impedances to determine R_g, R_s, and R_d.

An example of a (semi-successful) cold-FET measurement of a vertical nanowire MOSFET is presented in Fig. 5. As can be seen, the curves do not completely saturate at high frequencies, which can can be a sign of parasitic leakage paths, which are not taken into account in the simple model in Fig. 4. In this example, the obtained values for the resistances seem reasonable, which can be verified by comparison with DC measurements of the transistor. Sometimes, however, if parasitic leakages are too large, or the transistor does not turn off properly, the curves in Fig. 5 do not saturate at all. In that case, estimations from DC measurements are the only way to obtain approximate values for R_g, R_s, and R_d. In any case, comparison with the values obtained from DC measurements is good practice.

Once R_g, R_s, and R_d are determined, they can be subtracted from the complete measured z-matrix of the transistor just as $\boldsymbol{Z}_\mathbf{removed} = \boldsymbol{Z}_\mathbf{measured} - \boldsymbol{Z}_\mathbf{R}$. In the next step, the z-parameters $\boldsymbol{Z}_\mathbf{removed}$ without the resistances can be transformed to y-parameters, which now only consist of the parasitic capacitances in Fig. 4 as

$$\boldsymbol{Y}_\mathbf{removed} = \begin{bmatrix} \mathrm{j}\omega(C_\mathrm{gs,p} + C_\mathrm{gd}) & -\mathrm{j}\omega C_\mathrm{gd} \\ -\mathrm{j}\omega C_\mathrm{gd} & \mathrm{j}\omega(C_\mathrm{gd} + C_\mathrm{sd}) \end{bmatrix}. \tag{11}$$

Examples for determining the parasitic capacitances according to (11) are provided in Fig. 6. (For instructional purposes, this is not the same device as in Fig. 5, but a lateral nanowire MOSFET.) Once, values for the access resistances and the parasitic capacitances are obtained, the y-parameters from (8) can be used to determine the remaining small-signal parameters.

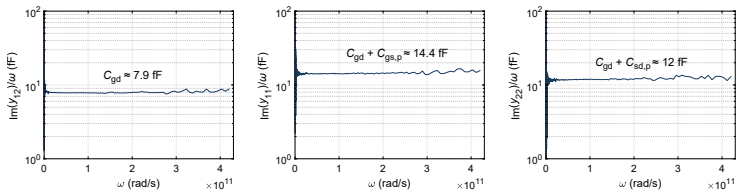


Figure 6: Imaginary parts of the y-parameters from (11) to determine the parasitic capacitances.

4.3 FITTING THE INTRINSIC PARAMETERS

With the access resistances and the parasitic capacitances known, the intrinsic y-parameters from (8) can now be fitted to the measured data. Fitting is vastly simplified by the observation that in (8a), $\omega^2 R_\mathrm{i}^2 C_\mathrm{gs,i}^2 \ll 1$ and in (8c), $\omega R_\mathrm{i} C_\mathrm{gs,i} \ll 1$, so that all parameters can be straightforwardly fitted to the linear or squared real and imaginary parts of the y-parameters. Once a parameter is determined, it can be

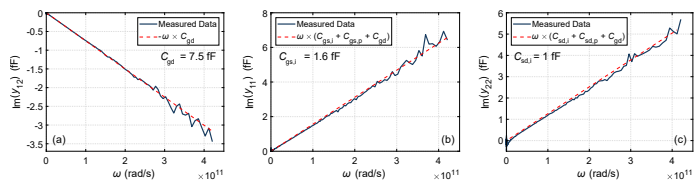


Figure 7: Fits to (a) $\mathrm{Im}(y_{12})$, (b) $\mathrm{Im}(y_{11})$, and (c) $\mathrm{Im}(y_{22})$ to obtain C_gd, $C_\mathrm{gs,i}$, and $C_\mathrm{sd,i}$, respectively.

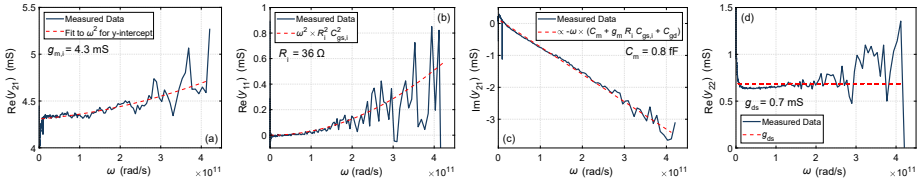


Figure 8: Fits to (a) Re(y_{21}), (b) Re(y_{11}), (c) Im(y_{21}), and (d) Re(y_{22}) to obtain $g_{\mathrm{m,i}}$, R_{i}, C_{m}, and g_{ds}, respectively.

used in subsequent fits to determine the remaining parameters. Examples for the fits to Im(y_{12}), Im(y_{11}), and Im(y_{22}) to obtain C_{gd} (again), $C_{\mathrm{gs,i}}$, and $C_{\mathrm{sd,i}}$, respectively, are provided in Fig. 7, and examples for the fits to Re(y_{21}), Re(y_{11}), Im(y_{21}), and Re(y_{22}) to obtain $g_{\mathrm{m,i}}$, R_{i}, C_{m}, and g_{gd}, respectively, are provided in Fig. 8. These examples are for the same lateral nanowire MOSFET as in Fig. 6.

When comparing the values for C_{gd}, $C_{\mathrm{gs,i}}$, and $C_{\mathrm{sd,i}}$ with the parasitic values from Fig. 6, it becomes obvious that the capacitances in these devices are largely dominated by parasitics. Especially C_{gd} is almost the same in the off-state (Fig. 6) and in the on-state (Fig. 7).

For R_{i} it can be noted that the value obtained from the fit is smaller than what would be expected from the approximative expression $R_{\mathrm{i}} \approx 1/(1.4 g_{\mathrm{m,i}})$. It is often difficult to accurately separate the effects of the extrinsic gate resistance R_{g} and the intrinsic R_{i} so that the "real" value for R_{i} probably lies somewhere between the fitted and the calculated value. In practice, the value can be varied between these two to obtain the best fits for the overall y-parameters.

4.4 BIAS DEPENDENCE OF INTRINSIC PARAMETERS

The intrinsic small-signal parameters depend on the bias condition of the transistor. The conductances $g_{\mathrm{m,i}}$ and g_{gd} essentially follow the same behaviour as in DC operation. For the capacitances, when the gate bias increases, the coupling between the channel and the source/drain terminals increases, so that the intrinsic $C_{\mathrm{gs,i}}$ and $C_{\mathrm{gd,i}}$ increase. The constant values for C_{gs} and C_{gd} below the threshold voltage are the parasitic contributions, which are not bias-dependent. For increasing drain bias (not shown), $C_{\mathrm{gs,i}}$ should remain approximately constant and $C_{\mathrm{gd,i}}$ should decrease, since the larger constant V_{GD} results in a relatively lower impact of the small-signal excitation at the drain. Examples of the gate bias dependences of $g_{\mathrm{m,i}}$ and g_{gd}, as well as C_{gd} and C_{gs} are provided in Fig. 9(a) and Fig. 9(b), respectively. These examples are for a vertical nanowire MOSFET, where the parasitic capacitances have been optimised to a certain extent, so that also the intrinsic C_{gd} emerges from the parasitic values. The changes in C_{gs} and C_{gd} can be qualitatively explained quite easily by thinking of the band structures along the channel at different V_{GS}.

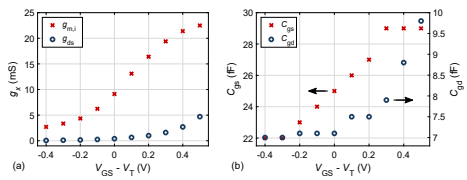


Figure 9: Gate bias dependence of (a) $g_{\mathrm{m,i}}$ and g_{ds} and (b) C_{gs} and C_{gd}. $g_{\mathrm{m,i}}$ and g_{ds} follow a similar behaviour as in DC operation. The constant values of the capacitances below V_{T} are parasitics, the increase results from the intrinsic capacitances.

4.5 *Y*-PARAMETERS INCLUDING THE RESISTANCES

The actual gains and maximum operation frequencies of a transistor will be affected by the access resistances as well. In the end, it is thus necessary to include these resistances.

The general transformation between y- and z-parameters is actually much more straightforward than between either of the two and the s-parameters, as in (4) and (10). The denominators Δy Δz in the following equations are $\Delta y = y_{11}y_{22} - y_{12}y_{21}$ and (same with z) $\Delta z = z_{11}z_{22} - z_{12}z_{21}$.

$$\boldsymbol{Z} = \begin{bmatrix} \frac{y_{22}}{\Delta y} & \frac{-y_{12}}{\Delta y} \\ \frac{-y_{21}}{\Delta y} & \frac{y_{11}}{\Delta y} \end{bmatrix} \quad (12) \qquad \text{and} \qquad \boldsymbol{Y} = \begin{bmatrix} \frac{z_{22}}{\Delta z} & \frac{-z_{12}}{\Delta z} \\ \frac{-z_{21}}{\Delta z} & \frac{z_{11}}{\Delta z} \end{bmatrix}. \quad (13)$$

The parameters can access resistances can then be added to $\boldsymbol{Z}$ simply in form of the matrix from (9).

4.6 GAINS AND MAXIMUM OPERATION FREQUENCIES

The intrinsic properties of a MOSFETs determine the performance, which it can deliver, e.g. in an amplifier. Here, briefly, some of these performance metrics will be reviewed.

Two common benchmark metrics for the RF performance of transistors are the cutoff frequency f_T and the maximum oscillation frequency f_max. These are the frequencies, where the forward current gain h_{21} and the unilateral power gain U, respectively, become equal to one. The two gains are defined as [6]

$$h_{21} = \frac{|y_{21}|}{|y_{11}|} \quad (14) \qquad \text{and} \qquad U = \frac{|y_{21} - y_{12}|^2}{4\left[\mathrm{Re}(y_{11})\mathrm{Re}(y_{22}) - \mathrm{Re}(y_{12})\mathrm{Re}(y_{21})\right]}, \quad (15)$$

where for U, z-parameters with the same indices can be used as well. We will now investigate expressions to describe h_{21}, f_T, U, and f_max based on the y-parameters. In general, the actual gain that a transistor can provide is reduced by the access resistances R_g, R_s, and R_d. We can thus not simply use the intrinsic y-parameters from (8) and the resulting expressions can become quite tedious. The expressions below will thus already be simplified by omitting higher-order terms in ω and by omitting small contributions. The expressions will simplify further, when R_s and R_d can be made small. Note, however, that the following expressions do not take into account the intrinsic NQS resistances R_i and R_j. R_i and R_j are instead lumped with the gate resistance R_g. For the calculation of h_{21}, f_T, U, and f_max, the difference does not matter much, since these values are 'what comes out of the total device'. Just keep in mind that for the analysis of the intrinsic parameters, R_i and R_j should be handled separately.

For $h_{21} = |y_{21}|/|y_{11}|$, we can obtain

$$h_{21} \approx \frac{g_\mathrm{m}}{\omega[C_\mathrm{gg} + (R_\mathrm{s} + R_\mathrm{d})(C_\mathrm{gg}\, g_\mathrm{ds} + C_\mathrm{gd}\, g_\mathrm{m})]}, \quad (16)$$

where C_gg is the total gate capacitance $C_\mathrm{gs,p} + C_\mathrm{gs,i} + C_\mathrm{gd}$. From this, an expression for f_T can be obtained by setting $|h_{21}| = 1$ and $\omega = 2\pi f_\mathrm{T}$ [7]:

$$f_\mathrm{T} \approx \frac{1}{2\pi}\left[\frac{C_\mathrm{gg,t}}{g_\mathrm{m}} + (R_\mathrm{s} + R_\mathrm{d})\left(C_\mathrm{gd,t} + \frac{g_\mathrm{ds}}{g_\mathrm{m}} C_\mathrm{gg,t}\right)\right]^{-1} \quad (17)$$

If R_s and R_d can be made (very) small, (16) and (17) simplify to

$$h_{21} \approx \frac{g_\mathrm{m}}{\omega C_\mathrm{gg}} \quad (18) \qquad \text{and} \qquad f_\mathrm{T} \approx \frac{g_\mathrm{m,i}}{2\pi C_\mathrm{gg}}, \quad (19)$$

which corresponds to expressions that can in fact be obtained from the intrinsic y-parameters including R_i and R_i as well. In order to arrive at (18) and (19) from the intrinsic y-parameters of (8), approximations have to be used that $\omega^2 R_\mathrm{i} C_\mathrm{gs,i}^2 \ll 1$ and $\omega^2 R_\mathrm{i}^2 C_\mathrm{gs}^2 \ll 1$ in y_{11} and the observation that in the on-state of a well-behaved MOSFET, g_m is the largest quantity in the numerator of y_{21}.

For U according to (15) it can already be divined that the expression with inserted parameters becomes quite complicated and not very employable for analytic calculations. Inserting the parameters is thus skipped here and instead, we continue with the more useful, if still clunky, expression for f_max, which, analogous to f_T, can be derived by setting $U = 1$ and $\omega = 2\pi f_\mathrm{max}$:

$$f_\mathrm{max} = \sqrt{\frac{f_\mathrm{T}}{8\pi R_\mathrm{g}\left(C_\mathrm{gd,t} + 2\pi f_\mathrm{T} \Psi\right)}} \quad (20)$$

with [8]

$$\Psi = \frac{g_\mathrm{ds}}{g_\mathrm{m}^2}\left[C_\mathrm{gg,t}^2 + \frac{R_\mathrm{s}}{R_\mathrm{g}}(C_\mathrm{gs,i} + C_\mathrm{gs,p})^2 + \frac{R_\mathrm{i}}{R_\mathrm{g}} C_\mathrm{gs,i}^2 + \frac{R_\mathrm{d}}{R_\mathrm{g}} C_\mathrm{gd,t}^2\right]. \quad (21)$$

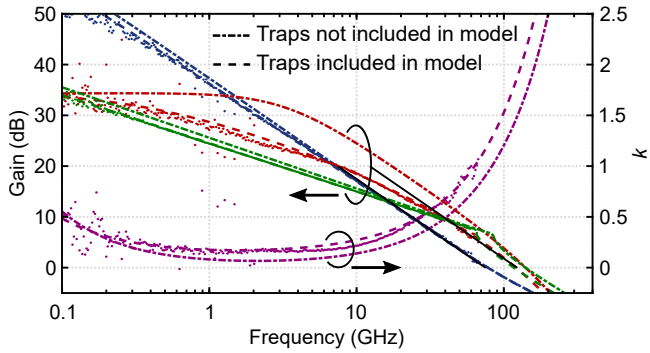


Figure 10: Forward current gain h_{21}, unilateral power gain U, maximum stable/available gain MSG/MAG, and stability factor k. Two models are provided to illustrate the limitations of the basic small-signal model and to anticipate the resulting extended model. h_{21} and U are extrapolated with -20 dB/decade to determine the cutoff frequency f_T and the maximum oscillation frequency f_max, respectively. The resistances R_s, R_d, and R_g are included in the calculation of the plotted quantities, so that the values for f_T and f_max are a bit lower than what the approximations in (19) and (22) would yield.

As before, if R_s and R_d can be made (very) small, the expression for f_max simplifies somewhat to

$$f_\mathrm{max} = \sqrt{\frac{f_\mathrm{T}}{8\pi R_\mathrm{g}\left(C_\mathrm{gd} + \frac{g_\mathrm{ds}}{g_\mathrm{m}} C_\mathrm{gg}\right)}}. \tag{22}$$

For further discussions about all expressions, see also [4] and [7].

All expressions demonstrate that for high gain as well as for high f_T and f_max, a high transconductance and low (especially parasitic) capacitances as well as a low output conductance are required. This also holds for the maximum stable gain MSG and the maximum available gain MAG, which are

$$\mathrm{MSG} = \frac{|y_{21}|}{|y_{12}|} \tag{23}$$

and

$$\mathrm{MSG} = \frac{|y_{21}|}{|y_{12}|}\left(k - \sqrt{k^2 - 1}\right). \tag{24}$$

k in (24) is the stability factor, which is defined as

$$k = \frac{2\mathrm{Re}(y_{11})\mathrm{Re}(y_{11}) - \mathrm{Re}(y_{12}y_{21})}{|y_{12}y_{21}|}. \tag{25}$$

From (18) and (with some algebra) from (15) it can be seen that with increasing frequencies, h_{21} and U roll off with -20 dB/decade. This is indicated in Fig. 10, where extrapolations of -20 dB/decade are used to determine f_T and f_max.

This figure then also leads over to the next section in that it already demonstrates one of the limitations of the basic small-signal model: The unilateral power gain U clearly changes its slope towards lower frequencies, which the model so far cannot describe. Upon closer inspection, it turns out that also MSG and h_{21} deviate a little from their ideal slopes of -10 and -20 dB/decade, respectively.

5 THE EXTENDED SMALL-SIGNAL MODEL

In this section, the basic small-signal model from above will be examined in more detail. It will turn out quite clearly that there are some features, which the basic model cannot properly describe. These effects can be traced back to the presence of defects related to the gate oxide, which can affect MOSFET operation even at gigahertz frequencies. Furthermore, impact ionisation (II) and band-to-band tunnelling (BTBT) on the drain side can have an impact on the y-parameters at sufficiently high gate-to-drain electric fields. Here, first, the effect of traps is discussed, followed by the effects of II and BTBT. These

effects are taken into account by modified equations for the respective parameters. Secondly, the revised model with all equations and an updated equivalent circuit are summarised in a dedicated section to gather everything in one place. Finally, the effects of the revised model on the gains and the stability factor are examined.

In measurements, so far, vertical nanowire MOSFETs turned out to be affected by oxide traps more than lateral nanowire MOSFETs or even planar ones. The following effects will thus be illustrated with measured data from vertical nanowire MOSFETs.

5.1 EXTENSION OF THE MODEL – OXIDE TRAPS

The effects of traps emerge when plotting the y-parameters on logarithmic scales, as illustrated in Fig. 12 (further on): $\mathrm{Re}(y_{11})$, $\mathrm{Re}(y_{12})$, $\mathrm{Re}(y_{21})$, $\mathrm{Im}(y_{11})$, and $\mathrm{Im}(y_{11})$ all exhibit a frequency dispersion, which cannot be recreated by the small-signal model without the inclusion of traps. In the following, it will be demonstrated, how these traps can be included.

5.1.1 MODELLING TRAPS BY A DISTRIBUTED *RC* NETWORK

Each trap is associated with a certain capture/emission time constant, which determines the frequency, up to which the respective trap can interact with carriers in the channel. The time constants depend on the distance x from the channel into the oxide and on an activation energy E_A due to local deformation upon capture or emission of charge. The time constants for the two processes can be asymmetric (due to different activation energies), but both are defined as

$$\tau_{c/e} = \tau_0 \exp\left(\frac{x}{\lambda}\right) \exp\left(\frac{E_\mathrm{A}}{k_\mathrm{k} T}\right) \tag{26}$$

with the Boltzmann constant k_B, the temperature T the tunnelling attenuation length λ according to the WKB approximation, and the pre-factor $\tau_0 = 1/(\sigma n v_\mathrm{th})$, which depends on the trap capture cross section σ, the thermal velocity v_th of the carriers, and the carrier concentration n. Since a wide range of values for σ can be found in literature, τ_0 varies a lot as well in different publications and values can be found to range from microseconds all the way to below picoseconds. For the small-signal models here, a value for τ_0 in the range of 1 ps usually describes the measured data best and as a sanity check, $\tau \approx 1$ ps results in very reasonable values of $\sigma \approx 10^{-15} - 10^{-14}$ cm^2.

As a side node, especially at low frequencies, such as in reliability or random telegraph noise measurements, the activation energy E_A usually dominates the time constant $\tau_\mathrm{c/e}$, so that any information about the position x is effectively hidden and the corresponding exponential term in (26) is merged with the pre-factor to an effective pre-factor τ_0'. Literatures values for this effective τ_0' are often in the range of nanoseconds, which means that for traps to be able to respond at gigahertz frequencies, E_A has to be rather small and the exchange mechanism approaches elastic tunnelling.

Irrespective of the exact charge exchange mechanism, a time constant can be modelled by a simple RC element and a distributed trap distribution can thus be modelled by a distributed RC network such as the one in Fig. 11. The capacitances can be assumed as constant and each time constant is defined by a corresponding resistance/conductance so that $\tau = \Delta C_\mathrm{t}/\Delta G_\mathrm{t}$ with the parameters as in Fig. 11. The index t denotes that these time constants are related to traps. Such a model was used e.g. in [9] to model the frequency dispersion in CV measurements, in [10] to model the frequency dispersion in the RF transconductance, in [11] to model conductive losses, and in [12] to model the trap capacitances. In [9] it was also demonstrated that for a certain frequency range, the total distributed network can be lumped into an equivalent conductance with a linear frequency dependence in parallel with a capacitance with a

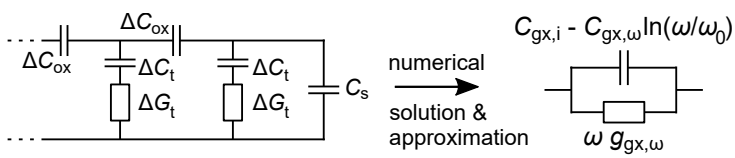


Figure 11: Distributed RC model and its lumped representation with corresponding frequency dependences.

logarithmic frequency dependence. This is illustrated in Fig. 11 as well and constitutes the basis for the models in [11] and [12]. The effects on the small-signal response are demonstrated in the following.

5.1.2 EFFECTS ON THE *Y*-PARAMETERS

If solved numerically, the distributed RC network results in a logarithmic variation of the real and imaginary part of the surface potential over C_s in Fig. 11. This logarithmic dependence has to be included in the transconductance g_m in the small-signal model. The examples in Fig. 12(a) demonstrate that this describes the measured data well: From below a few gigahertz down to 10 MHz, $\mathrm{Re}(y_{21}) \approx g_\mathrm{m}$ decreases in a logarithmic manner. A similar behaviour can be observed for the imaginary part, although at a smaller magnitude. The complete empirical expression for g_m, which takes this dispersion into account, becomes

$$g_\mathrm{m} = g_\mathrm{m,i}\left[1+\gamma_1 \ln(\omega/\omega_0) + \mathrm{j}\alpha\left(1+\gamma_2\ln(\omega/\omega_0)\right)\right] - \mathrm{j}\omega\left(C_\mathrm{m} - C_\mathrm{m,\omega}\ln(\omega/\omega_0)\right), \tag{27}$$

where $g_\mathrm{m,i}$ is the intrinsic transconductance without the effect of traps, γ_1 and γ_2 determine the strength of the frequency dispersion for the real and logarithmic part, respectively, α scales the imaginary part, since its magnitude is much smaller than the real part, C_m is the mutual differential capacitance as before, and $C_\mathrm{m,\omega}$ anticipates the logarithmic frequency dispersion in the capacitances, which will be discussed forthwith. The intrinsic $g_\mathrm{m,i}$ is reached, once no traps can respond anymore, i.e. above ω_0, and ω is thus the inverse of τ_0 in (26), i.e. the (angular) frequency, above no traps can respond anymore. Physically, a respective frequency dispersion should apply for g_ds as well. However, in all measurements, this was overshadowed by II and BTBT, which is discussed in Section 5.2 [12].

As anticipated, the numerical solution of the total impedance of the distributed RC network yields a logarithmic frequency dependence in the capacitances. Physically, this reflects that from ω_0 on downwards in frequency, more and more traps can interact with carriers in the channel. Since the capture and emission of charge as the consequence of a voltage variation constitutes a capacitance, this dispersion has to be added to the capacitances in the small-signal model. In any case this affects the intrinsic capacitances, but if the parasitic capacitances can be identified to include oxides, which potentially contain defects, the parasitic capacitances should include a similar frequency dispersion as the intrinsic capacitances. Fig. 12(b) demonstrates that this logarithmic frequency dispersion indeed reflects the measured behaviour of the capacitances.

Besides the logarithmic dispersion in the capacitances, the numerical solution of the distributed RC network yields a linear frequency dispersion in the real part of its impedance. This conductance finds its way into the small-signal model alongside the dispersion in the capacitance, as depicted in Fig. 11. The same discussion for intrinsic and parasitic elements applies as for the capacitances. The physical meaning of the conductances related to traps is a phase shift between the applied AC signals at the ports and the response of the traps. Due to their time constants, their interaction is not immediate with respect to the applied signal, which constitutes a relative phase shift and thus loss in the transistor.

Since the distributed RC network in total structurally replaces the gate oxide, the resulting model equations have to replace the intrinsic capacitances of the small-signal model. In the basic small signal

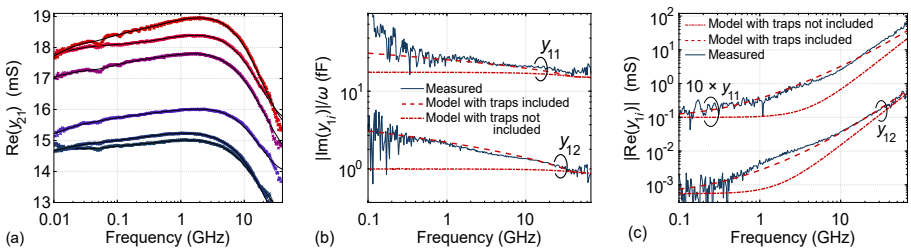


Figure 12: (a) Logarithmic frequency dispersion in the transconductance for different devices. (b) $\mathrm{Im}(y_{11})/\omega$ and $\mathrm{Im}(y_{12})/\omega$ measured and modelled without and with oxide traps. The model without the inclusion of traps results in a straight line, which is clearly not in agreement with the measured values. (c) $10\times\mathrm{Re}(y_{11})$ (to avoid overlap) and $\mathrm{Re}(y_{12})$ measured and modelled without and with oxide traps. (b) and (c) reproduced with permission from [12].

model this means first that the intrinsic $C_{\mathrm{gs,i}}$ is replaced by the equivalent admittance depicted in Fig. 11. As prefigured in Section 4, we will now also take into account the separation of the intrinsic and parasitic C_{gd} and the corresponding NQS resistance R_{j}. For both the gate-to-source and the gate-to-drain coupling, the total intrinsic admittances y_{gx}, with x either s or d for source or drain, respectively, in series with the corresponding intrinsic resistances, are now described by

$$y_{\mathrm{gx}} = \omega \times g_{\mathrm{gx},\omega} + \mathrm{j}\omega \left[C_{\mathrm{gx,i}} - C_{\mathrm{gx},\omega} \ln(\omega/\omega_0)\right]. \tag{28}$$

Here, $g_{\mathrm{gx},\omega}$ and $C_{\mathrm{gx},\omega}$ describe the strength of the frequency dispersion in the real and the imaginary part, respectively, $C_{\mathrm{gx,i}}$ is the intrinsic capacitance without the effect of traps, and all other parameters are as before. The same expression should apply for C_{sd} as well and indeed, frequency dispersions can sometimes be discerned in the measurements. As discussed before, however, the large g_{ds} often masks these effects and makes it difficult to deconvolute the intrinsic NQS resistance so that a combination of the simple model and (28) can be a satisfactory, if not completely physical, description. As mentioned, if the parasitic capacitances are likely to be affected by traps as well, they should be modelled by a corresponding equation. On top of (28), the parasitic expression can take into account DC gate leakage $g_{\mathrm{gx,l}}$ as

$$y_{\mathrm{gx,p}} = \omega \times g_{\mathrm{gxp},\omega} + \mathrm{j}\omega \left[C_{\mathrm{gxp},0} - C_{\mathrm{gxp},\omega} \ln(\omega/\omega_0)\right] + g_{\mathrm{gx,l}}. \tag{29}$$

All parameters correspond to the ones in (28), except for the parasitic contribution. While the constant parasitic $C_{\mathrm{gxp},0}$ can be determined from the off-state of the transistor as described in Section 4.2, the $g_{\mathrm{gxp},\omega}$ and $C_{\mathrm{gxp},\omega}$ are energy- and thus voltage-dependent due to their dependence on the trap energy distribution, so that they cannot be subtracted as constant values [12].

Equations (28), (29), and also (27) are defined for $\omega < \omega_0$. Above ω_0, no traps should be able to respond anymore. Physically, this means that above ω_0, the capacitances and the transconductance become constant at their intrinsic values and the linearly frequency-dependent conductive elements decrease to zero. For the capacitances and the transconductance, the transition between the dispersive and the constant region is reasonably smooth due to the logarithmic behaviour of the dispersion. For the conductances, however, the effect of traps increases with increasing frequency, so that some kind of transition function would be required for a physical, i.e. smooth, transition. Furthermore, recall that the approximation of the lumped model is only valid in a certain frequency range. The transition function should thus also take into account the increasing deviation from the actual numerical solution of the distributed RC network. So far, ω_0, where this transition should occur, is lying outside of the measurement range, or the deviation from the numerical solution has not become significant enough yet, so that the transition could not yet be investigated experimentally. For the same reason – that ω_0 is outside of the measurement range – this uncertainty about the transition of the conductive contribution does not yet affect the model.

5.2 EXTENSION OF THE MODEL – IMPACT IONISATION AND BAND-TO-BAND TUNNELLING

These two effects, impact ionisation (II) and band-to-band tunnelling are commonly observed in III-V MOSFETs. Their effect on the small-signal model can be included in the form of two current sources, which depend on v'_{gs}, just as g_{m}, and on v_{dg}. For an increasing v_{dg}, i.e. an increasing electric field between channel and drain, II and BTBT increase so that the (technical) current from drain to source increases. A higher concentration of electrons in the channel increases this effect further, thus the current source controlled by v'_{gs}. Note that in Fig. 13, this second current source is 'upside down'. This is because the increase of carriers in the channels due to v'_{gs} competes with another effect: Electrons undergoing BTBT leave behind holes in the channel, which decrease the channel barrier so that the current increases further. A higher v'_{gs}, however, leads to a more efficient transport of holes out of the channel, which in turn counteracts the hole feedback effect and the source-drain current decreases. Since the two current sources model both II and BTBT together, the source controlled by v'_{gs} can contribute both positively or negatively. It is thus drawn 'upside down' in Fig 13 to illustrate this. The effect of BTBT and II can be discerned in the lower to curves of $\mathrm{Re}(y_{22})$ Fig. 14, where they increase at the lowest frequencies. Since the equations for these two current sources on their own are not of much use, they are only listed in the summary below. Details about II can be found in [13] and about extraction of the parameters for the two current sources in the model here in [14].

5.3 COMPLETE MODEL – SUMMARY

With all the effects discussed above, the complete small-signal model looks like in Fig. 13. All equations for the different components as well as the equations for the revised y-parameters (from [12]) are

summarised here.

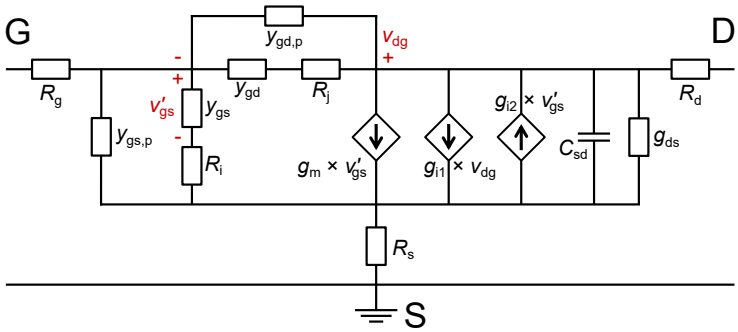


Figure 13: The complete small-signal model with all effects discussed in this overview. Reproduced, with permission, from [12].

The y-parameters are

$$y_{11} = y_{\mathrm{gs,p}} + \frac{y_{\mathrm{gs}}}{1 + y_{\mathrm{gs}} R_{\mathrm{i}}} - y_{12}, \tag{30a}$$

$$y_{12} = -y_{\mathrm{gd,p}} - \frac{y_{\mathrm{gd}}}{1 + y_{\mathrm{gd}} R_{\mathrm{j}}}, \tag{30b}$$

$$y_{21} = \frac{g_{\mathrm{m}}}{1 + y_{\mathrm{gs}} R_{\mathrm{i}}} + y_{12} - g_{\mathrm{i1}} - \frac{g_{\mathrm{i2}}}{1 + y_{\mathrm{gs}} R_{\mathrm{i}}} \tag{30c}$$

$$y_{22} = g_{\mathrm{ds}} + \mathrm{j}\omega C_{\mathrm{sd}} - y_{12} + g_{\mathrm{i1}}. \tag{30d}$$

The transconductance under the influence of traps is

$$g_{\mathrm{m}} = g_{\mathrm{m,i}} \left[1 + \gamma_1 \ln(\omega/\omega_0) + \mathrm{j}\alpha \left(1 + \gamma_2 \ln(\omega/\omega_0)\right)\right] - \mathrm{j}\omega \left(C_{\mathrm{m}} - C_{\mathrm{m},\omega} \ln(\omega/\omega_0)\right). \tag{31}$$

The intrinsic and parasitic admittances, which take into account oxide traps, are

$$y_{\mathrm{gx}} = \omega \times g_{\mathrm{gx},\omega} + \mathrm{j}\omega \left[C_{\mathrm{gx,i}} - C_{\mathrm{gx},\omega} \ln(\omega/\omega_0)\right], \tag{32a}$$

$$y_{\mathrm{gx,p}} = \omega \times g_{\mathrm{gxp},\omega} + \mathrm{j}\omega \left[C_{\mathrm{gxp},0} - C_{\mathrm{gxp},\omega} \ln(\omega/\omega_0)\right] + g_{\mathrm{gx,l}}. \tag{32b}$$

The intrinsic NQS resistances are

$$R_{\mathrm{i}} = \frac{1}{1.4\, g_{\mathrm{m,i}}} \tag{33a}$$

and

$$R_{\mathrm{j}} = \frac{1}{1.4\, g_{\mathrm{m,i}} \times C_{\mathrm{gd,i}}/C_{\mathrm{gs,i}}}, \tag{33b}$$

where R_{j} can be derived in the same manner as R_{i}, while just taking care of the correct capacitances. The expressions for the two current sources for BTBT and II are

$$g_{\mathrm{i1}} = \frac{g10}{1 + \mathrm{j}\omega\tau_{\mathrm{i}}} \tag{34a}$$

and

$$g_{\mathrm{i2}} = \frac{g20}{1 + \mathrm{j}\omega\tau_{\mathrm{i}}}, \tag{34b}$$

i.e. symmetric except for their magnitude g_{10}/g_{20}. The assumption of one common time constant for BTBT and II is an approximation to simplify the model, while still taking these effects into account. Note that τ_{i} is different from $\tau_{\mathrm{c/e}}$ and that the two current sources together model both BTBT and II, not one each.

In Fig. 12 the improvement in the y-parameters due to the inclusion of traps is demonstrated on a logarithmic scale. Fig. 14 demonstrates that the model fits over the whole measured frequency range as well.

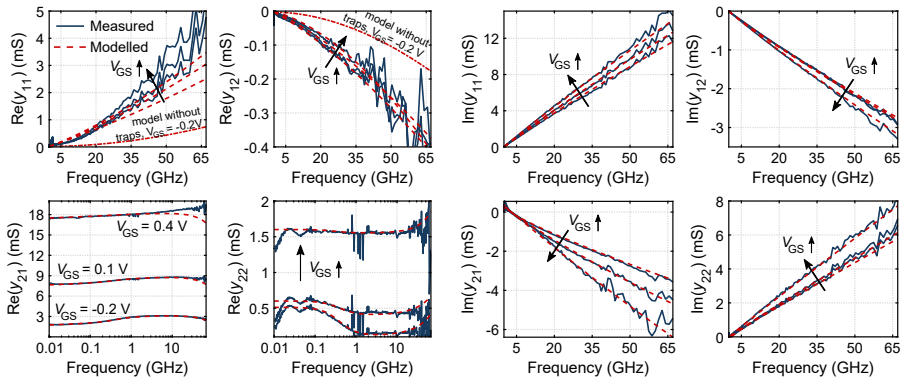


Figure 14: Measured and modelled y-parameters at different gate voltages. Modelled according to Fig. 13 with the respective equations. Only Re(y_{22}) plotted with a logarithmic frequency axis to demonstrate the effect of BTBT and II. Reproduced, with permission, from [12].

5.4 EFFECTS ON GAINS AND STABILITY

With these revised expressions for the y-parameters, as in [12], we can now readily explain the deviations in the measured gains from the ideal roll-offs. Beginning with U, which displays the most evident deviations, we can insert the respective y-parameters from (30) into the definition (15) and by examining, which parameters dominate at intermediate and low frequencies, arrive at

$$U \approx \frac{|g_\mathrm{m}|^2}{4\left[\omega\left(g_{\mathrm{gs},\omega}+g_{\mathrm{gsp},\omega}+g_{\mathrm{gd},\omega}+g_{\mathrm{gdp},\omega}\right)g_\mathrm{ds}+\omega\left(g_{\mathrm{gd},\omega}+g_{\mathrm{gdp},\omega}\right)\mathrm{Re}(g_\mathrm{m})\right]}. \tag{35}$$

In this expression, the linear frequency dispersions from the trap contribution dominates over the usually squared frequency dependence, which changes the roll-off to almost -10 dB/decade instead of -20 dB/decade. This can lead to a reduction of the unilateral gain of up to 10 dB at typical operation frequencies.

The dispersions in h_{21} and MSG in Fig. 10 are more subtle than in the case of U, but their logarithmic increase towards lower frequencies already indicate that this dispersion is related to the logarithmic contribution in the capacitances and indeed, inserting the y-parameters from (30) into the definitions (14) and (23) yields

$$h_{21} = \frac{|g_\mathrm{m}|}{|\omega\left(C_{\mathrm{gs,i}}-C_{\mathrm{gs},\omega}\times\ln(\omega/\omega_0)+C_{\mathrm{gd,i}}-C_{\mathrm{gd},\omega}\times\ln(\omega/\omega_0)\right)|} \tag{36}$$

and

$$\mathrm{MSG} = \frac{|g_\mathrm{m}|}{|\omega\left(C_{\mathrm{gd,i}}-C_{\mathrm{gd},\omega}\times\ln(\omega/\omega_0)\right)|}. \tag{37}$$

Indeed, it can be seen, that now logarithmic frequency dependencies dominate the denominators of these expression and as part of g_m also the numerators. In both cases, the numerator decreases logarithmically towards lower frequencies and the denominator increases, which results in an overall decrease of h_{21} and MSG towards lower frequencies. This matches the measured values.

As the last quantity, in the same manner, we can investigate the stability factor k. Inserting (30) into (25) and simplifying yields

$$k = \frac{2\omega\left(g_{\mathrm{gs},\omega}+g_{\mathrm{gsp},\omega}+g_{\mathrm{gd},\omega}+g_{\mathrm{gdp},\omega}\right)g_\mathrm{ds}+\omega\left(g_{\mathrm{gd},\omega}+g_{\mathrm{gdp},\omega}\right)\mathrm{Re}(g_\mathrm{m})}{|\left[-\omega(g_{\mathrm{gd},\omega}+g_{\mathrm{gdp},\omega})-\mathrm{j}\omega\left(C_{\mathrm{gd,i}}+C_{\mathrm{gd,p}}\right)\right]g_\mathrm{m}|}. \tag{38}$$

Here, the logarithmic frequency dependences of the capacitances were omitted for increased readability, just as the DC leakage contributions $g_{\mathrm{gx,l}}$. Note that the extensive terms in parentheses just contain the different $g_{\mathrm{gx},\omega}$ and $g_{\mathrm{gxp},\omega}$. In (38) it can be seen that from the mode without the effect of traps, k turns out to be lower than when traps are included. Again, this is in agreement with the measured data in Fig. 10.

REFERENCES

[1] W. Liu, *Fundamentals of III-V Devices*. Wiley Interscience, 1999. ISBN 978-0471297000

[2] P. Roblin and H. Rohdin, *High-speed heterostructure devices*. Cambridge University Press, 2002. ISBN 978-0-511-06903-1

[3] E. Lind, "Lecture slides High-Speed Devices," January 2018.

[4] W. Liu, *FET High-Frequency Properties*. Wiley Interscience, 1999, ch. 6, pp. 371–446. ISBN 978-0471297000

[5] P. Roblin and H. Rohdin, *Small- and large-signal AC models for the short-channel MODFET*. Cambridge University Press, 2002, ch. 12, pp. 384–411. ISBN 978-0-511-06903-1

[6] P. Roblin and H. Rohdin, *MODFET high-frequency performance*. Cambridge University Press, 2002, ch. 17, pp. 567–612. ISBN 978-0-511-06903-1

[7] E. Lind, "High frequency III–V nanowire MOSFETs," *Semiconductor Science and Technology*, vol. 31, no. 9, p. 093005, aug 2016. doi: 10.1088/0268-1242/31/9/093005

[8] E. Lind, M. Hellenbrand, Lindelöw, and S. Andric, "Improved small signal model n-type, demonstrated p-type small-signal models," Lund University, Tech. Rep., 2018.

[9] Y. Yuan, L. Wang, B. Yu, B. Shin, J. Ahn, P. C. McIntyre, P. M. Asbeck, M. J. W. Rodwell, and Y. Taur, "A distributed model for border traps in Al_2O_3 – InGaAs MOS devices," *IEEE Electron Device Letters*, vol. 32, no. 4, pp. 485–487, April 2011. doi: 10.1109/LED.2011.2105241

[10] S. Johansson, M. Berg, K. Persson, and E. Lind, "A high-frequency transconductance method for characterization of high-κ border traps in III-V MOSFETs," *IEEE Transactions on Electron Devices*, vol. 60, no. 2, pp. 776–781, Feb 2013. doi: 10.1109/TED.2012.2231867

[11] C. B. Zota, G. Roll, L. Wernersson, and E. Lind, "Radio-frequency characterization of selectively regrown InGaAs lateral nanowire mosfets," *IEEE Transactions on Electron Devices*, vol. 61, no. 12, pp. 4078–4083, Dec 2014. doi: 10.1109/TED.2014.2363732

[12] M. Hellenbrand, E. Lind, O.-P. Kilpi, and L.-E. Wernersson, "Effects of traps in the gate stack on the small-signal RF response of III-V nanowire MOSFETs," *Solid-State Electronics, manuscript under review with minor revisions pending*, May 2020.

[13] M. Isler and K. Schünemann, "Impact-ionization effects on the high-frequency behavior of HFETs," vol. 52, no. 3, pp. 858–863, March 2004. doi: 10.1109/TMTT.2004.823553

[14] S. Johansson, M. Egard, S. G. Ghalamestani, B. M. Borg, M. Berg, L. E. Wernersson, and E. Lind, "RF characterization of vertical InAs nanowire wrap-gate transistors integrated on Si substrates," *IEEE Transactions on Microwave Theory and Techniques*, vol. 59, no. 10, pp. 2733–2738, Oct. doi: 10.1109/TMTT.2011.2163076

PAPERS

Paper I

Paper I

Reproduced, with permission, from

S. Netsu, M. Hellenbrand, C. B. Zota, Y. Miyamoto, and E. Lind, "A Method for Determining Trap Distributions of Specific Channel Surfaces in InGaAs Tri-Gate MOSFETs," *IEEE Journal of the Electron Devices Society*, vol. 6, pp. 408–412, Apr. 2018, doi: 10.1109/JEDS.2018.2806487.

© 2018 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission.

Received 5 December 2017; revised 11 January 2018; accepted 6 February 2018. Date of publication 19 February 2018;
date of current version 9 April 2018. The review of this paper was arranged by Editor M. J. Kumar.

Digital Object Identifier 10.1109/JEDS.2018.2806487

A Method for Determining Trap Distributions of Specific Channel Surfaces in InGaAs Tri-Gate MOSFETs

SEIKO NETSU [1], MARKUS HELLENBRAND [2], CEZAR B. ZOTA [2], YASUYUKI MIYAMOTO [1] (Senior Member, IEEE), AND ERIK LIND [2]

1 Department of Physical Electronics, Tokyo Institute of Technology, Tokyo 152-8550, Japan
2 Department of Electrical and Information Technology, Lund University, 221 00 Lund, Sweden

CORRESPONDING AUTHOR: E. LIND (e-mail: erik.lind@eit.lth.se)

ABSTRACT We present a method for estimating the trap distributions on each of the surfaces in a multi-gate MOSFET. We perform *I–V* hysteresis measurements on InGaAs Tri-gate MOSFETs with various channel widths (25, 60, and 100 nm) from which top surface and side wall trap distributions are determined. We show that the total trap distribution of a device can be expressed as a linear combination of the top surface and side wall trap distributions. The results show that the minimum trap density of the top InGaAs (100) surface is smaller than that of the {110} side walls by almost an order of magnitude. Since the nanowire constituting the channel in these devices is selectively regrown, rather than etched out, the different trap distributions can be explained by the specific surface chemistries of two surfaces.

INDEX TERMS MOSFETs, high-κ, InGaAs, hysteresis, trap density, inter face trap, III-V, multi-gate, FinFETs.

I. INTRODUCTION

III-V compound semiconductors such as InGaAs have been widely investigated as a promising n-type MOSFET channel material for future CMOS technology because of their higher electron mobility [1], [2], which enables higher transconductance g_m and on-current at a given gate length. Although poor quality high-κ gate stacks have been considered as a crucial issue for realizing the use of III-V based MOSFETs in industry, several research groups have recently reported on $In_xGa_{1-x}As$ MOSFETs with g_m of about 3 mS/μm while achieving an acceptable high-κ gate stack quality with respect to the interface trap density [3]–[5]. Further improvement of the performance of III-V-based MOSFETs, the device reliability, for example Positive Bias Temperature Instability (PBTI) is another key research area for III-V based MOSFET [6]. Understanding the relative energies of the trap distribution in the high-κ oxide and charge carriers in the channel is a vital aspect of improving the device reliability and performance. However, in the case of multi-gate MOSFETs such as FinFETs, tri-gates, gate-all-around (GAA) and vertical nanowire (VNW) structures, it is difficult to determine the trap distributions separately for each of the channel surfaces. The trap distributions for each surface type can be assumed to be different due to surface chemistry and process-related roughness. In recent years, a simple trap evaluation method based on the hysteresis of *C-V* or *I-V* characteristics of MOSFETs has been proposed [7]. Even though this method is not suitable for estimating the trap distributions of different surface orientations separately, it is useful in investigating the total MOSFET gate stack properties.

In this paper, we propose a novel method for estimating the trap distributions of tri-gate MOSFET channel top surfaces [$N_{eff,top\ surface}\ (E) = N_{eff\_(100)}\ (E)$] and side walls [$N_{eff,\ side\ walls} = N_{eff\_\{110\}}\ (E)$] based on the *I-V* hysteresis dependence of the ratio between top surface and side wall length of the channel. We investigate the *I-V* hysteresis of devices with channel widths of 25 to 100 nm. From the hysteresis data, the total trap distributions are estimated by fitting the trap distributions of the top (100) surface and the {110} side walls. This method allows characterization of specific channel surfaces in MOSFETs without the need

2168-6734 © 2018 IEEE. Translations and content mining are permitted for academic research only.
Personal use is also permitted, but republication/redistribution requires IEEE permission.
See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

to fabricate several MOSCAPs, and offers a new way of identifying optimal channel surfaces.

II. DUT & HYSTERESIS MEASUREMENT

Fig. 1(a) shows a representative cross-sectional SEM image of nanowires used as the channel in these devices. These nanowires are taller than those used in the devices in order to clearly show the facet angles. The {110} surface planes are determined from the orientation of the nanowire and the 45 sidewall angle. Fig. 1 (b) shows a schematic image and the final dimensions of a nanowire used in the Tri-gate FETs evaluated in this work. The height of the channel is 7 nm, giving 10 nm total length for each of the {110} side walls. The gate length L_g is 150 nm, and the device channel consists of a single nanowire. A 1 + 4 nm /Al_2O_3 / HfO_2 bilayer gate stack (EOT $\approx$ 1.5 nm) was used in these devices; The relationship between the number of ALD cycles, the high-κ oxide thickness and the gate capacitance is known from our past studies based on C-V measurement and TEM image observation of MOSCAPs and MOSFETs [8]–[11]. Further details about device fabrication process and performance are shown in [5] and [12].

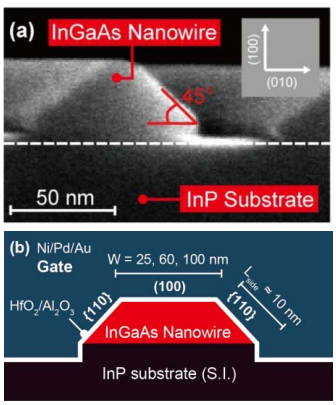


FIGURE 1. (a) Cross-sectional SEM image of the FinFET channel. The surface orientations are (100) and {110} for the top and the side, respectively. (b) Schematic image of the channel structure evaluated in this paper.

We measured *I-V* hysteresis for devices with channel widths of 25, 60 and 100 nm, respectively. The measurements were carried out by sweeping the gate voltage from a starting voltage (V_{start}) to an end voltage (V_{end}), and then back to V_{start}. Each measurement was done on the same device unlike [7]. In order to reduce the effect of residual hysteresis from previous measurement, we started our measurement with small V_{end}, and then gradually increased V_{end}. Furthermore, at least one minute interval was taken until next measurement for the same reason. We decided hold time of each measurement point to 20 ms with each sweep consisting of 2000 points, assuming that trap states can be fully charged. We set $V_{start} = V_{th}$, and V_{th} of the measured devices was about 200 mV. V_{end} varied as a parameter in the measurement. For each V_{end}, the maximum hysteresis voltage was extracted. The effective trapped charge ΔN_{eff} is proportional to the hysteresis voltage $\Delta V_{hysteresis} = q\Delta N_{eff}/C_{ox}$ as described in [13]. Note that the average and standard deviation of V_{th} of each channel width devices, $\overline{V_{th}}$ and σ_{Vth}, are as follows; 1) W = 100 nm device: $\overline{V_{th}} = 200$ mV and $\sigma_{Vth} = 30$ mV. 2) W = 25 nm device: $\overline{V_{th}} = 200.5$ mV and $\sigma_{Vth} = 25$ mV. Based on these small variation of V_{th}, we conclude that we can set V_{start} as a fixed $V_{th} = 200$ mV.

III. METHOD FOR SEPARATION OF TRAP DISTRIBUTIONS FROM I-V HYSTERESIS MEASUREMENT

Next we show how to estimate the trap distribution of the top (100) surface and the {110} side walls from the *I-V* hysteresis data.

First, two Gaussian distributions were assumed for expressing the trap distribution $N_{eff}(E|N_{peak}, \mu_1, \mu_2, \sigma_1, \sigma_2)$ at a surface of the channel in the same manner as in [7] and [14]. The unit of N_{eff} should be [$cm^{-3}eV^{-1}$] and contain a depth information of trap distribution, however, for simplicity, we represented a depth distribution of traps as an equivalent charge sheet representation ([$cm^{-2}eV^{-1}$]) located only at the interface between the semiconductor and the oxide. Since ΔN_{eff} at each V_{end} is already known from the experimental result of $\Delta V_{hysteresis}(V_{ov})$, $N_{eff}(E)$ can be derived by using equation (1) and (2).

$$\Delta N_{eff} = \int_{\varphi_{start}}^{\varphi_{end}} N_{eff}(E)dE \tag{1}$$

$$N_{eff}\left(E\,|N_{peak}, \mu_1, \mu_2, \sigma_1, \sigma_2\right) = N_{peak}\left(\frac{1}{\sigma_1\sqrt{2\pi}}\exp\left(\frac{-(E-\mu_1)^2}{2\sigma_1^2}\right) + \frac{1}{\sigma_2\sqrt{2\pi}}\exp\left(\frac{-(E-\mu_2)^2}{2\sigma_2^2}\right)\right) \tag{2}$$

where N_{peak} is the peak value of the trap distribution, μ its mean value and σ^2 its variance. Since the trap distribution has several sources, such as, e.g., dangling bonds, vacancies and defects in the oxide [14], the two Gaussian distributions differ. N_{eff}, μ and σ^2 are fitting parameters in reproducing the measured $\Delta V_{hysteresis} - V_{ov}$ relationship. φ_{start} and φ_{end} are the surface potentials which correspond to V_{start} and V_{end} respectively. The reference potential $\varphi_{start} = 0$ V was fixed to the threshold voltage. The surface potential for different surface orientation was assumed to be same, based on the fact that although the area ratio of (100) to {110} is different between channel width of 100 nm and 25 nm devices, there was no significant change in the V_{th}.

Second, since equation (1) and (2) can only express the trap distribution at one surface, several extensions are required and the following assumption are made; The total ΔN_{eff} is a linear sum of ΔN_{eff} from different orientations with a coupling based on the ratio of individual channel surface dimensions as shown in equation (3) and (4).

This assumption is based on the idea that the total hysteresis reflects the sum of the charges trapped on different orientations.

$$\Delta N_{eff\_100nm} = \int_{\varphi_{start}}^{\varphi_{end}} N_{eff\_100nm}(E)dE = \alpha \int_{\varphi_{start}}^{\varphi_{end}} N_{eff\_(100)}(E)dE + \beta \int_{\varphi_{start}}^{\varphi_{end}} N_{eff\_\{110\}}(E)dE \quad (3)$$

$$\Delta N_{eff\_25nm} = \int_{\varphi_{start}}^{\varphi_{end}} N_{eff\_25nm}(E)dE = \gamma \int_{\varphi_{start}}^{\varphi_{end}} N_{eff\_(100)}(E)dE + \delta \int_{\varphi_{start}}^{\varphi_{end}} N_{eff\_\{110\}}(E)dE \quad (4)$$

where alpha, beta, gamma, and delta are the ratio of each facet length to the total length of the device channel ($\alpha = 0.83$, $\beta = 0.17$, $\gamma = 0.56$, $\delta = 0.44$ in this study). Since the integration range of surface potential is the same, equation (3) and (4) can be written as follows from the linearity of the integral.

$$N_{eff\_100nm}(E) = \alpha N_{eff\_(100)}(E) + \beta N_{eff\_\{110\}}(E) \quad (5)$$

$$N_{eff\_25nm}(E) = \gamma N_{eff\_(100)}(E) + \delta N_{eff\_\{110\}}(E) \quad (6)$$

Based on equation (3)-(6), $\Delta N_{eff\_W=25\ nm}$ and $\Delta N_{eff\_W=100\ nm}$ can be reproduced by fitting parameters of $N_{eff\_(100)}(E)$ and $N_{eff\_\{110\}}(E)$.

In the simulation, we modeled the gate stack capacitance as the series combination of the oxide capacitance C_{ox} and the quantum capacitance C_{qu} [12]. In order to consider the effects of traps, a trap capacitance C_{Neff} was connected to C_{qu} in parallel. The surface potential and defect charge was subsequently calculated. The way of mapping V_{end} to the surface potential φ_{end} is as follows; V_{end} is given by the sum of the surface potential φ_{end} and the voltage drop across the gate oxide V_{ox} as shown in equation (7). V_{ox} can be calculated by using charge in C_{qu} and $C_{Neff}(=\Delta N_{eff}/\varphi_{end})$ as shown in equation (8). Note that a change of quantized energy level due to band bending was neglected in this calculation.

$$V_{end} = \varphi_{end} + V_{ox} \quad (7)$$

$$V_{ox} = \frac{\Sigma\, Q}{C_{ox}} = \frac{Q_{qu} + \Delta N_{eff}}{C_{ox}} \quad (8)$$

From the distributions $N_{eff\_(100)}(E)$ and $N_{eff\_\{110\}}(E)$ extracted from 25 nm and 100 nm devices, we then predicted the distribution $N_{eff,W=60nm}(E)$ for 60 nm devices. Like the distributions $N_{eff\_(100)}(E)$ and $N_{eff\_\{110\}}(E)$, $N_{eff,W=60nm}(E)$ can be expressed by equation (9).

$$N_{eff\_60nm}(E) = \varepsilon N_{eff\_(100)}(E) + \zeta N_{eff\_\{110\}}(E) \quad (9)$$

where epsilon and zeta are the ratio of each facet length to the total length of the device channel ($\varepsilon = 0.55$, $\zeta = 0.45$). Comparing these predictions with the measured results of such a device provides a means to verify the extracted distributions for different surface orientations.

TABLE 1. Parameters of trap distributions.

	Channel width = 100 nm	Channel width = 25 nm
Top surface (100)	$\mu_1 = 0.98$, $\sigma_1 = 0.17$ $\mu_2 = -1.00$, $\sigma_2 = 0.50$ $N_{peak} = 9.4\times10^{13}$	$\mu_1 = 1.00$, $\sigma_1 = 0.13$ $\mu_2 = -1.00$, $\sigma_2 = 0.58$ $N_{peak} = 2.4\times10^{13}$
Side Wall {110}	$\mu_1 = 0.70$, $\sigma_1 = 0.16$ $\mu_2 = -0.86$, $\sigma_2 = 0.56$ $N_{peak} = 1.2\times10^{13}$	$\mu_1 = 0.62$, $\sigma_1 = 0.19$ $\mu_2 = -0.80$, $\sigma_2 = 0.52$ $N_{peak} = 1.6\times10^{13}$

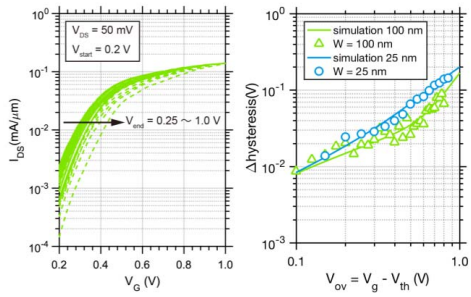


FIGURE 2. (a) Hysteresis characteristics of W = 100 nm FinFET (b) Measured hysteresis data for different channel widths (W = 25 nm and 100nm) against simulation data (solid line).

IV. RESULTS AND DISCUSSION

The *I-V* measurement of a 100 nm channel width tri-gate FET for different V_{end} is shown Fig. 2 (a). The *I-V* hysteresis increases as V_{end} increases, indicating an increase also in the number of trapped charges. Fig. 2 (b) displays the dependence of V_{end} on $\Delta V_{hysteresis}$ for channel widths of 25 and 100 nm. Solid lines were derived from best fits to the experimental data. The parameters used in this fitting are shown in Table 1.

As can be seen, the 25 nm device shows larger hysteresis than the 100 nm device over the entire range, suggesting the existence of a larger number of traps in the former. This was particularly noticeable in the ranges of higher overdrive voltages ($V_{ov}\sim$ 0.7) where the Fermi-level crosses the steepest part of the trap distribution.

Fig. 3 (a) shows the trap distributions for 25 nm and 100 nm channel width devices, obtained by the fitting shown in Fig. 2 (b). As expected from the hysteresis date of Fig. 2 (b), 25 nm channel width gate stack contains higher trap densities. Fig. 3 (a) also includes the trap distributions of the top surface and the side wall. The {110} side wall trap distribution is almost one order of magnitude higher than for the (100) top surface at 0.25 eV. This tendency agreed with the previous results from high-κ/InAs gate stacks [10]. Since the channel width of 25 nm device has more area of the side walls than the top surface relatively in comparison with that of 100 nm device, the side wall properties have a great influence on its total properties, causing a higher trap density as shown in Fig. 3 (a).

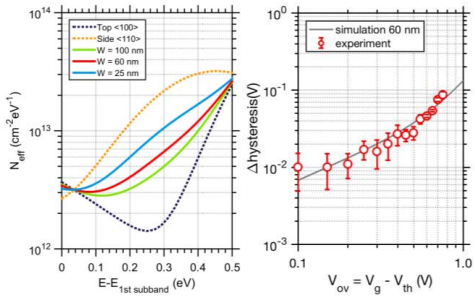


FIGURE 3. (a) Fitted trap distributions (W = 25 nm and 100 nm, solid line), separated trap distributions (top-well and side-well, dashed line) and predicted trap distribution (W = 60 nm) (b) Experimental hysteresis data plot of W = 60 nm device. The error bar indicates the standard deviation of the measurements of six devices. Solid line was predicted by using obtained top surface and side wall trap distributions.

The predicted hysteresis (solid line) of a 60 nm device and its measured hysteresis are shown in Fig. 3 (b). Predicted hysteresis agrees well with measured data, giving validity to our assumption that the hysteresis is explained by a linear combination of the trap distributions of the top and side wall surfaces. Note that in the simulation we fitted equations (5) and (6) separately, thus two sets of $N_{eff\_(100)}(E)$ and $N_{eff\_\{110\}}(E)$ were obtained. In reality, there should be only one set of $N_{eff\_(100)}(E)$ and $N_{eff\_\{110\}}(E)$ in the gate stack. Therefore, we took the average of them when $N_{eff,W=60\ nm}$ was reproduced.

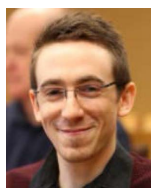
V. CONCLUSION

We have demonstrated a method for the evaluation of the trap energy distribution on different channel surfaces of InGaAs tri-gate MOSFETs. This method is based on a linear combination of the individual channel surface properties. Experimental observations confirmed the validity of this model. We derived the trap distributions of channel widths of 25, 60, 100 nm devices, respectively, from their I-V hysteresis and successfully modelled different distributions for different surface orientations. We also showed that the maximum difference between the trap density of the (100) surface and that of the {110} surface was approximately an order of magnitude, which agrees well with past studies [10]. Although the {110} surface showed 10 times larger trap density as compared to the (100) surface in this study, proposed method could be contributed to find an optimized surface treatment condition and/or selective regrowth condition which minimizes the trap density of the {110} surface in an actual multi-gate device.

REFERENCES

[1] J. A. del Alamo, "Nanometre-scale electronics with III-V compound semiconductors," *Nature*, vol. 479, no. 7373, pp. 317–323, Nov. 2011.
[2] S. Oktyabrsky and P. D. Ye, *Fundamentals of III-V Semiconductor MOSFETs*. New York, NY, USA: Springer, 2010.
[3] J. Lin, X. Cai, Y. Wu, D. A. Antoniadis, and J. A. Alamo, "Record maximum transconductance of 3.45 mS/μm for III-V FETs," *IEEE Electron Device Lett.*, vol. 37, no. 4, pp. 381–384, Apr. 2016.
[4] C. B. Zota, L.-E. Wernersson, and E. Lind, "Single suspended InGaAs nanowire MOSFETs," in *Int. Electron Devices Meeting Tech. Dig. (IEDM)*, Washington, DC, USA, Feb. 2016, pp. 31.4.1–31.4.4.
[5] C. B. Zota, L.-E. Wernersson, and E. Lind, "High-performance lateral nanowire InGaAs MOSFETs with improved on-current," *IEEE Electron Device Lett.*, vol. 37, no. 10, pp. 1264–1267, Oct. 2016.
[6] J. Franco *et al.*, "Demonstration of an InGaAs gate stack with sufficient PBTI reliability by thermal budget optimization, nitridation, high-k material choice, and interface dipole," in *Proc. Symp. VLSI Technol.*, Honolulu, HI, USA, 2016, pp. 1–2.
[7] A. Vais *et al.*, "Impact of starting measurement voltage relative to flatband voltage position on the capacitance-voltage hysteresis and on the defect characterization of InGaAs/high-k metal-oxide-semiconductor stacks," *Appl. Phys. Lett.*, vol. 107, no. 22, pp. 1–5, 2015.
[8] G. Roll, J. Mo, E. Lind, S. Johansson, and L.-E. Wernersson, "Defect evaluation in InGaAs field effect transistors with HfO2 or Al2O3 dielectric," *Appl. Phys. Lett.*, vol. 106, no. 20, 2015, Art. no. 203503.
[9] C. B. Zota, G. Roll, L.-E. Wernersson, and E. Lind, "Radio-frequency characterization of selectively regrown ingaas lateral nanowire MOSFETs," *IEEE Trans. Electron Devices*, vol. 61, no. 12, pp. 4078–4083, Dec. 2014.
[10] C. H. Wang *et al.*, "High-k dielectrics on (100) and (110) n-InAs: Physical and electrical characterizations," *AIP Adv.*, vol. 4, no. 4, 2014, Art. no. 047108.
[11] M. Egard *et al.*, "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, Mar. 2012.
[12] C. B. Zota, D. Lindgren, L.-E. Wernersson, and E. Lind, "Quantized conduction and high mobility in selectively grown $In_xGa_{1-x}As$ nanowires," *ACS Nano*, vol. 9, no. 10, pp. 9892–9897, 2015.
[13] J. Lin *et al.*, "An investigation of capacitance-voltage hysteresis in metal/high-k/In 0.53Ga0.47As metal-oxide-semiconductor capacitors," *J. Appl. Phys.*, vol. 114, no. 14, 2013, Art. no. 144105.
[14] C.-W. Cheng, G. Apostolopoulos, and E. A. Fitzgerald, "The effect of interface processing on the distribution of interfacial defect states and the C-V characteristics of III-V metal-oxide-semiconductor field effect transistors," *J. Appl. Phys.*, vol. 109, no. 2, 2011, Art. no. 023714.

SEIKO NETSU received the B.E. degree in electrical and electronics engineering from the Tokyo Institute of Technology, Tokyo, Japan, in 2015, where he is currently pursuing the M.E. degree with the Department of Physical Electronics.



MARKUS HELLENBRAND received the M.S. degree in engineering nanoscience from Lund University, Lund, Sweden, in 2015, where he is currently pursuing the Ph.D. degree with the Department of Electrical and Information Technology.

CEZAR B. ZOTA received the Ph.D. degree in electrical engineering from Lund University, Lund, Sweden, in 2017. He is currently a Post-Doctoral Researcher with IBM Research Zurich, Switzerland.

YASUYUKI MIYAMOTO (M'88–SM'08) received the D.Eng. degree in physical electronics from the Tokyo Institute of Technology Tokyo, Tokyo, Japan, in 1988.

He is currently a Professor with the Department of Electrical and Electronics Engineering, Tokyo Institute of Technology.

ERIK LIND received the Ph.D. degree in solid-state physics from Lund University, Lund, Sweden, in 2004, where he is currently an Associate Professor with the Department of Electrical and Information Technology.

Paper I © 2018 IEEE

Paper II

Paper II

Reproduced, with permission, from

M. Hellenbrand, O.-P. Kilpi, J. Svensson, E. Lind, and L.-E. Wernersson, "Low-frequency noise in nanowire and planar III-V MOSFETs," *Microelectronic Engineering*, vol. 215, pp. 110986, May 2019, doi: 10.1016/j.mee.2019.110986.

Microelectronic Engineering 215 (2019) 110986

Contents lists available at ScienceDirect

Microelectronic Engineering

journal homepage: www.elsevier.com/locate/mee

Research paper

Low-frequency noise in nanowire and planar III-V MOSFETs

Markus Hellenbrand*, Olli-Pekka Kilpi, Johannes Svensson, Erik Lind, Lars-Erik Wernersson

Department of Electrical and Information Technology, Lund University, Box 118, 221 00 Lund, Sweden

ARTICLE INFO

Keywords:
III-V
Nanowire (NW)
MOSFET
Low-frequency noise
Gate oxide defects
Border traps

ABSTRACT

Nanowire geometries are leading contenders for future low-power transistor design. In this study, low-frequency noise is measured and evaluated in highly scaled III-V nanowire metal-oxide-semiconductor field-effect transistors (MOSFETs) and in planar III-V MOSFETs to investigate to what extent the device geometry affects the noise performance. Number fluctuations are identified as the dominant noise mechanism in both architectures. In order to perform a thorough comparison of the two architectures, a discussion of the underlying noise model is included. We find that the noise performance of the MOSFETs in a nanowire architecture is at least comparable to the planar devices. The input-referred voltage noise in the nanowire devices is superior by at least a factor of four.

1. Introdution

The continuous improvement of the metal-oxide-semiconductor field-effect transistor (MOSFET) has been one of the major driving forces in electronics. In the future, further improvements of the MOSFET will require more innovative measures than the so far tremendously successful conventional geometrical scaling. Common suggestions include gate-all-around architectures and novel materials at the device level [4,5], and monolithic 3D integration at the system level [1–4]. III-V nanowire (NW) MOSFETs, especially in a vertical architecture, can facilitate all these approaches at once and they have already demonstrated competitive or even superior performance in comparison with planar MOSFETs or FinFETs both in simulation [6] and in experiment [7–10].

In this paper we compare planar and NW III-V MOSFETs with similar channel and gate-oxide materials and investigate to what extent the change in architecture affects the low-frequency noise (LFN) of the devices, a property which is important as a technology quality metric at the device level, and for analog circuit performance at the system level [11]. Although LFN has been studied in both planar [12] and in NW [13] III-V MOSFETs, the two architectures have not yet been compared directly. The first part of this article describes the measured devices and the measurement setup. The comparison of the two MOSFET architectures comprises two metrics: the gate oxide defect density N_{bt} and the input-referred gate voltage noise power S_{V_G}. In order to facilitate a detailed comparison of N_{bt}, its calculation is accompanied by a detailed discussion of the LFN model.

2. Devices and measurement setup

The following device descriptions focus on the channel and the gate stack as the two parts, which are most important for LFN. References are provided for further processing details.

For the vertical nanowire (vNW) devices, single nanowires were grown by metal-organic vapor-phase epitaxy on an n^+-InAs buffer layer integrated on Si. The nanowires were graded from intrinsic InAs on the source side to n^+-InGaAs on the drain side. During the InGaAs growth, an n^+-shell was formed around the InAs source to reduce the access resistance. After growth, the nanowires were covered with a top metal and a bottom spacer, so that the channel area could be thinned down by digital etching (alternating surface oxidation and oxide etching) to form a recess gate. After digital etching, the 1 nm Al_2O_3/4 nm HfO_2 bilayer gate oxide (equivalent oxide thickness (EOT) ≈ 1.5 nm) and the 60 nm W gate metal were applied by atomic layer deposition (ALD) and by sputtering, respectively. A schematic of the device structure is provided in Fig. 1(a) and details about the processing can be found in [14]. Fig. 1(b) provides transfer curves representative of the vNW MOSFETs along with the gate lengths and the channel widths of the different measured devices. For the vNW MOSFETs, the channel width corresponds to the nanowire circumference.

For the planar reference devices, the channel consisted of not intentionally doped, 10-nm-thick InGaAs, which was grown by molecular beam epitaxy. In order to investigate differences between different gate oxides, three different ALD high-κ gate oxides were used in the planar devices: 5 nm Al_2O_3, 4 nm HfO_2, and 6.5 nm HfO_2. Evaporated Ti/Pd/Au was used as the gate metal. A schematic of the planar device

* Corresponding author.
E-mail address: markus.hellenbrand@eit.lth.se (M. Hellenbrand).

https://doi.org/10.1016/j.mee.2019.110986
Received 15 April 2019; Received in revised form 6 May 2019; Accepted 13 May 2019
Available online 18 May 2019
0167-9317/ © 2019 Elsevier B.V. All rights reserved.

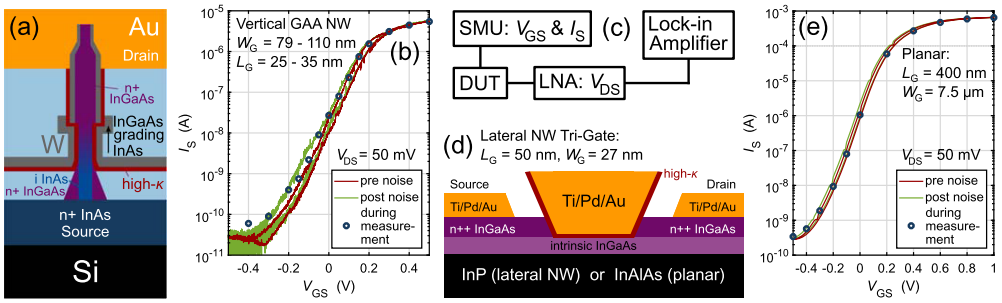


Fig. 1. Schematics of the measured structures (a) & (d), the measurement setup (c), and representative transfer curves for the vertical (b) and the planar devices (e). Similar materials were used for the different devices and all devices exhibit good gate control with typically at least four orders of magnitude current modulation.

structure is provided in Fig. 1(d) and details about the processing can be found in [15]. Fig. 1(e) provides a transfer curve representative of the planar MOSFETs along with the gate lengths and widths of the measured devices.

As an intermediate reference between the planar and the vNW MOSFETs, lateral NW MOSFETs from [16] are included in the analysis as well. Their structure was similar to that of the planar devices, although with a single lateral nanowire as the channel, formed by selective area growth, and with a gate oxide similar to that of the vNW devices. Prior to high-κ application, the nanowires were digitally etched, then sulfur-passivated and annealed in an N_2 atmosphere. Details about processing of the lateral NW devices can be found in [17] and their dimensions are included in Fig. 1(d).

Fig. 1(c) depicts the LFN measurement setup. The drain current of the device under test was measured by a low-noise amplifier (LNA), which also supplied the drain bias $V_{DS} = 50$ mV. The current noise power spectral density S_{I_D} was measured by a lock-in amplifier, which recorded the signal from the LNA. Gate and source bias were supplied by a Source/Measure Unit, which also measured the corresponding currents. For all measured devices, the gate current was at least one order of magnitude lower than the lowest source-drain current and in some cases it was below the measurement noise floor.

3. Results and analysis

LFN in MOSFETs is typically explained as either number fluctuations, caused by gate oxide defects, or as mobility fluctuations, often described with the Hooge model [18]. The two mechanisms can be identified by the dependence of the measured S_{I_D} on the device current I_S. In the case of mobility fluctuations, S_{I_D}/I_S^2 is expected to be proportional to $1/I_S$, which was not observed for any of the measured devices. Instead, the measured S_{I_D} was roughly proportional to the transconductance squared (Fig. 2(b) and (c)), which identified number fluctuations as the dominant LFN mechanism. When measured as a function of the frequency f, LFN typically exhibits a $1/f^\gamma$ shape with $\gamma \approx 1$, where, in the case of number fluctuations, the value of γ depends on the spatial distribution of gate oxide defects. The $1/f^\gamma$ dependence was verified in Fig. 2(a), which presents a measurement for a vNW MOSFET, which is representative for the lateral NW and the planar devices as well.

In the case of number fluctuations, electrons from the MOSFET channel tunnel into and out of gate oxide defects with a certain time constant that determines the corresponding frequency component of S_{I_D}. The changing charge states in the oxide cause fluctuations in the flatband voltage and in carrier scattering, which then cause fluctuations in the device current. The two effects are called number fluctuations and correlated mobility fluctuations (CMF), respectively, and the relation between S_{I_D} and the gate oxide defect density N_{bt} can be derived as [19]

$$S_{I_D} = \frac{q^2 k_B T \lambda N_{bt}}{f^\gamma L_G W_G C_g^2}\left(1 + \frac{\alpha \mu_{eff} C_g I_S}{g_m}\right)^2 g_m^2 = S_{V_{fb}}\left(1 + \frac{\alpha \mu_{eff} C_g I_S}{g_m}\right)^2 g_m^2, \quad (1)$$

where g_m is the transconductance, q the elemental charge, k_B the Boltzmann constant, T the temperature, λ the tunneling attenuation length (calculated by the WKB approximation), L_G and W_G are the gate length and width, respectively, and C_g is the oxide capacitance in series with the quantum capacitance, which has to be taken into account for III-V MOSFETs. The prefactor including N_{bt} in the left term is the detailed expression for the power spectral density of the flatband voltage noise $S_{V_{fb}}$. For the term in parentheses, I_S is the device current, μ_{eff} the effective carrier mobility, and α is the scattering parameter, which describes the change in mobility due to a change of charge in the gate oxide. To take into account screening due to charges in the channel, α decreases with increasing current. This can be expressed empirically as $\alpha = \alpha_0 - \alpha_1 \ln(N_q)$, where α_0 and α_1 are fitting parameters and N_q is the number of charges in the channel [20]. Here, instead of calculating N_q in detail, we express it as proportional to the device current I_S, which is an acceptable approximation, since $\ln(N_q)$ is multiplied by the fitting parameter α_1.

The derivation of (1), which relies on the assumption of elastic tunneling as the trapping/de-trapping mechanism, calls for a short side note. Evidence, in particular from reliability measurements [21], but also from some LFN measurements [22,23], point towards inelastic rather than elastic tunneling as the charge exchange mechanism. However, since tunneling is still part of the inelastic mechanism, the assumption of elastic tunneling constitutes a lower limit of the inelastic model and it is as such that it was used in this study. This lower limit assumption can affect the values for N_{bt}, calculated in the following, which thus also constitute a lower limit.

Results from different measurement techniques in literature [24–27] strongly suggest that N_{bt} varies with respect to energy with a minimum in vicinity to the InGaAs conduction band and increasing both above and below. Since different V_{GS} in the LFN measurement probe different energy levels in the gate oxide, N_{bt} is expected to vary with respect to the device current I_S. With two varying fitting parameters, N_{bt} and α, virtually arbitrary results could be obtained by fitting (1) to the measured data. Therefore, instead of choosing a random partitioning of the effects of a distributed N_{bt} and the CMF, we consider the two effects individually by studying the two extreme cases of (i) CMF with a constant N_{bt} and (ii) a varying N_{bt} without the influence of CMF (i.e. $\alpha = 0$).

Case (i) results in reasonable fits for the vertical and the planar devices, as demonstrated by the dark green, solid lines in Fig. 2(b) and

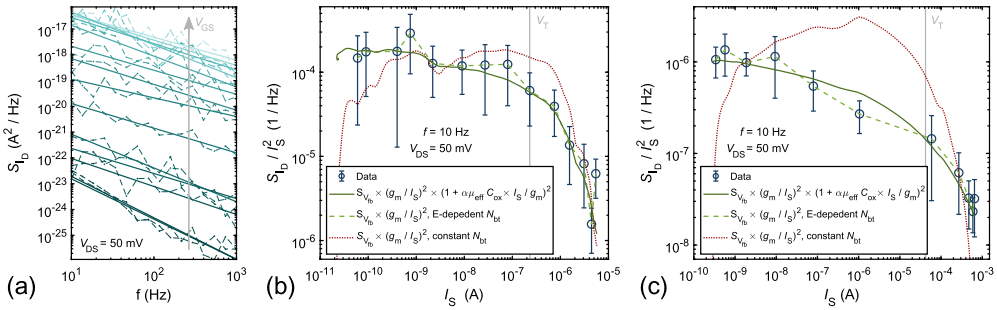


Fig. 2. (a) Noise current power spectral density S_{I_D} as a function of the frequency for a vNW MOSFET. The $1/f^{\gamma}$ dependence with γ close to one is typical of LFN. (b) and (c) S_{I_D}/I_S^2 at a fixed frequency of 10 Hz for a vertical and a planar device, respectively. Circles represent the measured S_{I_D}/I_S^2 with the measurement standard deviation as error bars, lines represent different models with dependences according to the legend entries. Dark green, solid: Number and correlated mobility fluctuations with a constant N_{bt}. Light green, broken: Pure number fluctuations with an energy-dependent N_{bt}. Red, dotted: Reference for the assumption of pure number fluctuations with a constant gate oxide defect density N_{bt}. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

(c). The corresponding constant N_{bt} are summarized in the inset of Fig. 3(a), where lines indicate the average N_{bt} for the respective architecture and shaded boxes indicate the standard deviation of the residuals. For the fits in Fig. 2(a) and (b), values of $\mu_{eff} = 1300\,cm^2/Vs$ [13] and $3500\,cm^2/Vs$ [28] were used for the vertical and the planar devices, respectively, C_g was typically about half the value of the geometrical oxide capacitance, and α typically varied between 2×10^4 and 10^3 Vs/C, which is in agreement with values from literature [11,29,30]. The simplest approach to (ii) is calculating N_{bt} pointwise from the measured S_{I_D} and (1) (with $\alpha = 0$). The resulting N_{bt} are presented in Fig. 3(a) and their shapes resemble the findings in literature. For both (i) and (ii), the values for the NW devices are comparable to those of the planar references and they are comparable to planar Si MOSFETs with HfO_2 gate oxides and a SiO_2 interface layer (EOT < 2 nm, values adapted from [30,31]). In case (ii), the NW devices actually achieve lower values than both of the planar HfO_2 references. For the lateral NW MOSFETs, (i) did not properly model the measured data over the whole measurement range, whereas (ii) resulted in good fits, so that the lateral NW devices are not included in the inset of Fig. 3(a).

Although in most cases, both (i) and (ii) reproduced the measured data reasonably well, neither of the two models can be assumed to be sufficient on their own. (i) does not take into account the distributed nature of N_{bt}, while (ii) yields unphysical values for the highest and lowest V_{GS} (see Fig. 3(a)). Thus, a combination of both models is likely so that the actual N_{bt} resembles the shape of (ii), but with a shallower distribution and slightly lower overall values. Irrespective of the exact contribution of either model, it can be concluded that the mere change from a conventional planar to a NW architecture does not cause a degradation of the gate oxide in terms of LFN performance. In the comparison of the different high-κ gate oxides in the planar architecture, the HfO_2 oxides, especially the thinner ones, tend to exhibit higher values for N_{bt} than the Al_2O_3 oxides, when CMF are disregarded (i). Upon inclusion of CMF (ii), the differences are diminished, although the thin (4 nm) HfO_2 still tends to yield the highest values. These trends are

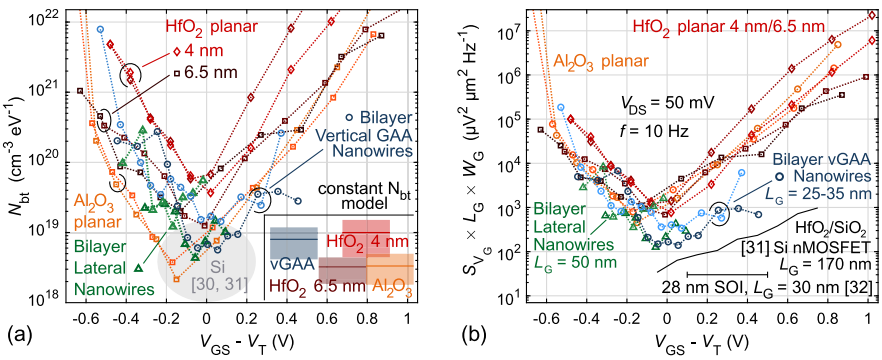


Fig. 3. (a) Gate oxide defect density N_{bt} calculated with two models. (i) Inset: Constant N_{bt}, correlated mobility fluctuations. The lines represent the average values for all measured devices and the shaded boxes the standard deviation of the residuals. (ii) Main: Energy-dependent N_{bt}, no correlated mobility fluctuations. The differences between the two models are larger for the planar devices. Shaded oval as reference for HfO_2 on Si with a thin SiO_2 interface layer from [30,31]. (b) Input-referred voltage noise power spectral density S_{V_G} normalized with the gate area. Black lines as references for [31] as in (a) and commercial SOI [32]. The literature values from [30,31] are shifted with respect to V_{GS} so that they coincide with the measured minima. The metrics in both (a) and (b) reveal that MOSFETs in a (vertical or lateral) nanowire architecture are not degraded when compared with planar MOSFETs. Especially for the input-referred noise S_{V_G}, the nanowire devices actually achieve better values.

most likely related to the ALD growth dynamics, but a more thorough investigation is beyond the scope of this study.

Since the differences in N_{bt} between (i) and (ii) are more prominent for the planar than for the NW devices (especially for HfO_2), it can be concluded that the relative contribution of CMF is stronger in the planar devices. This conclusion is supported by the comparison of the curves for (i) with the dark red, dotted reference curves in Fig. 2(b) and (c), which represent a constant N_{bt} without CMF. For the planar devices, it requires a larger contribution of CMF to merge the two curves than it does for the NW devices. Since the effect of CMF was smaller or even negligible on both the vertical and the lateral NW MOSFETs, the influence of CMF seems to be related to the dimensions of the transistors rather than the lateral or vertical architecture. With large gate areas of L_G = 400 nm and W_G = 7.5 µm in the planar MOSFETs, mobility and scattering play an important role in carrier transport. With much shorter L_G = 25–35 nm (vertical) and 50 nm (lateral), the NW MOSFETs are governed by quasi-ballistic carrier transport [28] and they are approaching the one-dimensional regime, which diminishes the influence of scattering. Furthermore, the larger surface-to-volume ratio in the nanowires is likely to entail a larger ratio of pure number fluctuations over correlated mobility fluctuations.

As a further comparison besides N_{bt}, Fig. 3(b) presents the input-referred gate voltage noise power spectral density $S_{V_G} = S_{I_D}/g_m^2$. This metric does not rely on a specific model, since it merely describes how much of a voltage fluctuation would be required at the gate terminal of a transistor without gate oxide defects to cause the same fluctuations in the device current that were measured for the actual transistor. Thus, S_{V_G} indicates the minimum size of a signal that could still be amplified by the transistor, if it was used e.g. as the input stage of an amplifier. Fig. 3(b) shows that in terms of S_{V_G}, the nanowire devices perform better than the planar references by typically at least a factor of four. For comparison with Si, Fig. 3(b) provides reference values for the same devices as Fig. 3(a) and furthermore for 28 nm silicon-on-insulator (SOI) MOSFETs, the S_{V_G} values of which were calculated from [32] and [33]. The values for the Si and the nanowire devices differ by only a factor three to five. With further optimization of the gate stack in the NW MOSFETs, it should be possible to overcome the remaining difference as well.

4. Conclusions

We measured and compared LFN in nanowire and in planar III-V MOSFETs. Number fluctuations were identified to dominate the LFN in both types of devices and two models were discussed in order to obtain values for the gate oxide defect distribution N_{bt}: number fluctuations with correlated mobility fluctuations for a constant N_{bt} and pure number fluctuations with an energy-dependent distribution of N_{bt}. The results for N_{bt} suggest that a combination of both models is likely, where NW devices are less susceptible to correlated mobility fluctuations. Both models resulted in comparable values for N_{bt} in the NW and the planar architecture, which demonstrates that the change from a planar to a NW architecture does not deteriorate the gate oxide in terms of low-frequency noise. This, together with the lower values for the input-referred gate voltage noise in NW MOSFETs, demonstrates that the performance of the NW MOSFET architecture is not inhibited by low-frequency noise.

Acknowledgements

We thank C. Möhle for the LFN data of the lateral NW MOSFETs, C. B. Zota for fabricating the lateral NW MOSFETs, and J. Mo for fabricating the planar MOSFETs.

This work was supported in part by the Swedish Research Council, in part by the Swedish Foundation for Strategic Research, and in part by the European Union Horizon 2020 Program INSIGHT under grant agreement number 688784.

Declaration of interests

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] H. Riel, L.-E. Wernersson, M. Hong, J.A. del Alamo, III–V compound semiconductor transistors—from planar to nanowire structures, MRS Bull. 39 (8) (2014) 668–677, https://doi.org/10.1557/mrs.2014.137.
[2] N. Loubet, T. Hook, P. Montanini, C.-W. Yeung, S. Kanakasabapathy, M. Guillom, T. Yamashita, J. Zhang, X. Miao, J. Wang, A. Young, R. Chao, M. Kang, Z. Liu, S. Fan, B. Hamieh, S. Sieg, Y. Mignot, W. Xu, S.-C. Seo, J. Yoo, S. Mochizuki, M. Sankarapandian, O. Kwon, A. Carr, A. Greene, Y. Park, J. Frougier, R. Galatage, R. Bao, J. Shearer, R. Conti, H. Song, D. Lee, D. Kong, Y. Xu, A. Arceo, Z. Bi, P. Xu, R. Muthinti, J. Li, R. Wong, D. Brown, P. Oldiges, R. Robison, J. Arnold, N. Felix, S. Skordas, J. Gaudiello, T. Standaert, H. Jagannathan, D. Corliss, M.-H. Na, A. Knorr, T. Wu, D. Gupta, S. Lian, R. Divakaruni, T. Gow, C. Labelle, S. Lee, V. Paruchuri, H. Bu, M. Khare, Stacked nanosheet gate-all-around transistor to enable scaling beyond FinFET, 2017 Symposium on VLSI Technology, 2017, pp. T230–T231, , https://doi.org/10.23919/VLSIT.2017.7998183.
[3] S. Salahuddin, K. Ni, S. Datta, The era of hyper-scaling in electronics, Nature Electronics 1 (8) (2018) 442–450, https://doi.org/10.1038/s41928-018-0117-x.
[4] F. Balestra, M. Graef, B. Huizing, Y. Hayashi, H. Ishiuchi, T. Conte, P. Gargini, The International Roadmap for Devices and Systems, 2017 edition, (2018) https://irds.ieee.org/roadmap-2017 , Accessed date: 12 April 2019.
[5] M.M. Shulaker, G. Hills, R.S. Park, R.T. Howe, K. Saraswat, H.-S.P. Wong, S. Mitra, Three-dimensional integration of nanotechnologies for computing and data storage on a single chip, Nature 547 (2017) 74–78, https://doi.org/10.1038/nature22994.
[6] D. Yakimets, G. Eneman, P. Schuddinck, T.H. Bao, M.G. Bardon, P. Raghavan, A. Veloso, N. Collaert, A. Mercha, D. Verkest, A.V. Thean, K. De Meyer, Vertical GAAFETs for the ultimate CMOS scaling, IEEE Transact. Electron Dev. 62 (5) (2015) 1433–1439, https://doi.org/10.1109/TED.2015.2414924.
[7] O. Kilpi, J. Svensson, E. Lind, L. Wernersson, Electrical properties of vertical InAs/InGaAs heterostructure MOSFETs, IEEE J. Electron Dev. Soc. 7 (2019) 70–75, https://doi.org/10.1109/JEDS.2018.2878659.
[8] S. Johansson, E. Memisevic, L.-E. Wernersson, E. Lind, High-frequency gate-all-around vertical InAs nanowire MOSFETs on Si substrates, IEEE Electron Device Letters 35 (5) (2014) 518–520, https://doi.org/10.1109/LED.2014.2310119.
[9] C.B. Zota, F. Lindelow, L.-E. Wernersson, E. Lind, InGaAs tri-gate MOSFETs with record on-current, 2016 IEEE International Electron Devices Meeting, 2016, https://doi.org/10.1109/IEDM.2016.7838336 pp. 3.2.1–3.2.4.
[10] C.B. Zota, F. Lindelöw, L.-E. Wernersson, E. Lind, High-frequency InGaAs tri-gate MOSFETs with f_{max} of 400 GHz, Electron. Lett. 52 (22) (2016) 1869–1871, https://doi.org/10.1049/el.2016.3108.
[11] M. von Haartman, M. Östling, Low-Frequency Noise in Advanced MOS Devices, Springer, 2007, https://doi.org/10.1007/978-1-4020-5910-0.
[12] T. Karatsori, M. Pastorek, C. Theodorou, A. Fadjie, N. Wichmann, L. Desplanque, X. Wallart, S. Bollaert, C. Dimitriadis, G. Ghibaudo, Static and low frequency noise characterization of ultra-thin body InAs MOSFETs, Solid State Electron. 143 (2018) 56–61, https://doi.org/10.1016/j.sse.2017.12.001.
[13] K.-M. Persson, B.G. Malm, L.-E. Wernersson, Surface and core contribution to 1/f-noise in InAs nanowire metal-oxide-semiconductor field-effect transistors, Appl. Phys. Lett. 103 (3) (2013), https://doi.org/10.1063/1.4813850 033508–1–033508–4.
[14] O.-P. Kilpi, J. Svensson, J. Wu, A.R. Persson, R. Wallenberg, E. Lind, L.-E. Wernersson, Vertical InAs/InGaAs heterostructure metal–oxide–semiconductor field-effect transistors on Si, Nano Lett. 17 (10) (2017) 6006–6010, https://doi.org/10.1021/acs.nanolett.7b02251.
[15] G. Roll, J. Mo, E. Lind, S. Johansson, L.-E. Wernersson, Defect evaluation in InGaAs field effect transistors with HfO_2 or Al_2O_3 dielectric, 106 (20) (2015), https://doi.org/10.1063/1.4921483 203503.
[16] C. Möhle, C. Zota, M. Hellenbrand, E. Lind, 1/f and RTS noise in InGaAs nanowire MOSFETs, Microelectron. Eng. 178 (2017) 52–55 special issue of Insulating Films on Semiconductors (INFOS 2017) https://doi.org/10.1016/j.mee.2017.04.038.
[17] C.B. Zota, L.-E. Wernersson, E. Lind, High-performance lateral nanowire InGaAs MOSFETs with improved on-current, IEEE Electron Device Letters 37 (10) (2016) 1264–1267, https://doi.org/10.1109/LED.2016.2602841.
[18] F. Hooge, 1/f noise is no surface effect, Phys. Lett. A 29 (3) (1969) 139–140, https://doi.org/10.1016/0375-9601(69)90076-0.
[19] G. Ghibaudo, O. Roux, C. Nguyen-Duc, F. Balestra, J. Brini, Improved analysis of low frequency noise in field-effect MOS transistors, Phys. Status Solidi A 124 (2) (1991) 571–581, https://doi.org/10.1002/pssa.2211240225.
[20] A. Pacelli, S. Villa, A.L. Lacaita, L.M. Perron, Quantum effects on the extraction of MOS oxide traps by 1/f noise measurements, IEEE Transact. Electron Dev. 46 (5) (1999) 1029–1035, https://doi.org/10.1109/16.760413.
[21] T. Grasser, Stochastic charge trapping in oxides: from random telegraph noise to bias temperature instabilities, Microelectron. Reliab. 52 (1) (2012) 39–70 2011 Reliability of Compound Semiconductors (ROCS)Workshop https://doi.org/10.1016/j.microrel.2011.09.002.
[22] J.P. Campbell, J. Qin, K.P. Cheung, L.C. Yu, J.S. Suehle, A. Oates, K. Sheng, Random telegraph noise in highly scaled nMOSFETs, 2009 IEEE International Reliability

Physics Symposium, 2009, pp. 382–388, , https://doi.org/10.1109/IRPS.2009.5173283.
[23] T. Nagumo, K. Takeuchi, T. Hase, Y. Hayashi, Statistical characterization of trap position, energy, amplitude and time constants by RTN measurement of multiple individual traps, 2010 International Electron Devices Meeting, 2010, pp. 2831–2834, , https://doi.org/10.1109/IEDM.2010.5703437.
[24] V. Putcha, J. Franco, A. Vais, S. Sioncke, B. Kaczer, Q. Xie, P. Calka, F. Tang, X. Jiang, M. Givens, N. Collaert, D. Linten, G. Groeseneken, BTI reliability of InGaAs nMOS gate-stack: On the impact of shallow and deep defect bands on the operating voltage range of III-V technology, 2017 IEEE International Reliability Physics Symposium (IRPS), 2017, pp. XT–8.1–XT–8.6, , https://doi.org/10.1109/IRPS.2017.7936422.
[25] B. Kaczer, J. Franco, P. Weckx, P. Roussel, V. Putcha, E. Bury, M. Simicic, A. Chasin, D. Linten, B. Parvais, F. Catthoor, G. Rzepa, M. Waltl, T. Grasser, A brief overview of gate oxide defect properties and their relation to MOSFET instabilities and device and circuit time-dependent variability, Microelectron. Reliab. 81 (2018) 186–194, https://doi.org/10.1016/j.microrel.2017.11.022.
[26] G. Brammertz, A. Alian, D.H. Lin, M. Meuris, M. Caymax, W. Wang, A combined interface and border trap model for high-mobility substrate metal–oxide–semiconductor devices applied to $In_{0.53}Ga_{0.47}As$ and InP capacitors, IEEE Transact. Electron Dev. 58 (11) (2011) 3890–3897, https://doi.org/10.1109/TED.2011.2165725.
[27] S. Johansson, M. Berg, K.M. Persson, E. Lind, A high-frequency transconductance method for characterization of high-κ border traps in III-V MOSFETs, 60 (2) (2013) 776–781, https://doi.org/10.1109/TED.2012.2231867.
[28] C.B. Zota, D. Lindgren, L.-E. Wernersson, E. Lind, Quantized conduction and high mobility in selectively grown $In_xGa_{1-x}As$ nanowires, ACS Nano 9 (10) (2015) 9892–9897, https://doi.org/10.1021/acsnano.5b03318.
[29] J. Jomaah, M. Fadlallah, G. Ghibaudo, Low Frequency Noise Analysis in Advanced CMOS Devices, in: Advances in Innovative Materials and Applications, Vol. 324 of Advanced Materials Research, Trans Tech Publications, 2011, pp. 441–444, https://doi.org/10.4028/www.scientific.net/AMR.324.441.
[30] D. Lopez, S. Haendler, C. Leyris, G. Bidal, G. Ghibaudo, Low-frequency noise investigation and noise variability analysis in high-k/metal gate 32-nm CMOS transistors, IEEE Transact. Electron Dev. 58 (8) (2011) 2310–2316, https://doi.org/10.1109/TED.2011.2141139.
[31] E. Simoen, A. Veloso, Y. Higuchi, N. Horiguchi, C. Claeys, On the oxide trap density and profiles of 1-nm EOT metal-gate last CMOS transistors assessed by low-frequency noise, IEEE Transact. Electron Dev. 60 (11) (2013) 3849–3855, https://doi.org/10.1109/TED.2013.2279892.
[32] E.G. Ioannidis, S. Haendler, A. Bajolet, T. Pahron, N. Planes, F. Arnaud, R.A. Bianchi, M. Haond, D. Golanski, J. Rosa, C. Fenouillet-Beranger, P. Perreau, C.A. Dimitriadis, G. Ghibaudo, Low frequency noise variability in high-k/metal gate stack 28nm bulk and FD-SOI CMOS transistors, 2011 International Electron Devices Meeting, 2011, pp. 18.6.1–18.6.4, , https://doi.org/10.1109/IEDM.2011.6131581.
[33] B.K. Esfeh, V. Kilchytska, V. Barral, N. Planes, M. Haond, D. Flandre, J.-P. Raskin, Assessment of 28nm UTBB FD-SOI technology platform for RF applications: figures of merit and effect of parasitic elements, Solid State Electron. 117 (2016) 130–137, https://doi.org/10.1016/j.sse.2015.11.020.

Paper III

Paper III

Reproduced, with permission, from

M. Hellenbrand, E. Memišević, M. Berg, O.-P. Kilpi, J. Svensson, and L.-E. Wernersson, "Low-Frequency Noise in III–V Nanowire TFETs and MOSFETs," *IEEE Electron Device Letters*, vol. 38, no. 11, pp. 1520–1523, Nov. 2017, doi: 10.1109/LED.2017.2757538.

© 2017 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission.

Low-Frequency Noise in III–V Nanowire TFETs and MOSFETs

Markus Hellenbrand, Elvedin Memišević, Martin Berg, Olli-Pekka Kilpi, Johannes Svensson, and Lars-Erik Wernersson

***Abstract*— We present a detailed analysis of low-frequency noise (LFN) measurements in vertical III–V nanowire tunnel field-effect transistors (TFETs), which help to understand the limiting factors of TFET operation. A comparison with LFN in vertical metal-oxide semiconductor field-effect transistors with the same channel material and gate oxide shows that the LFN in these TFETs is dominated by the gate oxide properties, which allowed us to optimize the TFET tunnel junction without deteriorating the noise performance. By carefully selecting the TFET heterostructure materials, we reduced the inverse subthreshold slope well below 60 mV/decade for a constant LFN level.**

***Index Terms*— Vertical nanowires, III–V, MOSFET, TFET, low-frequency noise.**

I. Introduction

LOW-FREQUENCY noise (LFN) is known to have a detrimental effect on both transistor and circuit performance [1]. At the same time, however, LFN can be used to gain insight into the material and transport properties of transistors and may be regarded as a technology quality metric. From a comparison of vertical nanowire metal-oxide semiconductor field-effect transistors (MOSFETs) with vertical nanowire tunnel field-effect transistors (TFETs) of a similar structure and the same channel material and gate oxide, we can deduce that the LFN properties in both device types are dominated by the gate oxide. This is further supported by comparison of TFETs with different tunnel junctions. LFN in TFETs was studied experimentally before in Si-based devices [2], [3], but knowledge about full III-V implementations is still limited [4]. Here, we use InAs MOSFETs as reference for the LFN analysis and show how the same LFN model can be applied to TFETs, even when the inverse subthreshold slope S is reduced below 60 mV/decade. This model and the experimental study of LFN in TFETs will help to better understand the limiting factors of TFET operation.

Manuscript received September 4, 2017; revised September 22, 2017; accepted September 25, 2017. Date of publication September 28, 2017; date of current version October 23, 2017. This work was supported in part by the Swedish Foundation for Strategic Research and the Swedish Research Council, in part by the European Union Seventh Framework Program E2SWITCH under Grant 619509, and in part by the European Union through the H2020 Program INSIGHT under Grant 688784. The review of this letter was arranged by Editor L. Selmi. *(Corresponding author: Markus Hellenbrand.)*

M. Hellenbrand, E. Memišević, O.-P. Kilpi, J. Svensson, and L.-E. Wernersson were with the Department of Electrical and Information Technology, Lund University, SE-221 00 Lund, Sweden (e-mail: markus.hellenbrand@eit.lth.se).

M. Berg was with the Department of Electrical and Information Technology, Lund University, SE-221 00 Lund, Sweden. He is now with Spallacatch AB, SE-211 20 Malmö, Sweden.

Color versions of one or more of the figures in this letter are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/LED.2017.2757538

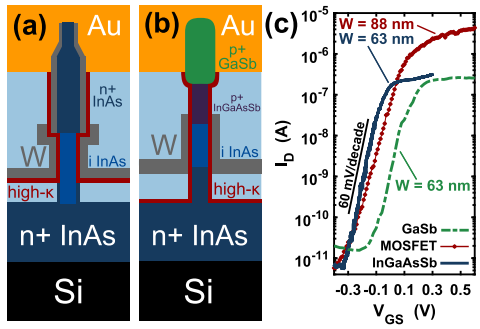


Fig. 1. (a) Schematic illustration of a MOSFET. (b) Schematic illustration of an InGaAsSb TFET. The GaSb structure looks the same, only without the InGaAsSb segment. (c) Transfer characteristics of a MOSFET and both types of TFET at room temperature with V_{DS} = 50 mV. All transistors consisted of a single nanowire. *W* in (c) corresponds to the gate width, which for gate all-around transistors is the circumference of the nanowire.

II. Device Structures and DC Characterization

The processing and final device structure (Fig. 1) for the studied MOSFETs and TFETs are very similar, and especially the resulting gate stack is the same. All nanowires were grown by metalorganic vapor phase epitaxy, using Au seed particles electron-beam-lithography-defined on a highly n-doped InAs buffer layer integrated on Si. For MOSFETs, the nanowires consisted of a 200-nm-long not intentionally doped InAs bottom segment, followed by a 400-nm-long highly n-doped top segment, the growth of which also resulted in a highly doped shell overgrown around the bottom segment. The TFET nanowires consisted of a highly n-doped InAs bottom segment and a not intentionally doped InAs channel segment, followed by a highly p-doped segment to form the tunnel junction. In the first generation, the junction was realized by switching to GaSb (S > 60 mV/decade) and in the second generation to $In_{0.1}Ga_{0.9}As_{0.88}Sb_{0.12}$, which resulted in S below 60 mV/decade. To maintain good top contacts, the InGaAsSb segment was followed by an additional GaSb segment and for the MOSFETs, a W/TiN top contact was applied immediately after growth. For all devices, the InAs channel region was digitally etched to reach final diameters of 28 nm for MOSFETs and 20 nm for TFETs. After etching, identical high-κ gate oxides

0741-3106 © 2017 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission.
See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

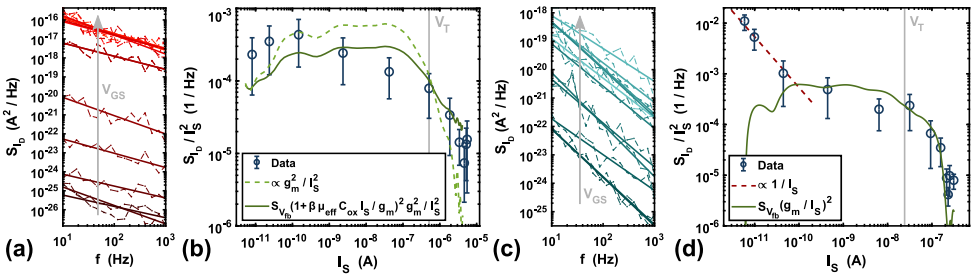


Fig. 2. Representative noise measurements. (a) Frequency sweep for a MOSFET. (b) MOSFET measurement at 10 Hz. The lines indicate that the noise is dominated not only by number fluctuations (broken line), but by both number and correlated mobility fluctuations (solid line). (c) and (d) Same graphs as (a) and (b), but for TFETs. (d) Number fluctuations are dominant over a large current range (solid line). In the off-state, however, mobility fluctuations of electrons injected into the channel by defect-assisted tunneling, take over (broken line). The error bars in (b) and (d) show the standard deviation of the measurement.

(1 nm/4 nm Al_2O_3/HfO_2 atomic layer deposition, effective oxide thickness $\approx$ 1.4 nm) and sputtered W metal gates were applied for all devices. Source, gate, and drain were separated by spacers, details on the processing schemes can be found in [5] for MOSFETs, and in [6] (GaSb) and [7] (InGaAsSb) for TFETs. Fig. 1(c) compares representative transfer characteristics of a MOSFET with both types of TFETs, where these TFET examples exhibit values of $S = 72$ mV/decade (GaSb) and $S = 52$ mV/decade (InGaAsSb). Other InGaAsSb devices on the same sample reached values of S below 50 mV/decade.

III. Measurement and Analysis

Each measurement consisted of a logarithmic frequency sweep between 10 Hz and 1 kHz and a more detailed measurement at 10 Hz, and all measurements were carried out for gate voltages V_{GS} covering the whole range from the off- to the on-state. A lock-in amplifier measured the current noise power spectral density (PSD) of the transistor current for a constant source-drain bias $V_{DS} = 50$ mV, which was supplied by a low-noise current preamplifier that was used to increase the measurement sensitivity. Fig. 2(a) and (b) show measurements of the noise current PSD representative for MOSFETs, Fig. 2(c) and (d) show the same measurements for InGaAsSb TFETs, and the GaSb TFET measuements showed the same behavior as well. With the exception of the noise for the lowest currents in Fig. 2(d), the complete noise behavior of these devices can be described by the so-called McWhorter model [8]. According to this model, electrons tunneling into and out of gate oxide defects cause fluctuations in the channel potential energy, which appears in the drain current as noise. This noise origin is commonly referred to as number fluctuations. For a long channel MOSFET, starting from the PSD of a single gate oxide defect and integrating over a spatial and energetic defect distribution N_{bt} [9] results in the expression

$$\frac{S_{I_D}}{I_S^2} = \frac{q^2 kT\lambda N_{bt}}{f^\gamma L_G W_G C_{ox}^2}\frac{g_m^2}{I_S^2} = S_{V_{fb}}\frac{g_m^2}{I_S^2}, \quad (1)$$

where I_S is the source current, q the elemental charge, k the Boltzmann constant, T the temperature, $\lambda = 0.13$ nm the tunneling attenuation length according to the Wentzel-Kramers-Brillouin approximation, f the frequency, L_G and W_G the effective gate length and width, respectively, $C_{ox} = 0.033$ F/m^2 is the gate area normalized gate oxide capacitance calculated as a cylindrical capacitor, and g_m the transconductance. $S_{V_{fb}}$, summarizing everything except for the transconductance, is called the flatband voltage noise PSD and was a fitting parameter in Fig. 2, whereas measured values were used for I_S and g_m. The frequency exponent γ empirically takes into account spatial nonuniformities in the defect density, where $\gamma = 1$ corresponds to a spatially uniform defect distribution.

Although the derivation of (1) in [9] was carried out for long channel MOSFETs, it holds for TFETs just as well. With the expression

$$I_{1D} = a\,(V_R - V_{th})\exp(-b/\xi) \quad (2)$$

from [10] for an ideal 1D TFET, electrons tunneling into and out of gate oxide defects will change both V_R, the reverse bias applied to the tunnel junction by the gate, and, since it depends on V_R, also ξ, the electric field across the junction. In (2), a and b summarize constants and material parameters, and V_{th} is ln(4) times the thermal voltage kT/q. V_R is related to the gate bias V_{GS} as $V_R = V_{GS}C_{ox}/(C_{ox}+C_q+C_{it})$ with the gate, the semiconductor, and the interface state capacitances C_{ox}, C_q, and C_{it}, respectively. Since all TFETs featured gate lengths of at least 260 nm, V_{DS}, in first order approximation, does not affect V_R anywhere close to the junction. The resulting fluctuation I_N in the current at each frequency (so that $I_N = \sqrt{S_{I_D}}$) due to a charge fluctuation ∂q_{ox} in the oxide can be calculated in a straightforward manner as

$$I_N = \frac{\partial I_{1D}}{\partial q_{ox}}\Delta q_{ox} = \frac{\partial I_{1D}}{\partial V_R}\frac{\partial V_R}{\partial q_{ox}}\Delta q_{ox} = g_m\frac{\partial V_R}{\partial q_{ox}}\Delta q_{ox}, \quad (3)$$

which also holds for a 3D expression of (2). Since the noise both in MOSFETs and in TFETs is caused by the same local flatband fluctuations due to charges in the oxide, it is justified to identify $(\Delta q_{ox}\partial V_R/\partial q_{ox})^2$ in (3) with $S_{V_{fb}}$ in (1), so that the same noise analysis can be applied both for MOSFETs and for TFETs. And indeed, the proportionalities in (1) fit very well with the results of both InGaAsSb TFETs (Fig. 2(d)) and GaSb TFETs (not shown). Furthermore, the application of the same model is supported by the unnormalized noise in Fig. 3(a), since all measurements follow the same trend. For some MOSFETs, as e.g. in Fig. 2(b), measurement and model do not agree as well, which we attribute to so-called correlated mobility fluctuations. This additional

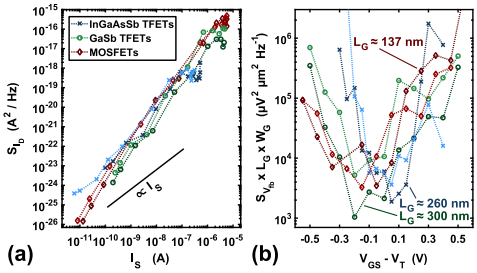


Fig. 3. (a) Measurements at 10 Hz like the ones in Fig. 2(b) and (d), but without I_S normalization. Each curve represents one device, using the same measurement conditions for all. S_{I_D} not being proportional to I_S (cp. (4)) supports Fig. 2 in that it is not mobility fluctuations that are dominant. (b) Equivalent input noise, same transistors and same colors as in Fig. 3(a). Despite the improved subthreshold slope, the InGaAsSb TFETs do not exhibit higher equivalent input noise than the GaSb TFETs or even the MOSFETs.

contribution in the intermediate current region results from increased Coulomb scattering from electrons trapped in gate oxide defects and can be modelled by including a factor $(1 + \beta\mu_{eff}C_{ox}I_S/g_m)^2$ in the right side terms of (1) [9]. Here, the scattering parameter β and the effective channel mobility μ_{eff} together were a fitting parameter, determined by applying a linear fit to $\sqrt{(S_{I_D}/g_m^2)}$ plotted versus I_S/g_m, which also determined $S_{V_{fb}}$ for this model. For Fig. 2(b), assuming $\mu_{eff} \approx 1300$ cm^2/Vs for a nanowire InAs channel [11], this fit resulted in $\beta \approx 2700$ Vs/C, which is reasonably close to values typically found for β [9].

The assumption of elastic tunneling as an important part of the capture and release mechanism was confirmed by carrying out LFN measurements at low temperatures (11 K, not shown), where the results were the same as in Fig. 2, except for a difference in magnitude, which directly corresponds to the explicit temperature dependence in the middle expression of (1). This reveals that the capture and release mechanism is not dominated by any activation energy.

Besides the largely dominant number fluctuations in Fig. 2(b) and (d), a second noise contribution shows up for very small currents as indicated in Fig. 2(d). In the subthreshold region, these so-called mobility fluctuations [9] can be empirically described as

$$\frac{S_{I_D}}{I_S^2} = \frac{\alpha_H \mu_{eff} 2kT}{f L_G^2 I_S}, \tag{4}$$

where α_H is the so-called Hooge parameter, and the other parameters are as defined before. The line in Fig. 3(a) proportional to I_S (cp. (4)) further supports that mobility fluctuations are only relevant for a few devices and for low currents. This separation of mobility and number fluctuations was observed before in nanowire MOSFETs and depends on a surface or core conduction path in the respective operation regime [11]. In the case of the TFETs here, it is known that the off-state was affected by thermionic emission through defect-assisted tunneling [6], [7]. Since the TFET off-current after the junction is governed by drift-diffusion, mobility fluctuations can occur in the TFET off-current as well. With the $1/I_S$ fit in Fig. 2(d) and $\mu_{eff} = 1300$ cm^2/Vs for an InAs channel as before [11], the Hooge parameter can be calculated as $\alpha_H = 3.5 \times 10^{-5}$, which is in very good agreement with [11]. It appears that the region right after the threshold voltage in Fig. 2(b) and (d) and up to saturation of the curves, which probably results from series resistance noise becoming dominant, could be described by (4) as well. However, for MOSFETs, mobility fluctuations are already taken into account as a correlated effect, and for the measured TFETs, the tunneling process of electrons from the source into the channel is not subject to the mobility. Since number fluctuations in contrast directly affect the tunneling window, they dominate the overall noise behavior.

The slopes in Fig. 2(a) and (c) (and thus γ in (1)) differ between MOSFETs and TFETs and γ was generally slightly smaller than one for MOSFETs and slightly larger for TFETs. This small variation around $\gamma \approx 1$ is – besides measurement uncertainties – most likely a cause of processing variations, leading to slightly different defect distributions around the probed depths. For some TFET measurements, however, γ even approaches values of two, which indicates that in those particular devices, there were only very few active defects. As was shown for example in [2] and [12], the LFN behavior in TFETs is dominated by the 10-20 nm of the gate/channel closest to the junction, which explains how only very few defects contribute to the LFN characteristics. In fact, in many of the TFETs studied here, for certain bias conditions we were able to measure Random Telegraph Signal (RTS) noise [7], which is caused by only one individual defect, and agrees well with the observation of γ approaching values of two. The small gate area affecting LFN in TFETs also explains why the levels of the gate area normalized equivalent input gate voltage noise PSD $S_{V_{fb}} \times L_G \times W_G$ for MOSFETs and TFETs in Fig. 3(b) are comparable. While, due to its exponential dependence on the applied bias, the tunnel junction in TFETs is much more sensitive to fluctuations in the energy bands than the thermionic barrier in MOSFETs, the gate area inducing LFN in TFETs is much smaller. This leads to comparable LFN levels. Besides the border trap density N_{bt}, $S_{V_{fb}} \times L_G \times W_G$ only includes constants (cp. (1)), which allows for directly calculating N_{bt} and thus evaluating the gate oxide quality. For all measured devices the minimum N_{bt} is in the order of 2×10^{20} cm^{-3}eV^{-1}, which is consistent with all devices having the same InAs channel and the same gate stack. For the TFETs, this is actually a conservative estimate, since the value was calculated with the whole physical gate length instead of the probably critical 10-20 nm closest to the junction, so that the actual value for N_{bt} in TFETs could be lower by up to a factor of ten.

IV. Conclusion

We analyzed in detail the LFN behavior in vertical III-V nanowire MOSFETs and TFETs and showed that the same noise formalism is applicable to both types of devices despite their different transport mechanisms. From this, we conclude that the gate oxide is the dominant source of LFN in our devices, which is reflected in a constant LFN level even in the case of a significant improvement of the TFET inverse subthreshold slope due to the optimization of the heterojunction.

Acknowledgments

The authors would like to thank J. Min and Prof. P. Asbeck from UC San Diego, Prof. A. Seabaugh from the University of Notre Dame, and Assoc. Prof. E. Lind from Lund University for valuable discussions.

REFERENCES

[1] S. Datta, H. Liu, and V. Narayanan, "Tunnel FET technology: A reliability perspective," *Microelectron. Rel.*, vol. 54, no. 5, pp. 861–874, Mar. 2014. [Online]. Available: http://dx.doi.org/10.1016/j.microrel.2014.02.002

[2] J. Wan, C. Le Royer, A. Zaslavsky, and S. Cristoloveanu, "Low-frequency noise behavior of tunneling field effect transistors," *Appl. Phys. Lett.*, vol. 97, no. 24, pp. 243503-1–243503-3, Nov. 2010, doi: 10.1063/1.3526722.

[3] Q. Huang, R. Huang, C. Chen, C. Wu, J. Wang, C. Wang, and Y. Wang, "Deep insights into low frequency noise behavior of tunnel FETs with source junction engineering," in *Symp. VLSI Technol., Dig. Tech. Papers*, Jun. 2014, pp. 1–2, doi: 10.1109/VLSIT.2014.6894371.

[4] R. Bijesh, D. K. Mohata, H. Liu, and S. Datta, "Flicker noise characterization and analytical modeling of homo and hetero-junction III–V tunnel FETs," in *Proc. 70th Device Res. Conf.*, Jun. 2012, pp. 203–204, doi: 10.1109/DRC.2012.6257032.

[5] M. Berg, O.-P. Kilpi, K.-M. Persson, J. Svensson, M. Hellenbrand, E. Lind, and L.-E. Wernersson, "Electrical characterization and modeling of gate-last vertical InAs nanowire MOSFETs on Si," *IEEE Electron Device Lett.*, vol. 37, no. 8, pp. 966–969, Aug. 2016, doi: 10.1109/LED.2016.2581918.

[6] E. Memišević, J. Svensson, M. Hellenbrand, E. Lind, and L.-E. Wernersson, "Scaling of vertical InAs–GaSb nanowire tunneling field-effect transistors on Si," *IEEE Electron Device Lett.*, vol. 37, no. 5, pp. 549–552, May 2016, doi: 10.1109/LED.2016.2545861.

[7] E. Memišević, M. Hellenbrand, E. Lind, A. R. Persson, S. Sant, A. Schenk, J. Svensson, R. Wallenberg, and L.-E. Wernersson, "Individual defects in InAs/InGaAsSb/GaSb nanowire tunnel field-effect transistors operating below 60 mv/decade," *Nano Lett.*, vol. 17, no. 7, pp. 4373–4380, Jun. 2017, doi: 10.1021/acs.nanolett.7b01455.

[8] E. Burstein, R. H. Kingston, and A. L. McWhorter, *Semiconductor Surface Physics*. Philadelphia, PA, USA: Univ. Pennsylvania Press, 1957.

[9] M. von Haartman and M. Östling, *Low-Frequency Noise in Advanced MOS Devices*. Dordrecht, The Netherlands: Springer, 2007, doi: 10.1007/978-1-4020-5910-0.

[10] A. C. Seabaugh and Q. Zhang, "Low-voltage tunnel transistors for beyond CMOS logic," *Proc. IEEE*, vol. 98, no. 12, pp. 2095–2110, Dec. 2010, doi: 10.1109/JPROC.2010.2070470.

[11] K.-M. Persson, B. G. Malm, and L.-E. Wernersson, "Surface and core contribution to 1/f-noise in InAs nanowire metal-oxide-semiconductor field-effect transistors," *Appl. Phys. Lett.*, vol. 103, no. 3, pp. 033508–1–033508–4, Jun. 2013, doi: 10.1063/1.4813850.

[12] R. Pandey, B. Rajamohanan, H. Liu, V. Narayanan, and S. Datta, "Electrical noise in heterojunction interband tunnel FETs," *IEEE Trans. Electron Devices*, vol. 61, no. 2, pp. 552–560, Feb. 2014, doi: 10.1109/TED.2013.2293497.

Paper IV

Paper IV

Reproduced, with permission, from

E. Memisevic, M. Hellenbrand, E. Lind, A. R. Persson, S. Sant, A. Schenk, J. Svensson, R. Wallenberg, and L.-E. Wernersson, "Individual Defects in InAs/InGaAsSb/GaSb Nanowire Tunnel Field-Effect Transistors Operating below 60 mV/decade," *Nanoletters*, vol. 17, no. 7, pp. 4373–4380, Jun. 2017, doi: 10.1021/acs.nanolett.7b01455.

NANO LETTERS

Letter

pubs.acs.org/NanoLett

Individual Defects in InAs/InGaAsSb/GaSb Nanowire Tunnel Field-Effect Transistors Operating below 60 mV/decade

Elvedin Memisevic,*[†] Markus Hellenbrand,[†] Erik Lind,[†] Axel R. Persson,[‡,§] Saurabh Sant,[∥] Andreas Schenk,[∥] Johannes Svensson,[†] Reine Wallenberg,[‡,§] and Lars-Erik Wernersson[†]

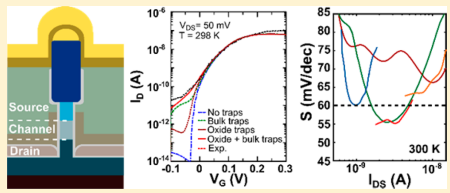
[†]Department of Electrical and Information Technology, Lund University, Lund 221 00, Sweden

[‡]Center for Analysis and Synthesis, Lund University, Box 124, 221 00 Lund, Sweden

[§]NanoLund, Lund University, Box 118, 22100 Lund, Sweden

[∥]Integrated Systems Laboratory, ETH Zürich, 8092 Zürich, Switzerland

ABSTRACT: Tunneling field-effect transistors (TunnelFET), a leading steep-slope transistor candidate, is still plagued by defect response, and there is a large discrepancy between measured and simulated device performance. In this work, highly scaled InAs/$In_xGa_{1-x}As_ySb_{1-y}$/GaSb vertical nanowire TunnelFET with ability to operate well below 60 mV/decade at technically relevant currents are fabricated and characterized. The structure, composition, and strain is characterized using transmission electron microscopy with emphasis on the heterojunction. Using Technology Computer Aided Design (TCAD) simulations and Random Telegraph Signal (RTS) noise measurements, effects of different type of defects are studied. The study reveals that the bulk defects have the largest impact on the performance of these devices, although for these highly scaled devices interaction with even few oxide defects can have large impact on the performance. Understanding the contribution by individual defects, as outlined in this letter, is essential to verify the fundamental physics of device operation, and thus imperative for taking the III−V TunnelFETs to the next level.



KEYWORDS: *TFET, nanowire, InAs, GaSb, InGaAsSb*

To address the power constraint and to augment the metal-oxide-semiconductor field-effect transistor (MOSFET) functionality for low-power applications, a novel device needs to operate well below the thermal limit (60 mV/decade), but also provide the same current levels although at lower drive voltages. Tunneling field-effect transistor (TunnelFET) is a device that uses band-pass filtering to remove high energy carriers and can be designed to operate below the thermal limit.[1,2] TunnelFETs fabricated from Si and III−Vs have demonstrated promising results.[3] Si TunnelFETs have shown the ability to operate well below 60 mV/decade, although not yet at any technically useful current levels.[4−6] III−V TunnelFETs have achieved high currents but without operation significantly below 60 mV/decade.[7−9] One of the main limitations to achieve a subthreshold swing (S) well below 60 mV/decade with III−V TunnelFETs are defects in the materials that enable defect-assisted tunneling,[10,11] which have had a detrimental effect on the subthreshold swing and the off-state currents. Thereby, to fully explore the potential of the III−V TunnelFETs, there is a need to understand what influence the different types of defects have on the performance. In this work, we present highly scaled nanowire III−V TunnelFETs with competitive performance, and with a subthreshold swing well below the thermal limit of 60 mV/decade at room temperature. This outstanding performance shows that these devices have a comparably low defect density, which allows studies of individual defects. Using transmission electron microscopy (TEM), the structure is characterized, determining the composition, crystal structure, and strain. Modeling is performed to understand the impact of defects in the channel, as well as in the oxide, on the performance of these devices. Finally, using Random Telegraph Signal (RTS) noise measurements, the influence of oxide defects was studied experimentally. The findings presented here allow for TunnelFET optimization, which can increase the energy efficiency of electronic systems. It may also enable new applications in areas where the energy budget is limited, for instance, in battery operated or energy harvesting applications related to the Internet of Things.

Nanowires were epitaxially grown using metal−organic vapor phase epitaxy (MOVPE) from Au seed particles patterned using electron beam lithography on a high resistivity Si(111) substrate with a 260 nm highly n-doped InAs layer.[12] The InAs and InGaAsSb segments were grown at 460 °C, while the GaSb segment was grown at 515 °C (see methods). The bottom part of the InAs segment was n-doped using Sn and the InGaAsSb

Received: April 7, 2017
Revised: May 28, 2017
Published: June 14, 2017

ACS Publications © 2017 American Chemical Society

DOI: 10.1021/acs.nanolett.7b01455
Nano Lett. 2017, 17, 4373−4380

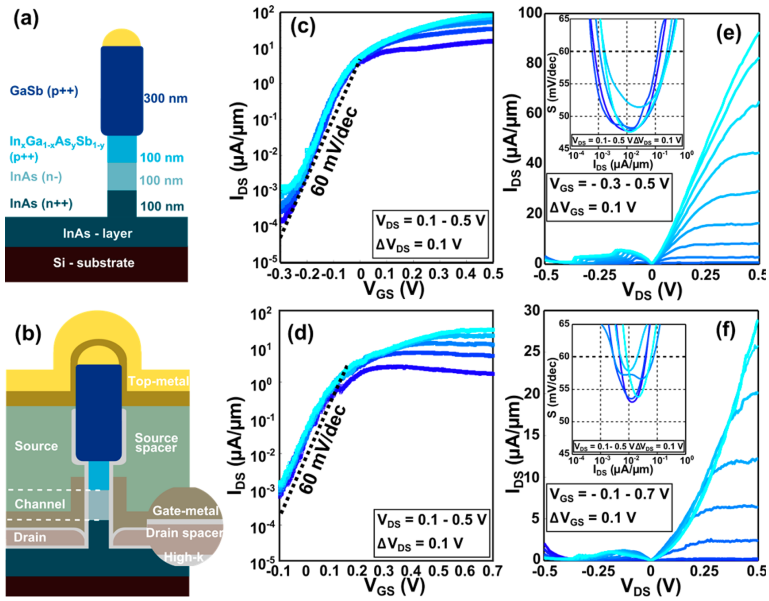


Figure 1. Nanowire structure and electrical measurements. (a) Schematic illustration of a nanowire with different sections marked with different colors (doping type and section lengths are indicated). (b) Schematic illustration of the vertical InAs/InGaAsSb/GaSb nanowire TunnelFET, showing all layers (right side) and regions (left side) of the transistor. High-κ is a bilayer of Al_2O_3/HfO_2 with EOT 1.4 nm, drain spacer is 15 nm SiO_x, gate metal is a 60 nm-thick tungsten layer, source spacer is an organic photoresist film, and the top metal is 15/150 nm Ni/Au film. Relative thickness of the layers in the image is exaggerated to enhance visibility. (c,d) Transfer data for drive voltages 0.1−0.5 V for device A and B, respectively. Devices show a small DIBL of 25/37 mV/V (A/B) and reaches an on-current of 10.3/3.9 μA/μm (A/B) at V_{DS} = 0.3 V and I_{off} = 1 nA/μm (e,f) Output data for device A and B, respectively. Device A shows larger on-currents than device B, reaching 92 μA/μm compared to the 29 μA/μm at $V_{DS} = V_{GS}$ = 0.5 V. This is mainly due to the source depletion observed in device B. Both devices have clear NDR peaks in backward direction reaching peak-to-valley current ratio of 14.6 and 14.2 for device A and B, respectively. Insets in both figures show a I_{DS} vs S graph, which confirms sub-60 mV/decade operation down to 48 and 53 mV/decade for device A and B, respectively.

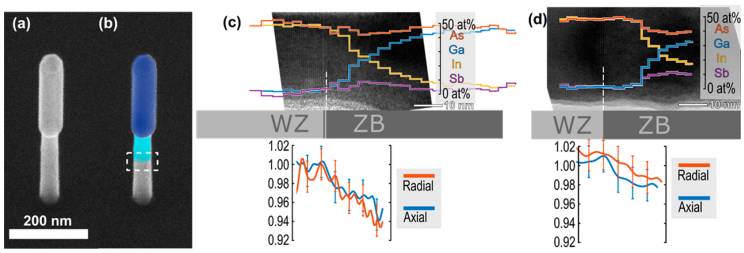


Figure 2. Characterization of nanowires. (a) SEM image of one nanowire from sample A taken directly after growth, prior to digital etching. The diameter of the nanowire increases when the GaSb segment is grown due to increased solubility of the group III materials in the Au-particle. (b) The same nanowire as in image (a) with InGaAsSb and GaSb segments colored. The marked box shows the section of main interest for the TEM study. (c,d) Illustrates the compositional and structural changes for sample A and B, respectively. The compositional changes in atomic % for respective element are shown in colored curves with legend and scale to the right, overlaying an HRTEM image of the transition. To ensure statistical significance, the curves represent the average of volume segments across the wire with a width of the individual steps. In the band across the top, the crystal structure segments are indicated, and at the bottom, the lattice measurements are indicated along with error bars. The radial measurements are normalized to InAs d_{112} for ZB (corresponding to InAs $d_{2\bar{1}10}$ for WZ).

DOI: 10.1021/acs.nanolett.7b01455
Nano Lett. 2017, 17, 4373−4380

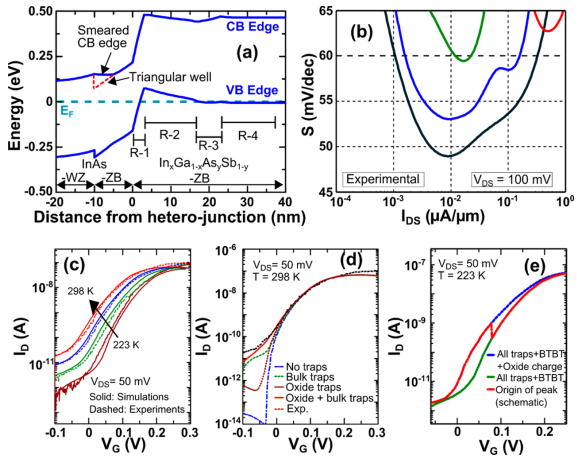


Figure 3. Simulations of the band structure. (a) Band edge diagram along the axis of the nanowire at V_{DS} = 0 V and V_{GS} = 0 V. A triangular-like quantum well is present at the InAs-ZB/InAs-WZ interface. (b) Current vs subthreshold swing for four different devices from same sample. All devices have one nanowire. (c) Comparison of experimental and simulated transfer characteristics for different temperatures. (d) Contribution for different type of the effects of InAs/oxide defects and bulk defects near the heterojunction. (e) Comparison of transfer characteristics with and without a trapped positive charge at the InAs/oxide interface 50 nm away from the heterojunction.

and GaSb segments were p-doped using Zn leaving a 100 nm-long not intentionally doped InAs channel segment (Figure 1a). To study how a change in the band lineup at the tunnel junction affects the electrical characteristics, two different InGaAsSb compositions were grown, $In_{0.1}Ga_{0.9}As_{0.88}Sb_{0.12}$ for sample A and $In_{0.32}Ga_{0.68}As_{0.72}Sb_{0.28}$ for sample B. Fabrication of the devices began with digital etching of the nanowires to remove a parasitic GaSb shell and to reduce the diameter of the InAs channel region down to 20 nm. A high-k layer was applied using atomic layer deposition followed by evaporation of a spacer layer to separate the drain and gate regions. The gate layer was formed by sputtering of a tungsten film, followed by definition of the physical length (L_g = 250 nm) using reactive ion etching and mask definition. A second spacer layer was applied to separate the gate and source regions followed by sputtering of Ni/Au top-metal and formation of probing pads. A detailed description of the fabrication process can be found in the methods section. A schematic image of a finished device is presented in Figure 1b. The transfer characteristics of a device from sample A (device A) and B (device B) are shown in Figure 1c,d, respectively. Both devices exhibit excellent electrostatics (low drain-induced barrier lowering), as expected for the gate-all-around geometry. Output characteristics for both devices shows good saturation (Figure 1e,f). Device B shows a superlinear current onset, typical for TunnelFETs. The well-behaved output data for both devices confirms the proper material choice for TunnelFETs. For an I_{off} = 1 nA/μm (low power logic) and V_{DS} = 0.3 V, the I_{on} current for device A and B is 10.6 and 3.9 μA/μm, respectively. Notably, both devices demonstrate a subthreshold swing below the thermal limit of 60 mV/decade, reaching 48 mV/decade for device A and 53 mV/decade for device B. Operation below 60 mV/decade occurs for both types of devices over a wide current range, close to 3 orders of magnitude for device A. This is exemplified by a record high I_{60} current (the current level at which the slope equals 60 mV/decade) of 0.31 μA/μm at 0.3 V for device A. These devices show a factor of almost 100 higher I_{60} as compared to other implementations approaching relevant current levels.[13] The high I_{60} originates from the use of a narrow bandgap channel, excellent electrostatics, and very low defect density in the InAs/InGaAsSb heterojunction. The TunnelFET operation is largely determined by the properties of the heterojunction. Consequently, the heterojunction of the nanowires, presented in Figure 2a,b, was examined in detail using transmission electron microscopy. The composition was determined using energy dispersive X-ray spectroscopy and corroborated by measurements of the lattice plane differences from high resolution images. Most of the InAs segment has a wurtzite (WZ) crystal structure, but the final part before the introduction of Ga, approximately 3−5 nm in sample A and 10 nm in sample B, has a zincblende (ZB) structure most likely formed during the switching sequence. The transition from InAs to InGaAsSb is graded over ∼25 nm, and both the InGaAsSb and the GaSb segment have pure zincblende crystal structure without any observed stacking faults, which is typical for Sb-containing nanowires.[14] The local lattice distances are reduced after the structural change from WZ to ZB (Figure 2c,d). Measurements are normalized to InAs(ZB),[15] d_{111} and d_{112} for the axial and radial measurements, respectively. The observed reduction is expected since the large amount of Ga lowers the average atomic radii. However, on both samples this reduction starts at the WZ−ZB transition, (Figure 2c,d) corresponding to 3−5 and 10 nm prior to the measured Ga increase in sample A and B, respectively. The InAs(ZB) is thereby compressed radially, which is expected since it tries to adapt to the smaller InGaAsSb lattice. However, Poisson's ratio would predict an expansion in the axial direction, which has not been observed. Exact strain is difficult to evaluate due to

DOI: 10.1021/acs.nanolett.7b01455
Nano Lett. 2017, 17, 4373−4380

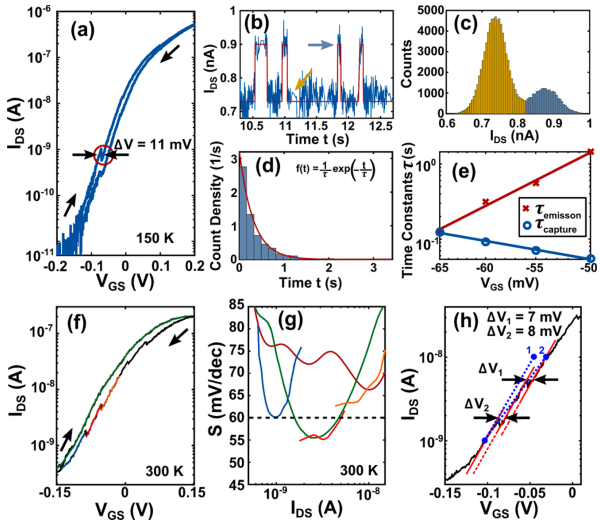


Figure 4. RTS measurements. (a) Transfer characteristic for a single nanowire transistor at 150 K. The effect of a single oxide defect is clearly visible as a distinct step. (b) Excerpt of an RTS measurement in time domain. Blue, measured signal; red, accentuation of RTS steps. The arrows indicate the upper and the lower current levels. (c) Representative histogram of the current for the entire measurement in (b). Two distinct current levels are clearly visible. The colors refer to the colors of the arrows in (b). (d) Representative histogram for the extraction of the emission time constant for the measurement in (b). The exponential distribution $f(t)$ fitted to extract the time constant τ is indicated in red. (e) Time constants varying close to the step indicated in (a). The same time constant for capture and emission of an electron occurs at exactly the step. (f) Transfer characteristics for a device with a number of distinct steeps in the backward sweep. (g) Subthreshold swings extracted in more detail for the different regions in (f). The individual slopes (blue, bright red, yellow) and the slope without visible steps (green) are at/below 60 mV/decade. Adding the influence of several individual defects results in an average slope (dark red), which is clearly degraded compared with the individual slopes. (h) The subthreshold region of the device presented in (f) where several distinct RTS steps are visible. Each step results from the discharge of an individual oxide defect. The individual slopes are almost parallel, and any average across the individual steps will degrade the total device slope significantly. The average slope is increased from 55 to 70 mV/decade.

nonlinearity of the resulting lattice for a given composition. However, assuming a linear relation between the composition and lattice distances and comparing to the measured lattice spacing, the local axial strain can be estimated to about 3% for sample A and 2% for sample B, both compressive, with its maximum in the InAs(ZB) region closest to the WZ−ZB transition and extending into the InGaAsSb region. It is known for similar compounds that can exist in both WZ and ZB polytypes that the WZ structure is the radially stiffer of the two (given a growth direction of [111]B or [0001]).[16] Applying this to InAs imposes restrictions on the ideal way of relaxation, and since the stiffer WZ is less prone to deformation, it retains its radial dimensions over a longer distance. This explains the steeper drop in axial lattice distance in the section immediately after the WZ−ZB interface for sample B. The more graded introduction of Ga into sample A provides a smoother transition, which is closer to the ideal adaption of the lattice to the final composition.

Based on the knowledge of the material composition and the strain profile, simulations of the InAs/InGaAsSb/GaSb nanowire TunnelFET were performed using the semiclassical simulator Sentaurus-Device. The actual device geometry and estimated doping profiles were used to analyze the performance of device B. All band structure parameters were set to the experimentally extracted values taken from ref 17. The InGaAsSb segment is divided into subregions based on its composition (Figure 3a). The formation of a quantization level in the triangular well at the InAs WZ-ZB interface was modeled by introducing a pseudograding of the CB edge (Figure 3a). The TEM images revealed a steep change of the mole fraction in the 3 nm-long InGaAsSb segment denoted by R-1. The device is most sensitive to defects within this region. Best fit to the measured data is acquired by choosing a defect density (N_t) of 1.6×10^{18} cm^{-3}, which would result in exactly one defect state in the segment R-1 on spatial integration. With so few defects, we would almost expect a digital-like variation between devices. Data from four devices with one nanowire from the same sample presented in Figure 3b shows that the minimal value for the subthreshold swing occurs at approximately the same I_{DS} although the devices show a strong variation in the subthreshold swing vs I_{DS}. Some of the variations could depend on some variations generated during processing, but part could also be from variations among position and energy of the few defects. Donor-like defects were assumed at the InAs/oxide interface. The comparison of simulated temperature-dependent transfer characteristics with the experimental data (Figure 3c) confirmed good agreement between the two, which validates the selected parameter set and simulation setup. Using a minimal number of fitting parameters allows to reliably analyze the impact of individual degradation mechanisms. The effect of

DOI: 10.1021/acs.nanolett.7b01455
Nano Lett. 2017, 17, 4373−4380

each defect type on the transfer characteristics is shown in Figure 3d. The simulations confirm that the subthreshold swing could be significantly improved if all defects were suppressed. Notably, introduction of defects at the oxide/semiconductor interface alone caused only a minor degradation of the subthreshold swing. This is due to the absence of surface inversion at the oxide/semiconductor interface as consequence of the small diameter, which disables the formation of a strong enough normal field required for defect-assisted tunneling (DAT). Due to the low doping in the InAs segment and the small diameter, the channel region is under accumulation leading to flat band conditions. Yet, a small subthreshold-swing degradation is observed when defects at the oxide/semiconductor interface are included, which is due to screening of the gate charge resulting in a weaker gate coupling. Additionally, oxide interface defects cause surface Shockley–Read–Hall (SRH) generation of electron–hole pairs by multiphonon excitation, which adds to the leakage current at very low current levels. Bulk defects present in region R-1 near the heterojunction are found mainly responsible for the degradation of the subthreshold swing. Therefore, suppressing these defects would have a stronger impact on the improvement of the TunnelFET performance compared with suppressing oxide interface defects. Nevertheless, suppressing oxide interface defects further would reduce the leakage current level. In the above analysis, the oxide interface defects were assumed to be uniformly distributed at the interface. Although a uniform distribution reliably models the average impact of these defects, a localized single defect may have a strong electrostatic effect, which can induce RTS type of effects on the I–V characteristics. Similar effects are also found in small area MOSFETs.[18] Transfer characteristics with and without a localized positive InAs/oxide interface charge located 50 nm away from the heterointerface are compared in Figure 3e. An acceptor defect at the InAs/oxide interface is positively charged in the off-state. At a certain gate bias, the charged defect captures an electron neutralizing this charge. This affects the electrostatics of the TunnelFET, and the on-current is sharply reduced after the positive charge has been removed. This gives rise to a peak in the transfer characteristics as schematically shown in Figure 3e.

To experimentally verify the role of individual defects within the gate oxide of the TunnelFETs with this small dimensions, the transfer characteristics of different transistors were scanned for RTS noise. Measuring the time constants and current step sizes provides information about the effect of single electron charges on the potential in the channel. Capture of an electron in an individual acceptor-type defect within the gate oxide reduces the current level as the channel potential energy is increased; emission of the electron increases the current. Since the current in TunnelFETs is most sensitive to potential fluctuations when operating in the subthreshold region, we expect to detect the strongest RTS noise below V_T. Experimentally, we observe relative RTS amplitudes $\Delta I_{DS}/I_{DS}$ of about 20% for the device in Figure 4 and even up to 50% for other devices not shown. The large RTS amplitude originates mainly from the small channel diameter, where a single defect can have a strong influence on the entire channel potential. Furthermore, the current through a TunnelFET is mainly limited by the potential variations close to the tunnel junction, which enhances RTS response from oxide defects in this region.[19] The effect of an individual defect on the subthreshold characteristics is presented in Figure 4a. An excerpt of a two-level RTS measurement at a constant bias point close to the step in Figure 4a is presented in Figure 4b. To be able to reliably extract the time constants, at least a few hundred (and up to a few thousand) current transitions were recorded for each bias point. The residual noise in Figure 4b could originate from 1/f-type noise from the channel region or from additional RTS noise too fast to resolve. The time constants, at a fixed bias, for capture and emission of an electron were determined by fitting an exponential distribution to the recorded capture/emission times as shown in Figure 4d. When the channel Fermi level is aligned with the defect energy level, the capture and emission time constants are equal and are set by the depth z of the defect from the channel into the oxide: $\tau_{c/e} = \tau_0 \times \exp(z/\lambda)$.[20] Here, $\tau_{c/e}$ is the capture/emission time constant, the constant τ_0 corresponds to capture into defects right at the MOS interface, and λ is the tunneling attenuation length $\lambda = \left[\frac{4\pi}{h}\sqrt{2m^*\Phi_B}\right]^{-1}$ according to the Wentzel–Kramers–Brillouin theory[20] with the Planck constant h, the oxide effective mass m^*, and the energy barrier height between the channel and the oxide Φ_B. With symmetrical time constants measured to range between 100 and 300 ms for different devices, the probed depth into the oxide amounts to approximately 2.8 nm for $\tau_0 = 10^{-10}$ s.[21] From the identified depths of the defects and with simplified electrostatic considerations, it can be estimated that the change from a single defect within the gate oxide is in the order of a few meV. This agrees well with the corresponding change of a few mV of gate voltage around the RTS step in the transfer curves of Figure 4a,f. Reducing the channel Fermi level below the defect energy level favors the unoccupied defect state, which increases the capture time constant and decreases the emission time constant (Figure 4e). Qualitatively, the strong effect of even individual oxide defects can be seen in Figures 4f–h. With the resolution of individual oxide defects, Figure 4g,h explains how these defects deteriorate the overall TunnelFET subthreshold swing. The slopes without current steps are well below or at 60 mV/decade, but any average over all individual defect current steps increases the overall slope significantly. A more detailed analysis of the slopes in Figure 4h reveals that, besides the dominant current steps, the change in electrostatic potential also slightly alters the tunneling dynamics, which can be seen from the deviation between the measured data in Figure 4g and the adjusted, parallel lines in Figure 4h. The overall impact of all effects on the average slope between $I_{DS} = 10^{-9}$ and 10^{-8} A is an increase from 55 to 70 mV/decade, as illustrated in Figure 4g.

We have demonstrated vertical III–V nanowire TunnelFETs with ability to operate well below the thermal limit of 60 mV/decade with currents in technically relevant range. This performance is achieved based on high quality and excellent electrostatics, due to high scaling and gate-all-around geometry. Yet, bulk defects in the proximity of the heterojunction still limits the device from reaching even lower subthreshold swing. Although, the origin of the defect level at present is unknown and it may originate from several contributions, like shallow or deep impurities, structural defects, or process damage, the presence of a high strain field at the heterojunction and the limited knowledge about dopant atoms in nanowire geometries suggest that these causes will first need to be investigated to identify the origin. Furthermore, the necessary scaling makes these devices also sensitive to the individual oxide defects. Our detailed investigations show that there is room for further improvement in device performance.

DOI: 10.1021/acs.nanolett.7b01455
Nano Lett. 2017, 17, 4373–4380

Methods. *Growth.* Growth was performed using arrays of Au discs with a thickness of 15 nm and diameter of 44 nm, which were patterned by EBL on high resistivity Si(111) substrates with a 260 nm highly doped InAs layer on top. The nanowires were grown using metal−organic vapor phase epitaxy (MOVPE) in an Aixtron CCS 18313 reactor with a total flow of 8000 sccm at a pressure of 100 mbar. After annealing at 550 °C, a 200 nm-long InAs segment was grown at 460 °C using trimethylindium (TMIn) and arsine (AsH_3) with a molar fraction of $X_{TMIn} = 6.1 \times 10^{-6}$ and $X_{AsH3} = 1.3 \times 10^{-4}$, respectively. The bottom part of the InAs segment was n-doped by triethyltin (TESn) ($X_{TESn} = 6.3 \times 10^{-6}$). The InAs segment was followed by a 100 nm (In)GaAsSb segment with different compositions for samples A and B using trimethylgallium (TMGa) ($X_{TMGa} = 4.9 \times 10^{-5}$), trimethylantimony (TMSb) ($X_{TMSb} = 1.2 \times 10^{-4}$), and AsH_3 (sample A, $X_{AsH3} = 2.7 \times 10^{-5}$; sample B, $X_{AsH3} = 1.3 \times 10^{-5}$) corresponding to a gas phase composition of $AsH_3/(AsH_3+TMSb) = 0.18$ for sample A and 0.094 for sample B. A 300 nm-long GaSb segment was subsequently grown while heating to 515 °C using TMGa ($X_{TMGa} = 4.9 \times 10^{-5}$) and TMSb ($X_{TMSb} = 7.1 \times 10^{-5}$). The GaAsSb and the GaSb segments were both p-doped during growth using diethylzinc (DEZn) ($X_{DEZn} = 3.5 \times 10^{-5}$).

Structural and Compositional Analyses. Structural and compositional analyses were performed using a JEOL 3000F transmission electron microscope (TEM), operated at 300 kV, with emphasis on the InAs/InGaAsSb transition. The wires were transferred to lacy carbon covered Cu-grids by pressing these onto the substrate in order to break off the wires. Both high-resolution TEM (HRTEM) and scanning TEM (STEM), employing a high angle annular dark-field (HAADF) detector, were used. For the HRTEM structural measurements, the wires were imaged in the $\langle 110\rangle/\langle 11\bar{2}0\rangle$ zone-axis. Compositional analysis was performed, using the STEM-HAADF mode in combination with the XEDS-detector to map the transition, data that were used for qualitative measurements such as length of transition and quantitative measurements for the composition along the wire. Small structural variation, such as axial and radial strain, can be measured directly in real space in HRTEM images[22] or, as used here, indirectly in Fourier space by geometric phase analysis (GPA).[23] The latter is implemented as a script in Digital Micrograph (Gatan Inc.), and the former is done manually in the same software.

Fabrication of the Devices. Fabrication of the devices started with digital etching of the nanowires using ozone plasma to oxidize the surface of the nanowires and citric acid to remove the oxide, reducing mainly the InAs diameter to 20 nm thereby improving the electrostatics of the devices.[24] The InGaAsSb segment was also thinned down although with a lower rate, and no visible etching of the GaSb could be observed. Using atomic layer deposition (ALD), the nanowires were covered with a high-k bilayer using five cycles of Al_2O_3 and 36 cycles of HfO_2 at temperatures of 300 and 120 °C, respectively. Estimated oxide thickness (EOT) for the high-k material is 1.4 nm. A 15 nm-thick SiO_x drain-gate spacer was applied using thermal evaporation with rotation and zero tilt. The flakes of the SiO_x on the sidewalls of the nanowires where removed with diluted HF followed by applying 12 cycles of HfO_2 at 120 °C to compensate the thinning. The gate was fabricated by first sputtering a 60 nm-thick tungsten (W) layer, followed by definition of the physical gate-length (L_g) using a resist etch-back process with O_2-plasma in a reactive ion etching (RIE) system. Tungsten was removed from the exposed sections with SF_6/Ar. Utilizing photoresist and UV-lithography followed by etching of W with RIE, the gate-pad was defined. A gate-source spacer was fabricated by a spin-on photoresist (S1800) followed by etch-back process in RIE to determine the final thickness. Gate-via and drain-via were fabricated using UV-photolithography and RIE. To realize drain and source contacts, the high-κ was removed on top of the nanowires and in the drain-via using HF. The final step was fabrication of the top-metal, by sputtering of 10 nm of Ni and 150 nm of Au followed by UV-lithography and wet-etching to define the contact pads.

Modeling of Band Edge and Defect Profiles. Modeling of band edge and defect profiles was carried out based on the knowledge on the material composition and the strain profile. The actual device geometry and estimated doping profiles were used to analyze the performance of device B. The diameter of the TunnelFET was set to 20 nm. As revealed from the TEM analysis, the channel consists of different segments of InGaAsSb with linearly varying compositions in the p+ doped source with a doping level of 10^{19} cm^{-3}. The unintentionally doped channel (background doping of 10^{17} cm^{-3}) consists of an InAs-WZ segment with an additional InAs-ZB segment adjacent to the heterointerface. The InGaAsSb segment is divided into subregions based on its composition (Figure 3a). In region R-1, alloy composition varies from InAs (to the left) to $In_{0.7}Ga_{0.3}As_{0.84}Sb_{0.16}$ (to the right), henceforth denoted as $In_{1\rightarrow 0.7}Ga_{0\rightarrow 0.3}As_{1\rightarrow 0.84}Sb_{0\rightarrow 0.16}$. Region R-2 consists of $In_{0.7\rightarrow 0.44}Ga_{0.3\rightarrow 0.56}As_{0.84\rightarrow 0.72}Sb_{0.16\rightarrow 0.28}$. Region R-3 is composed of $In_{0.44\rightarrow 0.32}Ga_{0.56\rightarrow 0.68}As_{0.72}Sb_{0.28}$, while R-4 consists of $In_{0.32}Ga_{0.68}As_{0.72}Sb_{0.28}$. The composition is assumed to vary linearly within each segment. Various band structure quantities such as effective mass values, band gaps, and electron affinities are required for a meaningful simulation of the TunnelFET. All band structure parameters have been set to the experimentally extracted values taken from ref 17. An interpolation formula suggested by Adachi[25] has been used to obtain the above quantities for the quaternary alloy $In_xGa_{1-x}As_ySb_{1-y}$ for all intermediate compositions. In this way, any explicit fitting of the band structure quantities and the band offsets have been avoided. As described in the previous section, uniaxial compressive strain is present at the heterointerface. The effect of strain on the band alignment has been modeled by the model-solid theoretical approach by Van de Walle. Simulations of the InAs/InGaAsSb nanowire TunnelFET were performed using the semiclassical simulator Sentaurus-Device. The inability to account for quantum confinement effects is a drawback of semiclassical simulations. The quantization of electronic states in the triangular-like potential well at the InAs WZ−ZB interface (Figure 3a) results in the formation of discrete energy levels. Our calculations suggest that the triangular quantum well will form only one bound state whose energy level is very close to the top of the finite barrier. This will effectively smear out the otherwise steep valley at the interface. Simulations showed that ignoring this effect causes a serious discrepancy between simulated and experimental data. Therefore, the above effect has been modeled by introducing a pseudograding of the CB edge at the interface thus making it continuous (Figure 3a). Effects related to geometrical confinement in radial direction start to emerge in III−V semiconductors when the diameter falls below 20 nm. Hence, they are not expected to significantly alter the band structure parameters (band gap, effective mass, etc.) in the source, channel, and drain segments and have been ignored. The

DOI: 10.1021/acs.nanolett.7b01455
Nano Lett. 2017, 17, 4373−4380

analysis of the TEM images reveals a steep change of the mole fraction in the 3 nm-long InGaAsSb segment R-1. This may induce defect states in that segment. The degradation of the TunnelFET characteristics due to defect states has been taken into account by creating a nonlocal mesh adjacent to the heterointerface and activating both direct and phonon-assisted nonlocal defect-assisted-tunneling (DAT) models. Choosing a peak N_t of 1.6×10^{18} cm^{-3} would result in exactly one defect state in the segment R-1 on spatial integration. Since the exact position of a single defect within the segment R-1 is unknown and cannot be probed, in a first approximation the N_t was assumed to be constant in the whole segment. The variation of the defect energy level within the band gap then leads to best agreement with the experimental data when it is located 0.1 eV above the VB edge. A simulation parameter called trap interaction volume, which is a measure of the coupling strength of the tunneling process and plays the role of a scaling factor for the generation rate in the nonlocal DAT model, was adjusted to 10 Å^3 to match the simulated DAT current level to the experimental one. Note that this parameter is related, but not identical to the volume of the localized wave function. In addition to defects near the heterojunction, defects at the oxide/semiconductor interface may degrade the TunnelFET performance. Donor-like defects were assumed and placed at the InAs/oxide interface. Their energetic distribution was adapted from an earlier study.[26] Defects at the InGaAsSb/oxide interface are screened by the central region of the nanowire due to its high doping concentration. Furthermore, the Fermi level in InGaAsSb is located below the VB edge, which leaves the defects unfilled and electrostatically inactive. Hence, defects at the InGaAsSb/oxide interface were not considered in the TCAD analysis.

RTS Noise. RTS noise was measured at temperatures of 11, 150, and 300 K using a setup consisting of a Lake Shore Cryotronics CRX-4K probe station and an Agilent B2912A SMU. Data with time resolutions between 0.2 and 2 ms was collected at several different fixed gate and drain voltages using measurement times long enough to obtain at least several hundred and up to a few thousand transitions.

The transition time constants (τ) were determined from the measured data by fitting the exponential distribution (eq 1) to a histogram of the capture/emission times.

$$f(t) = \frac{1}{\tau} e^{-t/\tau} \tag{1}$$

The distance z from the channel interface into the gate oxide was determined according to

$$z = \lambda \ln\left(\frac{\tau}{\tau_0}\right) \tag{2}$$

where λ is the tunneling attenuation length, τ the measured time constant, and τ_0 a constant. Here, we chose $\tau_0 = 10^{-10}$ s according to ref 21 as is often found in the literature. We are aware, however, that there does not seem to be a consensus on what this constant should be and also, e.g., $\tau_0 = 6.6 \times 10^{-14}$ can be found.[27] In any case, changing this number will not change the physical interpretation of our results but instead only shift the estimated depth of the defects in the gate oxide. The tunneling attenuation length λ is calculated according to

$$\lambda = \left(\frac{4\pi}{h}\sqrt{2m^*\phi_B}\right)^{-1} \tag{3}$$

where h is the Planck constant, m^* the effective electron mass in the gate oxide, and Φ_B the barrier height from the channel material to the gate oxide. Here, we used $m^* = 0.23\ m_0$ with the electron rest mass m_0 and $\Phi_B = 2.3$ eV,[28] which results in $\lambda = 0.13$ nm.

AUTHOR INFORMATION

Corresponding Author

*E-mail: elvedin.memisevic@eit.lth.se.

ORCID

Elvedin Memisevic: 0000-0003-0464-1123

Author Contributions

E.M. fabricated the devices, performed DC characterization, did data analysis, wrote: abstract, introduction, fabrication section, and conclusions, and produced Figure 1. M.H. performed the noise measurements, analyzed the noise characteristics, wrote most of the noise characterization text, and produced Figure 4. E.L. contributed to the analysis of the data. A.R.P performed the TEM analysis, analyzed the TEM data, wrote material characterization section, and produced Figure 2. S.S. performed simulations of heterostructures, helped analyze the data, wrote the modeling section, and produced Figure 3. A.S. supervised the modeling and helped to analyze the results. J.S. grew the nanowires and wrote the text about the growth. R.W. supervised the TEM characterization and helped to analyze the results. L.E.W. directed the project and contributed to the analyzing and writing of the manuscript. All authors discussed the data and commented on the manuscript.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

This work was supported by the Swedish Foundation for Strategic Research, Swedish Research Council, Knut and Alice Wallenberg, NanoLund, and the European Union Seventh Framework Program E2SWITCH under Grant 619509.

REFERENCES

(1) Ionescu, A. M.; Riel, H. Tunnel field-effect transistors as energy-efficient electronic switches. *Nature* **2011**, *479*, 329−337.

(2) Seabaugh, A. C.; Zhang, Q. Low-Voltage Tunnel Transistors for beyond CMOS Logic. *Proc. IEEE* **2010**, *98*, 2095−2110.

(3) Lu, H.; Seabaugh, A. C. Tunnel Field-Effect Transistors: State-of-the-Art. *IEEE J. Electron Devices Soc.* **2014**, *2*, 44−49.

(4) Knoll, L.; et al. Inverters With Strained Si Nanowire Complementary Tunnel Field-Effect Transistors. *IEEE Electron Device Lett.* **2013**, *34*, 813−815.

(5) Gandhi, R.; et al. CMOS-Compatible Vertical-Silicon-Nanowire Gate-All-Around p-Type Tunneling FETs With ≤ 50-mV/decade Subthreshold Swing. *IEEE Electron Device Lett.* **2011**, *32*, 1504−1506.

(6) Tomioka, K. et al. Tunnel Field-Effect Transistors using III-V Nanowire/Si Heterojunction. In *IEEE Symposium on VLSI Technology*; IEEE, 2012; pp 47−48.

(7) Dewey, G. et al. Fabrication, characterization, and physics of III-V heterojunction Tunneling Field-Effect Transistors (H-TFET) for steep sub-threshold swing. In *IEEE International Electron Device Meeting*; IEEE, 2011; pp 785−788.

(8) Ahn, D. H. et al. Performance improvement of $In_xGa_{1-x}As$ Tunnel FETs with Quantum Well and EOT scaling. In *IEEE Symposium on VLSI Technology*; IEEE, 2016; pp 224−225.

(9) Rajamohanan, B.; et al. 0.5 V Supply Voltage Operation of $In_{0.65}Ga_{0.35}As/GaAs_{0.4}Sb_{0.6}$ Tunnel FET. *IEEE Electron Device Lett.* **2015**, *36*, 20−22.

DOI: 10.1021/acs.nanolett.7b01455
Nano Lett. 2017, 17, 4373−4380

Paper IV © 2017 American Chemical Society

(10) Moselund, K. E.; et al. Lateral InAs/Si p-Type Tunnel FETs Integrated on Si − Part 1: Experimental Devices. *IEEE Trans. Electron Devices* **2016**, *63*, 4233−4239.
(11) Sant, S.; et al. Lateral InAs/Si p-Type Tunnel FETs Integrated on Si − Part 2: Simulation Study of the Impact of Interface Traps. *IEEE Trans. Electron Devices* **2016**, *63*, 4240−4247.
(12) Ghalamestani, S. G.; et al. High quality InAs and GaSb thin layers grown on Si (111). *J. Cryst. Growth* **2011**, *332*, 12−16.
(13) Memisevic, E.; et al. Vertical InAs/GaAsSb/GaSb Tunneling Field-Effect Transistor on Si with S = 48 mV/decade and I_{on} = 10 μA/μm for I_{off} = 1nA/μm at V_{DS} = 0.3 V. In *IEEE International Electron Device Meeting*; IEEE: 2016; pp 500−503.
(14) Borg, B. M.; Wernersson, L.-E. Synthesis and properties of antimonide nanowires. *Nanotechnology* **2013**, *24*, 1−18.
(15) Kriegner, D.; et al. Unit cell structure of crystal polytypes in InAs and InSb nanowires. *Nano Lett.* **2011**, *11*, 1483−1489.
(16) Hanada, T. Basic properties of ZnO, GaN, and related materials. *Adv. Mater. Res.* **2009**, *12*, 1−19.
(17) Vurgaftman, I.; et al. Band parameters for III-V compound semiconductors and their alloys. *J. Appl. Phys.* **2001**, *89*, 5815−5875.
(18) Toledano-Luque, M.; et al. Fast Ramped Voltage Characterization of Single Trap Bias and Temperature Impact on Time-Dependent V_{TH} Variability. *IEEE Trans. Electron Devices* **2014**, *61*, 3139−3144.
(19) Pandey, R.; et al. Electrical Noise in Heterojunction Interband Tunnel FETs, IEEE. *IEEE Trans. Electron Devices* **2013**, *61*, 552−560.
(20) Christensson, S.; et al. Low-Frequency Noise in MOS Transistors − I Theory. *Solid-State Electron.* **1968**, *11*, 797−812.
(21) von Haartman, M.; Östling, M. *Low-Frequency Noise in Advanced MOS Devices*; Springer: 2007; pp 47, 68.
(22) Carlsson, A.; et al. Strain state in semiconductor quantum dots on surfaces: a comparison of electron microscopy and finite element calculations. *Surf. Sci.* **1998**, *406*, 48−56.
(23) Hÿtch, M. J.; et al. Quantitative measurement of displacement and strain fields from HREM micrographs. *Ultramicroscopy* **1998**, *74*, 131−146.
(24) Memisevic, E.; et al. Scaling of Vertical InAs-GaSb Nanowire Tunneling Field-Effect Transistors on Si. *IEEE Electron Device Lett.* **2016**, *37*, 549−552.
(25) Adachi, S. Band gaps and refractive indices of AlGaAsSb, GaInAsSb, and InPAsSb: Key properties for a variety of the 2−4 mm optoelectronic device applications. *J. Appl. Phys.* **1987**, *61*, 4869−4876.
(26) Wu.; et al. Low Trap Density in InAs/High-*k* Nanowire Gate Stacks with Optimized Growth and Doping Conditions. *Nano Lett.* **2016**, *16*, 2418−2425.
(27) Lundström, L.; Svensson, C. Tunneling to traps in insulators. *J. Appl. Phys.* **1972**, *43*, 5045−5047.
(28) Li, N.; et al. Properties of InAs metal-oxide-semiconductor structures with atomic-layer-deposited Al_2O_3 dielectric. *Appl. Phys. Lett.* **2008**, *92*, 14−16.

DOI: 10.1021/acs.nanolett.7b01455
Nano Lett. 2017, 17, 4373−4380

 Paper IV © 2017 American Chemical Society

Paper V

Paper V

Reproduced, with permission, from

M. Hellenbrand, E. Memišević, J. Svensson, E. Lind, and L.-E. Wernersson, "Random Telegraph Signal Noise in Tunneling Field-Effect Transistors with S below 60 mV/decade," *European Solid-State Device Research Conference (ESSDERC)*, pp. 38–41, Sep. 2017, doi: 10.1109/ESSDERC.2017.8066586.

© 2017 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission.

Random Telegraph Signal Noise in Tunneling Field-Effect Transistors with S below 60 mV/decade

Markus Hellenbrand, Elvedin Memišević, Johannes Svensson, Erik Lind, and Lars-Erik Wernersson
Department of Electrical and Information Technology
Lund University, Box 118
S-22100 Lund, Sweden

***Abstract*—Single gate oxide defects in strongly scaled Tunneling Field-Effect Transistors with an inverse subthreshold slope well below 60 mV/decade are investigated by Random Telegraph Signal (RTS) noise measurements. The cause for RTS noise are electrons being captured in and released from individual defects in the gate oxide. Under the assumption that elastic tunneling is the underlying capture and emission mechanism, the measured RTS time constants vary with the relative position of the channel Fermi level and the defect energy level while the amplitudes – independent of the capture and release mechanism – follow the inverse of the inverse subthreshold slope.**

Keywords—Tunneling Field-Effect Transistors, Nanowires, Below 60 mV/decade, Random Telegraph Signal Noise, Elastic Tunneling

I. Introduction

Transistor dimensions are an important scaling parameter to continue Moore's law. Also for devices beyond Moore's law, certain transistor dimensions remain important metrics for optimized device performance. For vertical nanowire transistors, one of the key dimensions is the diameter of the nanowires, strongly affecting the electrostatic control of the gate over the channel [1]. With excellent electrostatic control and a carefully designed III-V heterojunction, we demonstrated Tunneling Field-Effect Transistors (TFETs) with an inverse subthreshold slope (S) well below 60 mV/decade [2]. At the same time, while improving the electrostatic gate control, scaling down transistor dimensions to a few nanometers limits the number of gate oxide defects to a few individual locations. Capturing and releasing individual electrons in and from these defects changes the channel potential energy by a discrete amount and thus – specific for TFETs with a dominant defect close to the junction – also the reverse bias of the tunneling junction. This leads to discrete steps in the device current, so-called Random Telegraph Signal (RTS) noise, and results in degradation of transistor and circuit performance [3]. In TFETs, this effect has been simulated thoroughly [3, 4], but rarely been observed experimentally [5, 6] and explained in detail even less. Here, we study RTS noise in TFETs with S well below 60 mV/decade, which allows estimations of both the spatial as well as energetical position of the dominant defect and its electrostatic effect on the channel potential.

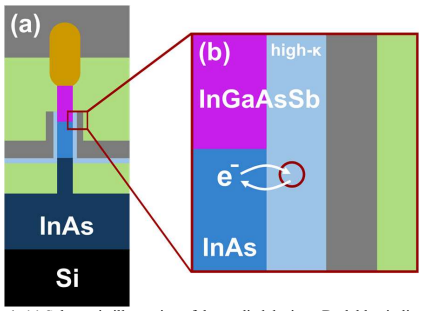


Fig. 1: (a) Schematic illustration of the studied devices. Dark blue indicates InAs, doped in the bottom, intrinsic in the channel region, purple indicates the InGaAsSb segment, yellow the GaSb segment, light blue indicates the gate oxide, grey are metals, green spacers. (b) Schematic of the tunneling process causing RTS noise. The red circle indicates a single defect in the gate oxide.

II. Device and Measurement Specifications

A. Device Fabrication

To clearly separate and identify all effects, only transistors made of one single nanowire were studied for this work. A schematic illustration is shown in Fig. 1. All nanowires were grown by metalorganic vapor phase epitaxy (MOVPE) from an Au seed particle defined by electron beam lithography (EBL). The substrate consisted of a MOVPE-grown InAs buffer layer on top of Si. The InAs bottom segment of the nanowire was n-doped with Sn to reduce series resistance. In order to achieve a gate control as good as possible, the channel segment on top was not intentionally doped. The tunneling junction was created

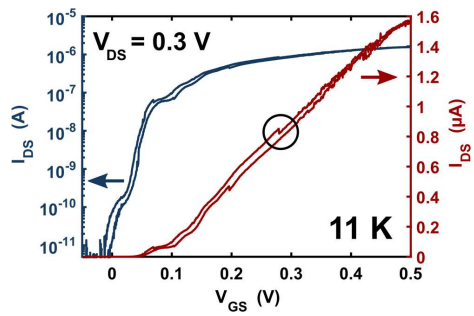


Fig. 2: Transfer characteristics for an example device at 11 K with V_{DS} = 0.3 V. Blue shows a logarithmic scale (left), red a linear scale (right). The circle indicates the RTS step more closely examined in this article. For the highest currents in the linear scale there are many more RTS steps.

978-1-5090-5978-2/17/$31.00 ©2017 IEEE

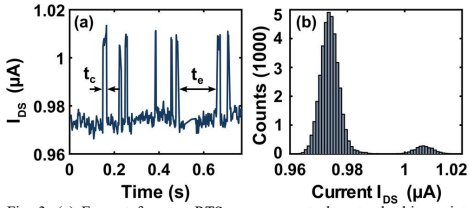


Fig. 3: (a) Excerpt from an RTS measurement close to the bias point indicated in Fig. 2. t_c and t_e denote the capture and the emission times, respectively. (b) Current histogram for the complete measurement in (a), clearly indicating two current levels.

by changing from the un-doped InAs channel segment to p-doped (using Zn) InGaAsSb. For good source contacts, the InGaAsSb segment was followed by a p-doped GaSb top segment. After growth, the InAs parts of the nanowire were thinned down to a final diameter of 20 nm by first oxidizing the surface with ozone plasma and then etching this oxide with citric acid. Immediately afterwards, a high-κ gate oxide consisting of approximately 1 nm of Al_2O_3 and 3.5 nm of HfO_2 was applied by atomic layer deposition (ALD). The estimated equivalent oxide thickness (EOT) is 1.4 nm. A SiO_x bottom spacer was deposited to separate the drain from the gate contact and a 60-nm-thick W gate metal was sputtered over the whole sample. To define the gate length from this coverage, another (temporary) spacer was applied and etched back so that all W beyond the intended gate contact could be removed from the nanowire. The gate pad in the lateral direction was defined by UV-lithography. After exchanging the temporary spacer for the final photoresist top spacer, the high-κ oxide was removed from the top of the nanowire and a 10 nm Ni/150 nm Au source metal contact was deposited. As a last step, drain and gate vias were etched and the top metal source, drain, and gate pads were defined, all by UV-lithography. The transfer characteristics of a representative device is shown in Fig. 2. For consistency, all data presented here are from this same device and the analysis is focused on the RTS step marked in Fig. 2. Further details on fabrication can be found in [2].

B. Measurement Setup and Technique

The measurement setup consisted of a Lake Shore Cryotronics CRX-4K probe station and an Agilent B2912A

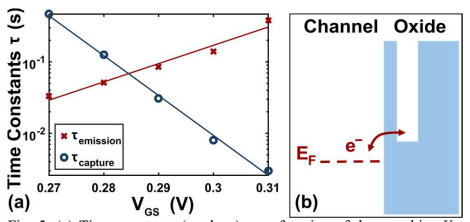


Fig. 5: (a) Time constants (markers) as a function of the gate bias V_{GS} together with exponential fits (lines). A decreasing capture time constant for an increasing emission time constant was observed for all TFETs. (b) Schematic illustration of the capture and emission process for a single oxide defect. The trends in (a) can be explained by the difference between the channel Fermi level and the defect energy level.

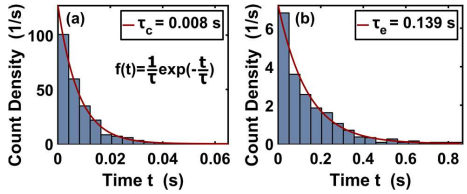


Fig. 4: Histograms of the capture (a) and emission (b) times for the measurement from Fig. 3. The red lines indicate the exponential distribution (equation in (a)) fitted to the data.

Source/Measure Unit. RTS noise measurements were carried out at room temperature, at 150 K, and at 11 K. However, due to the high sensitivity of the noise signal to even small changes in the device behavior and due to threshold voltage shifts at different temperatures, it was not possible to identify and measure with certainty the same defect at different temperatures. In each measurement like the one presented in Fig. 3, the drain current (InAs side) of a transistor was measured in sections of up to 200 seconds at constant gate and drain bias, while the source (GaSb top contact) was grounded. For each defect and each bias point, at least a hundred and up to 2200 transitions between current levels were recorded with measurement resolutions between 0.2 ms and 2 ms. The capture and emission time constants were determined by fitting an exponential distribution to the histograms of the times spent in the high current states (capture times) and the low current states (emission times), respectively (Fig. 4). The data presented here are from a measurement at 11 K, but are representative for measurements at all temperatures.

III. Results and Discussion

Both the measured capture and emission time constants τ_c and τ_e vary with varying gate bias (Fig. 5(a)), which can be qualitatively explained with the help of the process illustrated in Fig. 5(b) under the assumption that elastic tunneling is the capture and emission mechanism. If the channel Fermi level is increased towards the defect energy level from below, electron tunneling into the oxide defect exhibits a rather long time constant due to the energy difference between the two states. The emission time constant in this configuration is rather short since the electron can always tunnel out of the defect and relax to the energy corresponding to the channel Fermi level. Increasing the channel Fermi level above the trap energy level reverses this behavior. Since the effect of the defect considered here shows up in the on-state, the defect energy level must be located quite far below the channel conduction band in the unbiased state.

Besides the energy difference explained above, the tunneling time constant is determined by the depth z of the defect from the channel into the oxide. According to standard quantum-mechanical tunneling the probability of a state in a forbidden area – and thus the tunneling time constant τ – decays exponentially as

$$\tau_{c/e} = \tau_0 \exp(z/\lambda) \qquad (1)$$

where τ_0 is a constant, and λ the tunneling attenuation constant according to the Wentzel-Kramers-Brillouin approximation [7] given by

$$\lambda = \left(\frac{4\pi}{h}\sqrt{2m^*\phi_B}\right)^{-1} = 0.13 \text{ nm}, \qquad (2)$$

where h is the Planck constant, $m^* = 0.23 \cdot m_0$ the effective mass in the gate oxide (m_0 is the electron rest mass), and $\phi_B = 2.3$ eV the barrier height from the channel to the gate oxide [8]. With the assumptions made above, z can be determined from the biasing point where $\tau_c = \tau_e$ since in this configuration the channel Fermi level and the defect energy level are aligned and there is no energy difference, which distorts the time constants. With $\tau_{e/c} \approx 6 \cdot 10^{-2}$ s from Fig. 5(a), $\tau_0 = 4 \cdot 10^{-11}$ s from [9], and (1) rearranged to $z = \lambda \ln(\tau/\tau_0)$, the depth z amounts to approximately 2.7 nm from the channel interface into the gate oxide and thus to approximately halfway between the channel and the gate metal. From this calculation it is obvious that z depends on τ_0, for which different values can be found in literature, ranging from 10^{-10} s [7] to $6.6 \cdot 10^{-14}$ s [10]. Between these two extreme values, the position of the defect in the oxide would change by up to 1 nm, which does, however, not influence the underlying tunneling model. The value of $\tau_0 = 4 \cdot 10^{-11}$ s chosen here is taken from our own high-frequency characterizations of Metal Oxide Semiconductor Field-Effect Transistors (MOSFETs) similar to the presented TFET structures and corresponds to tunneling into defects which are very close to the channel/gate oxide interface.

Independent of the exact capture and emission mechanism, the RTS amplitude can be used to estimate the change in the channel potential energy that results from a single electron captured in a defect in the oxide. With the simplified expression

$$I_{1D} = a\,(V_R - V_T)\exp\left(-\frac{b}{\xi}\right) \qquad (3)$$

for a 1D TFET [11], it can be readily shown that the change in the TFET current from one single charge q_{ox} in the gate oxide amounts to

$$\frac{\partial I_{1D}}{\partial q_{ox}} = \frac{\partial V_R}{\partial q_{ox}} g_m, \qquad (4)$$

which also holds for a 3D expression of the current. In (3) and (4), a and b are constants summarizing elementary constants and material parameters, V_R is the reverse bias applied to the tunneling junction, V_T is the thermal voltage kT/q, ξ denotes the electrical field across the tunneling junction and g_m is the transconductance. From (4) it is clear that the change in the channel potential energy can be determined from the transfer characteristics. For this purpose, a more detailed illustration of the encircled RTS step in Fig. 2 is provided in Fig. 6.

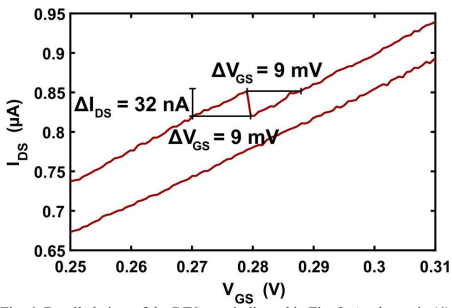


Fig. 6: Detailed view of the RTS step indicated in Fig. 2. As shown in (4), the change of the channel potential energy can be directly determined from the step.

According to a simple model of the surface potential Ψ_S, it depends on the gate voltage V_{GS} as

$$\Delta\Psi_S = \Delta V_{GS}\frac{C_{ox}}{C_{ox} + C_q + C_{it}}, \qquad (5)$$

where C_{ox} is the gate oxide capacitance, C_q is the semiconductor capacitance resulting from a finite density of states in the channel, and C_{it} represents interface defects between channel and gate oxide. The centroid capacitance C_c is neglected for the sake of simplicity, and furthermore it is masked by the small C_q. Since V_R denotes the change of Ψ_S due to the change in V_{GS}, (5) can be applied to V_R in exactly the same way as to Ψ_S.

The step ΔI_{DS} in the current indicated in Fig. 6 is caused by an abrupt change $\partial V_R/\partial q_{ox}$ in the tunnel junction reverse bias V_R due to a single captured charge q_{ox} in the gate oxide. To obtain the same ΔI_{DS} by changing the gate bias V_{GS}, a ΔV_{GS} of 9 mV would be required, as obtained from Fig. 6. To estimate the actual $\partial V_R/\partial q_{ox}$ from (5), we need to calculate C_{ox}, C_q, and C_{it}. C_{ox} can be simply calculated as a cylindrical capacitor, and C_q can be calculated as

$$C_q = \sqrt{\frac{2m^*}{E_F - E_1(0)}}\frac{q^2}{\pi\hbar} \qquad (6)$$

for the approximation of a 1D channel with a large C_{ox} [12], where $m^* = 0.023 \cdot m_0$ is the effective mass in the channel, $E_F - E_1(0)$ denotes the position of the Fermi level above the first subband, $\hbar$ the reduced Planck constant, and q the elementary charge. C_{it} can be calculated as $C_{it} = q^2 D_{it}$, where $D_{it} \approx 5 \cdot 10^{12}$ eV^{-1} cm^{-2} close to the conduction band edge [13]. With values of $C_{ox} \approx 0.033$ F/m^2, $C_q \approx 0.002$ F/m^2 ($E_F - E_1(0) \approx 0.15$ eV from Fig. 2 and (5) and C_q divided by the channel circumference for equal units), $C_{it} \approx 0.008$ F/m^2, and $\Delta V_{GS} = 9$ mV from Fig. 6, (5) results in $\partial V_R/\partial q_{ox} \approx 6.9$ mV.

With this value for the change in the channel potential energy, the plausibility of the calculated depth of the defect z can be validated further by electrostatic calculations. Approximating the cylindrical nanowire by a square nanowire with the same cross-sectional area and modelling a single charge at the calculated depth of 2.7 nm with the method of image charges results in a change in the channel potential energy of around 2.2 meV. Although this value is somewhat lower than the experimental one, it is still very close, given the uncertainties in some of the parameters used. Possible explanations are as follows. First of all, the assumption of a 1D density of states might be too optimistic and a larger 2D C_q would yield even more similar results for the calculated and the measured values ($\partial V_R/\partial q_{ox} \approx 5$ meV for C_q in 2D). Also, C_{it} might increase from the value from [13] when moving the bands further into the on-state, or the single dominant defect could be located closer to the channel interface than 2.7 nm. This would imply that τ_0 is larger than our choice of $4 \cdot 10^{-11}$ s. Furthermore, of course, the simulated model of image charges might be too simple. Again, however, with these sources of uncertainty taken into account,

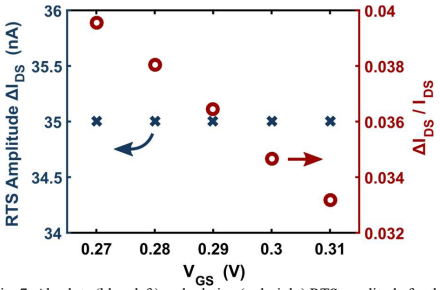


Fig. 7: Absolute (blue, left) and relative (red, right) RTS amplitude for the same RTS step as shown in the other figures. In accordance with (7), the relative RTS amplitude decreases with increasing gate voltage. Note that in accordance with (4), g_m is constant in this area.

both the experimental and the simulated values agree very well, which strongly supports the model of simple elastic tunneling as the mechanism for the observed RTS noise.

Finally, independent of the capture and release mechanism, the behavior of the relative RTS amplitude $\Delta I_{DS}/I_{DS}$ for the whole operation range can be described by extending (4) to

$$\frac{\Delta I_{DS}}{I_{DS}} = \frac{\partial V_R}{\partial q_{ox}} \frac{\ln(10)}{S}, \qquad (7)$$

where ΔI_{DS} corresponds to $\partial I_{ID}/\partial q_{ox}$, I_{DS} to I_{ID} from (3) and S denotes the subthreshold slope calculated from (3) as well. Just as with (4), (7) also holds for the 3D case. This expression suggests a peak of the relative RTS amplitude in the subthreshold region, where S is steepest. Before and after this region it suggests a decay. And indeed, for the measurement series presented here, $\Delta I_{DS}/I_{DS}$ shows a monotonous decrease with increasing gate voltage, depicted in Fig. 7. For other devices and bias points, we were able to measure the other regions described by (7) as well, both increasing relative RTS amplitudes far into the subthreshold region and values around the expected peak. In the on-state, where the measurement series presented here was carried out, the relative RTS amplitude is as low as 3 – 4 %, but measurements in the subthreshold region showed values of up to 50 %. This large relative amplitudes strongly suggest a defect location close to the tunneling junction since defects far away from the junction will not affect the tunneling probability in a similarly strong way [4].

IV. Conclusions

We presented a detailed explanation of RTS noise in III-V nanowire TFETs with inverse subthreshold slopes well below 60 mV/decade. The origin of RTS noise are electrons captured in and released from individual gate oxide defects close to the tunneling junction and the behavior of the RTS amplitude can be explained by the expression of the TFET tunneling current in a straight forward manner. In the case presented here, the energy level of the defect is located far below the channel conduction band. For the explanation of the RTS time constants we assume elastic tunneling as the capture and emission process and the resulting depth of the oxide defect, $z \approx 2.7$ nm from the channel interface, agrees well with electrostatic calculations.

Acknowledgments

This work was supported in part by the Swedish Foundation for Strategic Research, Swedish Research Council, and the European Union Seventh Framework Program E2SWITCH under Grant 619509.

References

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

[2] E. Memišević, J. Svensson, M. Hellenbrand, E. Lind, and L.-E. Wernersson, "Vertical InAs/GaAsSb/GaSb tunneling field-effect transistor on Si with S = 48 mV/decade and Ion = 10 μA/μm for Ioff = 1 nA/μm at Vds = 0.3 V," in *2016 IEEE International Electron Devices Meeting (IEDM)*, 2016, pp. 19.1.1-19.1.4.

[3] R. Pandey, B. Rajamohanan, H. Liu, V. Narayanan, and S. Datta, "Electrical Noise in Heterojunction Interband Tunnel FETs," *IEEE Transactions on Electron Devices,* vol. 61, no. 2, pp. 552-560, 2 2014.

[4] M. L. Fan, V. P. H. Hu, Y. N. Chen, P. Su, and C. T. Chuang, "Analysis of Single-Trap-Induced Random Telegraph Noise and its Interaction With Work Function Variation for Tunnel FET," *IEEE Transactions on Electron Devices,* vol. 60, no. 6, pp. 2038-2044, 2013.

[5] J. Wan, C. L. Royer, A. Zaslavsky, and S. Cristoloveanu, "Low-frequency noise behavior of tunneling field effect transistors," *Applied Physics Letters,* vol. 97, no. 24, p. 243503, 2010.

[6] Q. Huang *et al.*, "Deep insights into low frequency noise behavior of tunnel FETs with source junction engineering," in *2014 Symposium on VLSI Technology (VLSI-Technology): Digest of Technical Papers*, 2014.

[7] M. v. H. a. M. Östling, *Low-Frequency Noise in Advanced MOS Devices*. Springer, 2007.

[8] N. Li *et al.*, "Properties of InAs metal-oxide-semiconductor structures with atomic-layer-deposited Al2O3 Dielectric," *Applied Physics Letters,* vol. 92, no. 14, p. 143507, 2008.

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

[10] I. Lundström and C. Svensson, "Tunneling to traps in insulators," *Journal of Applied Physics,* vol. 43, no. 12, pp. 5045-5047, 1972.

[11] A. C. Seabaugh and Q. Zhang, "Low-Voltage Tunnel Transistors for Beyond CMOS Logic," *Proceedings of the IEEE,* vol. 98, no. 12, pp. 2095-2110, December 2010.

[12] M. Lundstrom and J. Guo, *Nanoscale Transistors*. Springer, 2006.

[13] J. Wu *et al.*, "Low Trap Density in InAs/High-k Nanowire Gate Stacks with Optimized Growth and Doping Conditions," *Nano Letters,* vol. 16, no. 4, pp. 2418-2425, 2016.

Paper VI

Paper VI

Reproduced, with permission, from

M. Hellenbrand, E. Lind, O.-P. Kilpi, and L.-E. Wernersson, "Effects of traps in the gate stack on the small-signal RF response of III-V nanowire MOSFETs," *Solid-State Electronics*, under review with minor revisions pending, May 2020.

© 2020 M. Hellenbrand. Personal use is permitted, but republication/redistribution requires permission from the authors.

Effects of traps in the gate stack on the small-signal RF response of III-V nanowire MOSFETs

Markus Hellenbrand*, Erik Lind, Olli-Pekka Kilpi, Lars-Erik Wernersson

Department of Electrical and Information Technology, Lund University, Ole Römers Väg 3, 223 63 Lund, Sweden

Abstract

We present a detailed study of the effect of gate-oxide-related defects (traps) on the small-signal radio frequency (RF) response of III-V nanowire MOSFETs and find that the effects are clearly identifiable in the measured admittance parameters and in important design parameters such as h_{21} (forward current gain) and MSG (maximum stable gain). We include the identified effects in a small-signal model alongside results from previous investigations of III-V RF MOSFETs and thus provide a comprehensive physical small-signal RF model for this type of transistor, which accurately describes the measured admittance parameters and gains. We verify the physical basis of the model assumptions by calculating the oxide defect density from the measured admittances.

Keywords: Border Traps, Gate Oxide Defects, Interface Defects, III-V, MOSFET, RF, Small-Signal Model

1. Introduction

Metal-oxide-semiconductor field-effect transistors (MOSFETs) are essential components in high frequency electronics. III-V MOSFETs, due to their narrow band gap and high electron mobility, can achieve higher operation frequencies [1] than corresponding Si MOSFETs, which dominate the digital market. For digital applications, even after several years of investigations into the III-V/high-κ material system, III-V MOSFETs are severely impaired by scalability and defects related to the gate oxide, so-called border and interface traps, which degrade reliability and lifetime of the transistors [2]. Two of the most common characterization techniques for traps are bias temperature instability (BTI) measurements in the time domain [3] and capacitance voltage (CV) measurements in the frequency domain [4]. Both techniques have greatly advanced our understanding of defects and led to significant improvement of transistor material qualities, which is reflected in nearly ideal inverse subthreshold slopes and high transconductances [5]. In both techniques, however, the direct observation of traps is typically limited to the megahertz regime and faster traps can only be accessed by decreasing the measurement temperature [6]. The effects of traps in the gate stack on the MOSFET radio frequency (RF) response have thus not yet been investigated in much detail. The few publications which take into account traps in RF investigations to a certain extent, usually focus on a single aspect only and simplify the small-signal model, in which the traps are studied, e.g. by disregarding non-quasi-static (NQS) effects [7–9]. A comprehensive description of the small-signal RF model, which takes into account all different aspects of the contributions of gate-oxide-related traps to the admittance parameters (y-parameters), including both real and imaginary parts, and which does not disregard NQS effects, is yet to be developed.

Here, we demonstrate that despite the improvements in recent years, frequency dispersion in III-V MOSFETs at room temperature due to traps can be considerable even at gigahertz frequencies and we investigate the resulting effects on the small-signal y-parameters, the RF gains, and the stability factor. Based on our measurements, we provide a comprehensive small-signal model, which takes into account the observed dispersion and which accurately models the measurements. This model also takes into account previous results on small-signal analyses and can be applied at all bias points in the on-state of a transistor and in the off-state before the onset of minority carrier effects. We choose a small-signal approach for this analysis rather than a large-signal model to deconvolute the effect of traps on the RF response from their effect on the quiescent point. This simplifies the analysis and allows us to investigate the effects more clearly. As such, the presented model is not intended for circuit simulations, but it can be used e.g. for small-signal-based amplifier design.

The paper is structured as follows. First, in Section 2, we introduce the devices, which were used to develop and verify the small-signal model and explain the measurement setup and data processing. In Section 3, we briefly discuss the different kinds of defects, which can affect MOSFETs and point out assumptions, which enter into our analysis. In Section 4, we develop expressions for all small-signal model components and demonstrate the effects of traps with measured data in Sections 5 and 6. Lastly, in Section 7, we verify that the model assumptions, which take into account oxide traps, are physically meaningful by calculating the trap densities from the measured admittances.

*Corresponding author

Email address: markus.hellenbrand@eit.lth.se (Markus Hellenbrand)

2. Devices and measurement

The development of the small-signal model presented here was based on measured data from a set of vertical III-V nanowire gate-all-around MOSFETs. The channels consisted of arrays of 180–300 parallel vertical nanowires of InAs graded to InGaAs on the drain side. The nanowire diameters varied between 25 and 30 nm. The high-κ gate oxide consisted of a 1 nm/4 nm Al_2O_3/HfO_2 bilayer and gate lengths varied between 100 nm and 220 nm. Details about processing can be found in [10]. For consistency, all measured data presented in this paper are from the same device.

For all investigated devices, scattering parameters (*s*-parameters) were measured with a vector network analyzer between 10 MHz and 67 GHz and for the subsequent analysis the *s*-parameters were converted to *y*-parameters. The measurement setup was calibrated off-chip by a load-reflect-reflect-match (LRRM) calibration procedure; on-chip open and short de-embedding structures were used to remove the effect of the pad parasitics. The parameters of the developed small-signal model were then determined by fitting the model to the measured *y*-parameters in Matlab.

The model was developed with a data set measured on the sample described above and verified by subsequent modeling of about thirty additional III-V devices on four different samples. Of the four samples, one was similar to the one described, one contained lateral nanowire MOSFETs, one contained MOSFETs with planar channels, and one contained steep-slope Tunnel FETs. Details about the samples can be found in [11–13]. While the measured parameters differed between samples due to their different geometries, their trends as a function of different bias points were consistent. Furthermore, the effects of traps on the gains, which will be described in Sections 5 and 6, can be discerned in examples from the literature as well [1, 14, 15]. Together with the different modeled samples, this demonstrates the general applicability of the model.

3. Traps in III-V MOSFET gate stacks

The application of high-κ oxides on III-V semiconductors can cause many different material defects in the oxide or at the interface between oxide and semiconductor. Conventionally, these defects are often divided into so-called border traps and interface traps, where border traps are commonly associated for instance with oxygen vacancies in the 'bulk' of the oxygen layer, i.e. at least a few monolayers away from the oxide/channel interface [16, 17]. Interface traps are believed to be related to disorder, such as dangling bonds, dimers, antisites, or oxygen vacancies at the oxide/channel interface and possibly reaching a few monolayers into the oxide [18, 19]. Density functional theory calculations reveal that all of these defects can create electrically active states in the semiconductor band gap as well as in the conduction band [16–19] in a way that all of these defects can potentially affect MOSFET performance at different biases and time scales. Despite their different chemical natures, the capture and emission of charge carriers by both types of defect can be described by general capture/emission time constants $\tau_{c/e}$, which depend on the tunneling distance and/or the activation energy associated with the charge exchange process [20, 21]:

$$\tau_{c/e} = \tau_0 \exp\left(\frac{x}{\lambda}\right) \exp\left(\frac{E_A}{k_B T}\right). \quad (1)$$

Here, x is the tunneling distance, λ the tunneling attenuation length according to the WKB approximation, E_A is the activation energy of a certain defect or defect population, k_B is the Boltzmann constant, and T the temperature. For the effective prefactor τ_0 it is usually assumed that $\tau_0 = 1/(\sigma n v_{th})$ with the defect capture cross section σ, the carrier concentration n, and the thermal velocity v_{th} of the charge carriers [20]. In correspondence with a large variation of reported values for σ, literature values for τ_0 can be found to range from microseconds [22] all the way [6, 23] to below picoseconds [24]. Physically, and in accordance with (1), τ_0 is the time constant below which none of the defects can respond any longer. In our measurements here, we observe frequency dispersions, which we attribute to traps, even at gigahertz frequencies so that $\tau_0 \approx 0.5$ ps (corresponding to an angular cutoff frequency $\omega_0 = 1/\tau_0 = 2\pi \times 300$ GHz) is required to model these dispersions. This is well within the range of values for τ_0 reported in literature and with $v_{th} = 5 \times 10^7$ cm/s and $n = 5 \times 10^{18}$ cm^{-3} [25] it results in $\sigma = 8 \times 10^{-15}$ cm^{-2}, which is a typical value for III-V materials [4, 26, 27]. As a convenient simplification, we typically treated τ_0 as a constant with respect to different bias points, but we point out that this is not a restriction, which is required to achieve agreement between measured values and our developed model.

Results from BTI measurements for example, have demonstrated that at relatively long time scales, the second exponential function in (1) usually dominates the overall time constant [28]. Eq. (1) then becomes $\tau_{c/e} = \tau_0' \times \exp(E_A/(k_B T))$ instead and the information about the location of the probed defects is hidden in the effective prefactor τ_0' with typical values for τ_0' in the nanosecond range [6, 29]. This, together with $\lambda = 0.13$ nm (for an InAs/Al_2O_3 interface), our assumption of $\tau_0 \approx 0.5$ ps, and $\tau_0' = \tau_0 \times \exp(x/\lambda)$ indicates that the RF measurements here are probing defects at a depth of about 1 nm from the channel interface into the oxide. At the same time, with the measurement frequencies from 10 MHz to 67 GHz, the corresponding time constants τ are already in the nanosecond range or below even without the exponential energy term. Thus, in the wide distribution of activation energies E_A that is assumed in BTI analyses [2, 6, 30], our measurements here seem to be probing the very smallest E_A so that the inelastic BTI models almost behave like an elastic model. In a purely elastic formulation (i.e. $E_A = 0$), the given measurement frequencies would probe depths in the oxide between 1.3 nm and 0.2 nm, see (1). Based on these considerations, we expect to probe a blend of the aforementioned interface and border traps as well as a combination of elastically and inelastically responding traps.

In the following, we will demonstrate that at radio frequencies, this blend of electrically active traps can be successfully modeled by a simple distributed RC circuit. In CV analyses, sometimes, additional elements are added to this network to

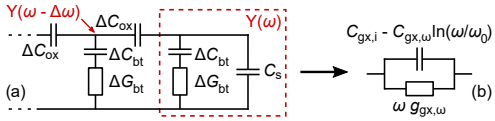


Figure 1: (a) Schematic of the distributed RC network, which models the effect of traps in a MOSFET. ΔC_{ox} are the incremental parts of the gate oxide, ΔC_t the trap capacitances, ΔG_t the conductances, which set the corresponding time constants for the charge exchange, and C_s is the semiconductor capacitance. $Y(\omega - \Delta\omega)$ and $Y(\omega)$ are used for the derivation in Section 7. (b) Lumped representation of the circuit in (a) with corresponding frequency dependences.

separate interface defects as a distinct factor [22, 31, 32]. Since a distributed RC network simply models time constants without assumptions about the chemical nature or location of the defects in question, this addition proved to be unnecessary in our analysis. We do not consider this as a claim for or against either of the two kinds of defects, however, and just note that the transition between the two is probably fluent. This makes nomenclature ambiguous for the defects probed at radio frequencies so that for the remainder of the paper we will refer to the blend of probed defects simply as 'traps'. The purpose of this paper is the inclusion of the effect of all kinds of traps in a small-signal model for the high frequency response of III-V MOSFETs.

4. Small-signal model components

The wide distribution of traps with different time constants, which affect a MOSFET, can be modeled by a distributed network of RC elements as illustrated in Fig. 1(a), where the ΔC_t model the capacitances associated with a certain trap population and the ΔG_t in series determine the corresponding time constants. The different contributions are separated by incremental parts ΔC_{ox} of the gate oxide and the RC branch next to the semiconductor capacitance C_s takes into account traps right at the interface. Carrier recombination far in the off-state is not taken into account in this model. Since the RC elements only model time constants, no assumptions are required at this point about the exact location, the chemical nature, or the exact capture mechanism of the corresponding defects.

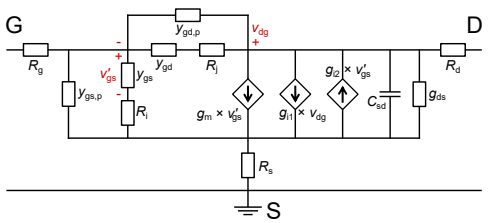


Figure 2: Complete small-signal model. $y_{gs,p}$, $y_{gd,p}$, y_{gs}, and y_{gd} each represent a lumped circuit as in Fig. 1(b). Detailed expressions for the different components are given by (1)–(6) and example values are listed in Table 1.

With a numerical solution of the voltage over C_s in the distributed network in Fig. 1(a) it can be shown that in the frequency range measured here, the distributed network can be approximated by a lumped admittance as depicted in Fig. 1(b), where the real part is linearly dependent on the angular frequency ω, and the imaginary part is linearly dependent on $-\ln(\omega/\omega_0)$ with ω_0 as discussed in Section 3. In a physical interpretation, the real part models losses due to a phase shift between the applied signal and the trap response, and the imaginary part models the capacitance due to traps capturing and emitting charges. The validity of this approximation is demonstrated by the measured data in Sections 5 and 6 and is illustrated in Fig. 4 and Fig. 6. Furthermore, similar RC networks have been successfully employed in BTI [33] and CV investigations [23] and the latter approach led to the same conclusion about the lumped frequency dependences that we use here.

In order to include the distributed RC network in the MOSFET small-signal model, equivalent lumped admittances like the one in Fig. 1(b) replace the otherwise constant intrinsic gate-to-source and gate-to-drain capacitances $C_{gs,i}$ and $C_{gd,i}$, respectively. In Fig. 2, the equivalent admittances are denoted y_{gs} and y_{gd} and with the explanation above, for $\omega < \omega_0$, they can be expressed as

$$y_{gx} = \omega \times g_{gx,\omega} + j\omega\left(C_{gx,i} - C_{gx,\omega}\ln(\omega/\omega_0)\right), \qquad (2)$$

where 'x' stands for either 's' or 'd' (source or drain), $g_{gx,\omega}$ and $C_{gx,\omega}$ are the frequency-dependent components of the conductance and the capacitance, respectively, and the $C_{gx,i}$ are the respective intrinsic capacitances without the effect of traps. While linearly frequency-dependent conductances were employed in small-signal models before [8, 9], none of those models included the trap contribution to the capacitances.

Furthermore, those previous models disregarded the NQS channel resistances R_i and R_j in Fig. 2 in series with y_{gs} and y_{gd}, respectively, which take into account the delay of charge carriers moving in the channel. This means that no distinction was made between intrinsic (y_{gs} and y_{gd} in Fig. 2) and parasitic

Table 1: Example parameters fitted to the small-signal model in Fig. 2 and used to calculate the modeled Y-parameters and gains in Fig. 5.

Parameter	Value	Parameter	Value
V_{GS}, V_{DS}	0.4 V, 0.5 V	τ_i	500 ps
V_T	0.1 V	g_{10}	50 µS
R_g	5 Ω	g_{20}	0[†]
R_s	4 Ω	ω_0	$2\pi \times 300$ GHz
R_d	16 Ω	R_i, R_j	38 Ω, 229 Ω
$g_{gs,l}$	10 µS	$g_{gd,l}$	0.5 µS
$g_{gs,\omega}$	2.5 fS/(rad/s)	$g_{gd,\omega}$	0.12 fF/(rad/s)
$C_{gs,p}$	20 fF	$C_{gd,p}$	6 fF
$C_{gs,i}$	6.0 fF	$C_{gd,i}$	1.0 fF
$C_{gs,\omega}$	0.35 fF	$C_{gd,\omega}$	0.2 fF
$C_{gsp,\omega}$	0.35 fF	$C_{gdp,\omega}$	0.2 fF
$g_{m,i}$	18.7 mS	C_m	1.0 fF
α	0.03	$C_{m,\omega}$	0.2 fF
γ_1	6×10^{-3}	γ_2	80×10^{-3}
g_{ds}	1.55 mS	C_{sd}	10 fF

[†] g_{20} is zero at this bias point, but differs from zero at lower V_{GS}.

($y_{gs,p}$ and $y_{gd,p}$ in Fig. 2 – see next paragraph) components of those models. In order to keep the complexity of the equations for our model to a minimum, we derived analytical expressions for R_i and R_j based on an RC relaxation time model [34, 35]. The expressions

$$R_i = \frac{1}{1.4\, g_{m,i}} \quad \text{and} \quad R_j = \frac{1}{1.4\, g_{m,i} \times C_{gd,i}/C_{gs,i}}, \tag{3}$$

can be obtained by equating the time it takes charge carriers to move through half of the channel with the RC constant consisting of $R_i(R_i)$ and $C_{gs,i}(C_{gd,i})$. $g_{m,i}$ is the intrinsic transconductance, i.e. without the effect of traps.

For most MOSFETs even the intrinsic device, i.e. after de-embedding pad parasitics, is subject to some parasitic elements, which in Fig. 2 are denoted $y_{gs,p}$ and $y_{gd,p}$. Typically, these parasitics consist of overlaps of the gate structure with the source and drain access regions and if they are dominated by structures, which contain oxides, they are expected to exhibit the same frequency dependence as (2). Furthermore, DC gate leakage can be considered as a parasitic effect and is included in $y_{gs,p}$ and $y_{gd,p}$ as well, so that for $\omega < \omega_0$,

$$y_{gx,p} = \omega \times g_{gxp,\omega} + j\omega\left(C_{gxp,0} - C_{gxp,\omega} \ln(\omega/\omega_0)\right) + g_{gx,l}. \tag{4}$$

Here again, 'x' stands for 's' or 'd' (source or drain), $g_{gx,l}$ takes into account the DC gate leakage, $C_{gxp,0}$ is the parasitic capacitance without the effect of traps, and all other parameters are analogous to (2). While the constant parasitic $C_{gxp,0}$ can be determined from the off-state of the transistor, the $g_{gxp,\omega}$ and $C_{gxp,\omega}$ are energy- and thus voltage-dependent due to their dependence on the trap energy distribution so that they cannot be subtracted as constant values. As a feasible approximation, we have assumed the same values for the intrinsic and the parasitic frequency dependences. This approach works well enough for modelling the y-parameters and gains and it only adds to the limitation of the quantitative analysis of the trap densities in Section 7.

Besides the admittances from (2) and (4), the transconductance g_m is also affected by traps. The numerical solution for the voltage over C_s in the distributed network in Fig. 1(a) reveals that the resulting dispersion affects both the real and the imaginary part. The complete expression for g_m below ω_0 is then

$$g_m = g_{m,i}\left[1 + \gamma_1 \ln(\omega/\omega_0) + j\alpha\,(1 + \gamma_2 \ln(\omega/\omega_0))\right] - j\omega\,(C_m - C_{m,\omega} \ln(\omega/\omega_0)), \tag{5}$$

where $g_{m,i}$ is the intrinsic transconductance, γ_1 and γ_2 determine the strength of the frequency dispersions of the real and the imaginary part, respectively, and α scales the imaginary part of this dispersion, since it is significantly smaller than the real part. A derivation of this transconductance-frequency (g_m-f) dispersion can be found in [7]. C_m in the purely imaginary part of (5) is the mutual differential capacitance $C_m = C_{dg} - C_{gd}$, which balances the charge in the channel. Since in the presence of traps, C_{dg} and C_{gd} are both frequency-dependent according to (2), C_m in (5) exhibits the respective frequency dependence as well. In all our measurements, however, this frequency dependence of C_m was masked by the γ_2 term in (5) so that we set $C_{m,\omega} = C_{gd,\omega}$.

Physically, equations like (2)–(5) should also apply to the source-drain elements C_{sd} and g_{ds}, since y_{22} is subject to the effect of traps as well. A corresponding resistance in series with C_{sd}, however, would always be masked by the large g_{ds} in par-

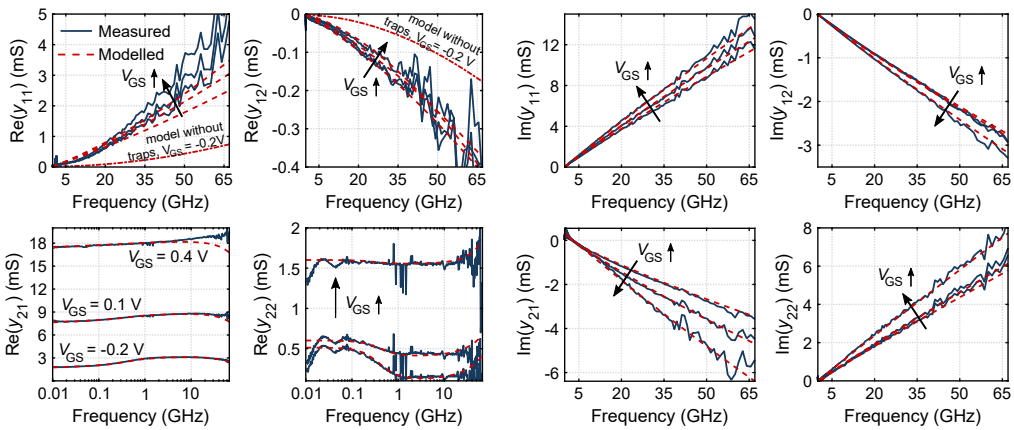


Figure 3: All y-parameters, measured (solid, blue lines) and modeled according to Fig. 2 (dashed, red lines), for three representative gate voltages V_{GS} = -0.2 V, 0.1 V, and 0.4 V. V_T = 0.1 V. Re(y_{21}) and Re(y_{22}) are plotted with logarithmic frequency axes to illustrate the frequency dispersion and the effects of impact ionization and band-to-band tunneling, respectively. In the subfigures for Re(y_{11}) and Re(y_{12}), example curves are provided for the model without traps at V_{GS} = -0.2 V and the deviation is evident. All other modeled curves fit well with the measured data. For these figures, R_s and R_d have been subtracted, but not R_g. Complete expressions for the modeled y-parameters are provided in the Appendix.

allel with C_{sd} and a g_{ds}-f dispersion corresponding to (5) was not observed. Instead, in virtually all our measurements we observed the onset of band-to-band tunneling (BTBT) or impact ionization (II) already at moderate electric fields between gate and drain. BTBT and II are well known effects in narrow bandgap MOSFETs even in DC measurements. They are not related to traps, but should be included in a comprehensive III-V MOSFET small-signal model. In Fig. 2, BTBT and II are modeled together by the current sources g_{i1} and g_{i2} with

$$g_{ik} = \frac{g_{k0}}{1 + \mathrm{j}\omega\tau_\mathrm{i}}. \tag{6}$$

Here, g_{k0} (with $k = 1, 2$) determines the magnitude and τ_i is the characteristic time constant. As an approximation, a common τ_i was used for both effects. In fact, BTBT and II also affect $\mathrm{Re}(y_{21})$ (cf. y_{21} in the Appendix), but it requires higher gate-to-drain electric fields than in $\mathrm{Re}(y_{22})$ for these effects to emerge from the g_m-f dispersion in $\mathrm{Re}(y_{21})$. Since the time constant for BTBT and II is different from the dispersion time constant of the transconductance, the effects can be separated despite appearing in parallel in the model in Fig. 2. In Fig. 3, BTBT and II can be identified by the changing slope of $\mathrm{Re}(y_{21})$ for V_GS = -0.2 V and V_GS = 0.1 V at about 0.1 GHz and by the increase of the two lower curves of $\mathrm{Re}(y_{22})$ below about 1 GHz. Here it should be noted that a change in the RF output conductance $\mathrm{Re}(y_{22})$ can also be caused by self-heating. This would result in a distinct stepwise increase in $\mathrm{Re}(y_{22})$ at about a few megahertz and the effect should increase with increasing V_GS [36]. We did not observe this in any of our measurements so that a convolution of our analysis with effects of self-heating is unlikely. Further details about BTBT and II in the small-signal model can be found in [37, 38].

The frequency-dependent expressions (2), (4), and (5) were provided with the constraint that $\omega < \omega_0$. Above ω_0, none of the traps can respond anymore and the expressions become drastically simpler. The transconductance (5) becomes $g_{\mathrm{m,i}} - \mathrm{j}\omega C_\mathrm{m}$ and the intrinsic and parasitic conductances (2) and (4), respectively, become $y_\mathrm{gx} = \mathrm{j}\omega C_\mathrm{gx,i}$ and $y_\mathrm{gx,p} = \mathrm{j}\omega C_\mathrm{gxp,0} + g_\mathrm{gx,l}$, respectively. For a physical model, this transition from the expressions below ω_0 to the ones above needs to be continuous. For the logarithmic parts, acceptable continuity is provided by the shape of the logarithm itself. For the real parts of the y_gx and $y_\mathrm{gx,p}$, a transition function would be required, which describes the decrease of the modeled losses back to zero. Since the physical ω_0 seems to lie at higher frequencies than what we were able to measure, we could not investigate this transition of the real part. This does not affect the modeled y-parameters, however, since such a transition function would only affect the model close to ω_0 and thus outside of the measurement range.

The complete expressions for the y-parameters (below ω_0) are provided in the Appendix and an example of a complete set of small-signal parameters is provided in Table 1.

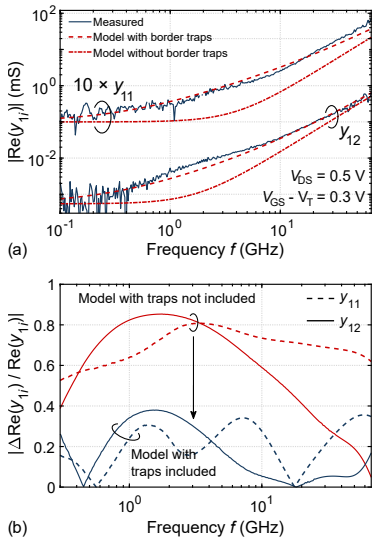


Figure 4: (a) $10 \times \mathrm{Re}(y_{11})$ (to avoid overlap) and $\mathrm{Re}(y_{12})$ on a logarithmic frequency axis. A purely quadratic model (i.e. without traps) cannot recreate the measured values properly. The model, which also includes the linear contribution from the traps, fits the data much better. (b) Relative deviation between measured and modeled $\mathrm{Re}(y_{11})$ and $\mathrm{Re}(y_{12})$ for the models with and without traps (same bias point as (a)). The inclusion of traps clearly reduces the deviations.

5. Linear components and their effect on the unilateral power gain and the stability factor

The importance of the linear contributions in (2) and (4) at high frequencies is demonstrated in Fig. 3 by the curves of $\mathrm{Re}(y_{11})$ and $\mathrm{Re}(y_{12})$ at V_GS = -0.2 V. Without it, the modeled $\mathrm{Re}(y_{11})$ is clearly lower and the modeled $\mathrm{Re}(y_{12})$ is clearly higher than the respective measured values. At first glance, on the linear scale of Fig. 3, better fits for $\mathrm{Re}(y_{11})$ and $\mathrm{Re}(y_{12})$ could be achieved by increasing their quadratic contributions, for example by increasing the NQS resistances or the gate resistance R_g. On the logarithmic scale of Fig. 4(a), however, it becomes evident that such a purely quadratic function cannot describe the frequency dependence of $\mathrm{Re}(y_{11})$ or $\mathrm{Re}(y_{12})$ correctly. The inclusion of the linear trap contribution greatly alleviates the deviation between the measured values and the purely quadratic model, as demonstrated in Fig. 4(b), where the relative differences for the models with and without traps are compared. Furthermore, the linear contributions to $\mathrm{Re}(y_{11})$ and $\mathrm{Re}(y_{12})$ are typically of different magnitudes, so that they cannot be corrected by a single common model parameter such as the gate resistance.

The linearly frequency-dependent contributions in $\mathrm{Re}(y_{11})$ and $\mathrm{Re}(y_{12})$ also readily explain the change of the slope of the unilateral power gain U, which is evident in Fig. 5. This can be understood from the equation describing U, when traps are taken into account according to (2) and (4). Starting from the general definition of U [39], at low and intermediate frequencies ($\leq$ 10 GHz in Fig. 5), where second-order frequency terms

are small, U can be written as

$$U = \frac{|y_{21} - y_{12}|^2}{4\,[\mathrm{Re}(y_{11})\,\mathrm{Re}(y_{22}) - \mathrm{Re}(y_{12})\,\mathrm{Re}(y_{21})]} \approx \frac{|g_\mathrm{m}|^2}{4\left[\omega\,(g_{\mathrm{gs},\omega} + g_{\mathrm{gd},\omega})\,g_\mathrm{ds} + \omega\,g_{\mathrm{gd},\omega}\mathrm{Re}(g_\mathrm{m})\right]}, \tag{7}$$

where the $g_{\mathrm{gx},\omega}$ contain the frequency dependences of both the intrinsic and the parasitic elements, i.e. from (2) and (4). Since (7) aims to highlight the importance of the linear frequency dependence (and for better readability), $g_{\mathrm{gs,I}}$ and $g_{\mathrm{gd,I}}$ from (4) are omitted in this approximation. In (7), the linearly frequency-dependent contributions in $\mathrm{Re}(y_{11})$ and $\mathrm{Re}(y_{12})$ dominate over the usually squared frequency dependences, which would have resulted in the well-known expression $U \propto g_\mathrm{m}^2/(4\omega^2 C_\mathrm{gg}^2)$ with the total gate capacitance C_gg. The now predominantly linear frequency dependence in the denominator of (7) changes the roll-off of U from the typical -20 dB/decade to almost -10 dB/decade. This can be observed in examples from literature as well [14, 15]. When compared with the model without traps, in a certain frequency range this leads to a decrease of the unilateral power gain by almost 10 dB. For the design of unilateral amplifiers, which are limited by U, this is a severe penalty. The constant level of U at low frequencies is caused by the constant leakage contributions $g_{\mathrm{gs,I}}$ and $g_{\mathrm{gd,I}}$ in (4). Furthermore, it can be noted that U is affected by the logarithmic frequency dependence of g_m. However, the linear components in (7) largely dominate over this logarithmic contribution.

Besides the roll-off of U, the linear components in $\mathrm{Re}(y_{11})$ and $\mathrm{Re}(y_{12})$ affect the stability factor k, as illustrated in Fig. 5 as well. The modeled k without the inclusion of traps is always lower than the measured values. In the equation for k, the effect of the linear contributions can be seen clearest at frequencies, where k is almost constant, i.e. between 1 GHz and 10 GHz in Fig. 5. At these frequencies, starting from its general definition [39], k can be written as

$$k = \frac{2\,\mathrm{Re}(y_{11})\,\mathrm{Re}(y_{22}) - \mathrm{Re}(y_{12}\,y_{21})}{|y_{12}\,y_{21}|} \approx \frac{2\omega\,(g_{\mathrm{gs},\omega} + g_{\mathrm{gd},\omega})\,g_\mathrm{ds} + \omega\,g_{\mathrm{gd},\omega}\mathrm{Re}(g_\mathrm{m})}{\left|\left(-\omega\,g_{\mathrm{gd},\omega} - \mathrm{j}\omega\,(C_{\mathrm{gd,i}} + C_{\mathrm{gd,p}})\right)g_\mathrm{m}\right|}, \tag{8}$$

where again, the $g_{\mathrm{gx},\omega}$ contain the frequency dependences of both the intrinsic and the parasitic elements and $g_{\mathrm{gs,I}}$ and $g_{\mathrm{gd,I}}$ from (4) are omitted for increased readability. In this approximation it is obvious that without the effect of traps ($g_{\mathrm{gx},\omega}$), k would become zero. Indeed, in Fig. 5, for the model without traps, k almost vanishes in the range between 1 GHz and 10 GHz, so that by the presence of traps, k is increased by a factor of up to five. Since the shift of k appears to be mostly constant in Fig. 5, the relative difference decreases towards higher frequencies, but can still amount to up to 30 % at 60 GHz. In Fig. 5, together with the effect of the traps on the gain, this lowers the k-point, i.e. the point, where MSG changes to the maximum available gain MAG, by 15 to 20 GHz. The remaining small difference of k from zero in the measured values stems from the constant leakage contributions $g_{\mathrm{gs,I}}$ and $g_{\mathrm{gd,I}}$ in $y_{\mathrm{gs,p}}$ and $y_{\mathrm{gd,p}}$, which were omitted from (8). The lower k-point decreases the frequency range available for circuit design.

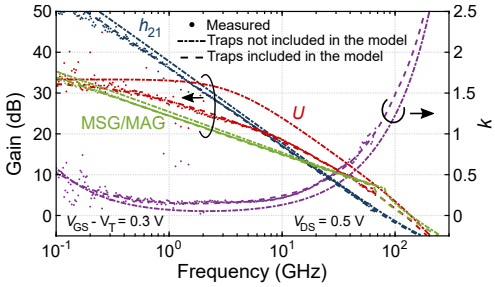


Figure 5: Gains and stability factor as a function of frequency. For each parameter, two different modeled curves are provided: The dashed lines (on top of the measured values) represent the complete model with traps, while the dash-dotted lines represent the model without the inclusion of traps.

6. Logarithmic components and their effect on the forward current gain and the maximum stable gain

Logarithmic frequency dispersions due to traps are present in the transconductance, (5), and in the intrinsic and parasitic capacitances in (2) and (4), respectively. The frequency dispersion in the real part of the transconductance is visible in $\mathrm{Re}(y_{21})$ in Fig. 3. The further increase of the measured $\mathrm{Re}(y_{21})$ at the higher V_GS beyond approx. 20 GHz, as well as the further increase of $\mathrm{Re}(y_{11})$ above approx. 35 GHz, is most likely due to the shortcomings of the open-short de-embedding at high frequencies.

Frequency dependences in capacitances due to traps have been resolved in CV measurements before [23], but not in transistor measurements. In Fig. 6(a), the measured $\mathrm{Im}(y_{11})/\omega$ and $\mathrm{Im}(y_{12})/\omega$, which approximately correspond to the capacitances in (2) and (4), are compared with the models with and without the effect of traps. The model without traps results in horizontal lines, which corresponds to ordinary capacitances, whereas the measured values clearly deviate from this behavior. Upon inclusion of the logarithmic frequency dependences in (2) and (4), the deviation readily disappears. To quantify this, the relative differences of the measured and modelled $\mathrm{Im}(y_{11})$ and $\mathrm{Im}(y_{12})$ are plotted in Fig. 6(b). At high frequencies, less traps can respond, so that the capacitances of the model, which includes traps, are reduced to the intrinsic capacitances and the difference between the models with and without traps vanishes. Towards lower frequencies, the agreement between measured and modeled values is improved by up to 20 % by the inclusion of traps in the model.

The same logarithmic deviation between measured and modeled values, which is observed in $\mathrm{Im}(y_{11})$ and $\mathrm{Im}(y_{12})$, can be identified in the forward current gain h_{21} and the maximum stable gain MSG in Fig. 5. Just as for the y-parameters, the deviation disappears upon inclusion of the frequency-dependent capacitances. Again, the difference between the two models can

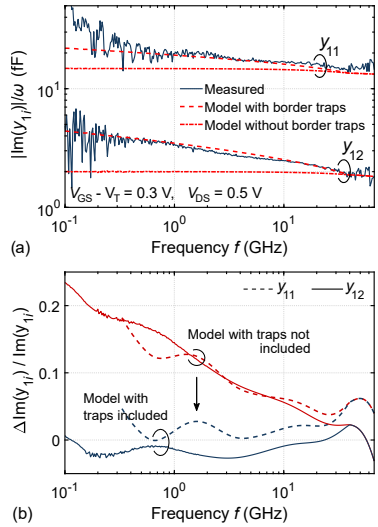


Figure 6: (a) $\mathrm{Im}(y_{11})$ and $\mathrm{Im}(y_{12})$ divided by ω to illustrate the logarithmic contribution to the capacitances. Constant capacitances result in horizontal lines and cannot reproduce the measured results, whereas capacitances with logarithmic frequency dependences according to (2) and (4) fit well with the measured results. (b) Relative deviation between measured and modeled $\mathrm{Im}(y_{11})$ and $\mathrm{Im}(y_{12})$ for the models with and without traps (same bias point as (a)). The inclusion of traps in the model clearly reduces the deviations.

be explained by the equations of the two quantities. Starting with its general definition [39], h_{21} can be written as

$$h_{21} = \frac{|y_{21}|}{|y_{11}|} \approx \frac{|g_\mathrm{m}|}{\left|\mathrm{j}\omega\left(C_\mathrm{gs} + C_\mathrm{gs,p} + C_\mathrm{gd} + C_\mathrm{gd,p}\right)\right|}, \tag{9}$$

where all four capacitances denote the total frequency-dependent capacitances from (2) and (4), e.g. $C_\mathrm{gd} = C_\mathrm{gd,i} - C_\mathrm{gd,\omega} \times \ln(\omega/\omega_0)$. In the approximation in (9), typically, the real parts of y_{11}, which are important for e.g. U and k, are significantly smaller than the imaginary parts. Without the trap contribution, in a Bode plot, (9) would yield the typical -20 dB/decade roll-off. When the effects of traps are included, logarithmic frequency dependences are added to both the numerator (decrease towards lower frequencies) and to the denominator (increase towards lower frequencies), which causes the lower slope for h_{21} in Fig. 5. The same applies for MSG, which, again, starting from its general definition [39], can be written as

$$\mathrm{MSG} = \frac{|y_{21}|}{|y_{12}|} \approx \frac{|g_\mathrm{m}|}{\left|\mathrm{j}\omega(C_\mathrm{gd} + C_\mathrm{gd,p})\right|}. \tag{10}$$

As in (9), both capacitances include the frequency dependences from (2) and (4) and as for U, the deviations of the slopes of h_{21} and MSG can be observed in examples in literature as well [1, 14]. MSG is an important metric for the design of small-signal amplifiers and the dispersion in the transistor MSG is directly translated to the amplifier. Since the dispersion in MSG is much smaller than in U, the resulting penalty in the amplifier should be acceptable as long as the design is not dependent on single decibels of gain. In any case, it is important to be aware of the dispersion, so that the amplifier design will not be based on an idealized and thus incorrect transistor gain.

7. Physical origin of trap model parameters

In the following, the trap density N_t will be calculated from the measured admittances to verify that the small-signal parameters, which model the oxide traps, are physically meaningful. The approach to deriving an expression for N_t is indicated in Fig. 1. At a frequency ω, the part of the distributed network, which responds at frequencies up to this ω, can be seen as an admittance $Y(\omega)$. The change $\Delta Y(\omega)$ for a change $\Delta\omega$ can then be expressed as $\Delta Y(\omega) = Y(\omega) - Y(\omega - \Delta\omega)$, where $Y(\omega - \Delta\omega)$ is the admittance corresponding to an incrementally lower ω. The trap density N_t, as detailed in [23], in included in this model as $\Delta C_\mathrm{t} = q^2 N_\mathrm{t} \Delta x$, where q is the elemental charge, and the ΔC_t are the incremental parts of the trap population within a thickness Δx in the gate oxide. The time constants $\tau = \Delta C_\mathrm{t}/\Delta G_\mathrm{t}$ associated with each ΔC_t are modeled by adding corresponding conductances ΔG_t in series with the incremental capacitances. At this stage, the ΔG_t and thus the different τ can still be modeled independently of the position Δx in the oxide. To be able to calculate N_t from the measured admittances, the angular frequency ω, and thus the time constants τ, are related to Δx via elastic tunneling as $\Delta x = -\lambda \Delta \ln(\omega)$, where $\lambda = 0.13$ nm is the tunneling attenuation length according to the WKB approximation as in Section 3. As discussed in Section 3 as well, elastic tunneling should be a viable approximation at gigahertz frequencies. With this, a change in the admittance $Y(\omega)$ of the distributed RC network in Fig. 1 (y_{11} or y_{12} in the small-signal model) for a change in the angular frequency ω can be expressed as

$$\frac{\Delta Y(\omega)}{\Delta \ln(\omega)} = \frac{1 + \mathrm{j}}{2}\left(\omega q^2 \lambda N_\mathrm{t} L_\mathrm{G} W_\mathrm{G}\right), \tag{11}$$

where the term $\lambda\, \Delta\ln(\omega)\, Y(\omega)$, which appears during the derivation, is sufficiently small to be dropped from the expression. In (11), L_G and W_G are the gate length and width of the transistor, respectively. Examples for N_t as calculated from the measured y_{12} are presented in Fig. 7 alongside a calculation of N_t based on the g_m-f dispersion as in [7]. The values for N_t calculated from y_{11} exhibited a similar behavior as the ones calculated from y_{12}.

For the calculations according to (11), first, the elements surrounding y_{11} and y_{12} in the small-signal model in Fig. 2 have to be subtracted until only the contribution of the y_gd or y_gs remains. The resulting uncertainties, indicated by the shaded areas in Fig. 7, reveal that an exact quantitative resolution of N_t, e.g. as a function of the gate voltage and thus of energy, is limited by the measurement accuracy of the s parameters and the subsequent deconvolution of the y-parameters. For the largest parts of the bias range, however, the average values for N_t calculated from the different components are within an order of magnitude of each other and they are in good agreement with results

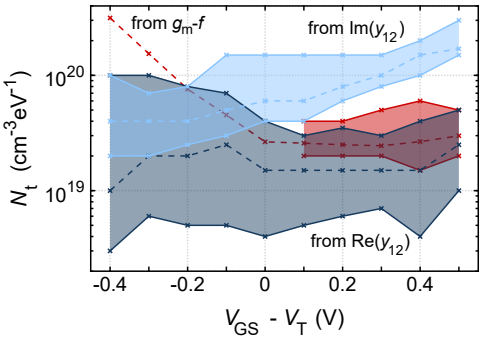


Figure 7: N_t as a function of V_{GS}. Dashed lines indicate the average values for N_t in approximately the first nanometer of the gate oxide from the channel. The shaded areas reflect the uncertainty due to the measurement accuracy. N_t from g_m-f below $V_{GS} - V_T = 0.1$ V was calculated from the modeled y_1 due to strong II and BTBT at these bias points.

that both we and others have observed with different characterization techniques [2, 40–43]. The results are thus sufficiently accurate to verify the physical origin of the model assumptions.

8. Conclusions

The effects of traps related to the gate oxide are clearly discernable in the RF y-parameters of III-V MOSFETs. The inclusion of traps in y_{11}, y_{12}, and y_{21} is thus essential to accurately describe a small-signal model for these devices. Although the resulting effects on design parameters such as h_{21} and MSG are small, the effects have to be identified in the first place to be able to draw this conclusion with certainty. The calculation of the trap density N_t based on the presented model can be limited by the measurement accuracy, but it yields values for N_t, which are in agreement with values calculated by other techniques. This demonstrates that the model parameters reasonably reflect physical assumptions.

Acknowledgements

We thank J. Svensson for growing the nanowires for the vertical nanowire MOSFETs, F. Lindelöw and J. Mo for providing us with samples of lateral nanowire MOSFETs and of planar MOSFETs, respectively, and S. Andrić for helpful discussions. This work was supported in part by the Swedish Foundation for Strategic Research and the Swedish Research Council, and in part by the European Union through the H2020 Program INSIGHT under Grant 688784.

Appendix

Here we provide the complete expressions for the y-parameters for the model in Fig. 2. For convenience, we repeat the expressions for all parameters, which enter into the y-parameters:

$$
\begin{aligned}
y_{gx} &= \omega \times g_{gx,\omega} + j\omega\left(C_{gx,i} - C_{gx,\omega}\ln(\omega/\omega_0)\right),\\
y_{gx,p} &= \omega \times g_{gxp,\omega} + j\omega\left(C_{gxp,0} - C_{gxp,\omega}\ln(\omega/\omega_0)\right)\\
&\quad + g_{gx,l},\\
g_m &= g_{m,i}\left[1 + \gamma_1\ln(\omega/\omega_0) + j\alpha\left(1 + \gamma_2\ln(\omega/\omega_0)\right)\right]\\
&\quad - j\omega\left(C_m - C_{m,\omega}\ln(\omega/\omega_0)\right),\\
g_{ik} &= \frac{g_{k0}}{1 + j\omega\tau_i},\\
R_i &= \frac{1}{1.4\,g_{m,i}} \quad \text{and} \quad R_j = \frac{1}{1.4\,g_{m,i}\times C_{gd,i}/C_{gs,i}}.
\end{aligned}
$$

With these, the intrinsic y-parameters (i.e. after subtracting the resistances R_s, R_d, and R_g) are

$$
\begin{aligned}
y_{11} &= \left.\frac{i_1}{v_1}\right|_{v_2=0} = y_{gs,p} + \frac{y_{gs}}{1 + y_{gs}R_i} - y_{12},\\
y_{12} &= \left.\frac{i_1}{v_2}\right|_{v_1=0} = -y_{gd,p} - \frac{y_{gd}}{1 + y_{gd}R_j},\\
y_{21} &= \left.\frac{i_2}{v_1}\right|_{v_2=0} = \frac{g_m}{1 + y_{gs}R_i} + y_{12} - g_{i1} - \frac{g_{i2}}{1 + y_{gs}R_i},\\
y_{22} &= \left.\frac{i_2}{v_2}\right|_{v_1=0} = g_{ds} + j\omega C_{sd} - y_{12} + g_{i1},
\end{aligned}
$$

References

[1] A. Tessmann, A. Leuther, F. Heinz, F. Bernhardt, L. John, H. Massler, L. Czornomaz, T. Merkle, 20-nm $In_{0.8}Ga_{0.2}As$ MOSHEMT MMIC technology on silicon, IEEE Journal of Solid-State Circuits 54 (9) (2019) 2411–2418. doi:10.1109/JSSC.2019.2915161.

[2] J. Franco, V. Putcha, A. Vais, S. Sioncke, N. Waldron, D. Zhou, G. Rzepa, P. J. Roussel, G. Groeseneken, M. Heyns, N. Collaert, D. Linten, T. Grasser, B. Kaczer, Characterization of oxide defects in InGaAs MOS gate stacks for high-mobility n-channel MOSFETs (invited), in: 2017 IEEE International Electron Devices Meeting (IEDM), 2017, pp. 751–754. doi:10.1109/IEDM.2017.8268347.

[3] B. Kaczer, T. Grasser, J. Roussel, J. Martin-Martinez, R. O'Connor, B. J. O'Sullivan, G. Groeseneken, Ubiquitous relaxation in BTI stressing—new evaluation and insights, in: 2008 IEEE International Reliability Physics Symposium, 2008, pp. 20–27. doi:10.1109/RELPHY.2008.4558858.

[4] R. Engel-Herbert, Y. Hwang, S. Stemmer, Comparison of methods to quantify interface trap densities at dielectric/III-V semiconductor interfaces, Journal of Applied Physics 108 (12) (2010) 124101. doi:10.1063/1.3520431.

[5] H. Riel, L.-E. Wernersson, M. Hong, J. A. del Alamo, III–V compound semiconductor transistors—from planar to nanowire structures, MRS Bulletin 39 (8) (2014) 668–677. doi:10.1557/mrs.2014.137.

[6] V. Putcha, J. Franco, A. Vais, S. Sioncke, B. Kaczer, D. Linten, G. Groeseneken, On the apparent non-arrhenius temperature dependence of charge trapping in III-V/high-k MOS stack, IEEE Transactions on Electron Devices 65 (9) (2018) 3689–3696. doi:10.1109/TED.2018.2851189.

[7] S. Johansson, M. Berg, K. Persson, E. Lind, A high-frequency transconductance method for characterization of high-κ border traps in III-V MOSFETs, IEEE Transactions on Electron Devices 60 (2) (2013) 776–781. doi:10.1109/TED.2012.2231867.

[8] M. Egard, L. Ohlsson, M. Ärlelid, K. Persson, B. M. Borg, F. Lenrick, R. Wallenberg, E. Lind, L.-E. Wernersson, High-frequency performance of self-aligned gate-last surface channel $In_{0.53}Ga_{0.47}As$ MOSFET, IEEE Electron Device Letters 33 (3) (2012) 369–371. doi:10.1109/LED.2011.2181323.

[9] C. B. Zota, G. Roll, L.-E. Wernersson, E. Lind, Radio-frequency characterization of selectively regrown InGaAs lateral nanowire MOSFETs, IEEE Transactions on Electron Devices 61 (12) (2014) 4078–4083. doi:10.1109/TED.2014.2363732.

[10] O. Kilpi, J. Svensson, E. Lind, L. Wernersson, Electrical properties of vertical InAs/InGaAs heterostructure MOSFETs, IEEE Journal of the Electron Devices Society 7 (2018) 70–75. doi:10.1109/JEDS.2018.2878659.

[11] C. B. Zota, F. Lindelöw, L.-E. Wernersson, E. Lind, High-frequency InGaAs tri-gate MOSFETs with f_{max} of 400 GHz, Electronics Letters 52 (22) (2016) 1869–1871. doi:10.1049/el.2016.3108.

[12] G. Roll, J. Mo, E. Lind, S. Johansson, L.-E. Wernersson, Defect evaluation in InGaAs field effect transistors with HfO_2 or Al_2O_3 dielectric, Applied Physics Letters 106 (20) (2015) 203503. doi:10.1063/1.4921483.

[13] M. Hellenbrand, E. Memisevic, J. Svensson, A. Krishnaraja, E. Lind, L. Wernersson, Capacitance measurements in vertical III–V nanowire TFETs, IEEE Electron Device Letters 39 (7) (2018) 943–946. doi:10.1109/LED.2018.2833168.

[14] J. Wu, Y. Fang, B. Markman, H. Y. Tseng, M. J. W. Rodwell, L_g = 30 nm InAs channel MOSFETs exhibiting f_{max} = 410 GHz and f_t = 357 GHz, IEEE Electron Device Letters 39 (4) (2018) 472–475. doi:10.1109/LED.2018.2803786.

[15] C. B. Zota, C. Convertino, M. Sousa, D. Caimi, K. Moselund, L. Czornomaz, High-frequency quantum well InGaAs-on-Si MOSFETs with scaled gate lengths, IEEE Electron Device Letters 40 (4) (2019) 538–541. doi:10.1109/LED.2019.2902519.

[16] K. Xiong, J. Robertson, M. C. Gibson, S. J. Clark, Defect energy levels in HfO_2 high-dielectric-constant gate oxide, Applied Physics Letters 87 (18) (2005) 183505. doi:10.1063/1.2119425.

[17] J. L. Gavartin, D. Muñoz Ramo, A. L. Shluger, G. Bersuker, B. H. Lee, Negative oxygen vacancies in HfO_2 as charge traps in high-k stacks, Applied Physics Letters 89 (8) (2006) 082908. doi:10.1063/1.2236466.

[18] J. Robertson, Y. Guo, L. Lin, Defect state passivation at III-V oxide interfaces for complementary metal–oxide–semiconductor devices, Journal of Applied Physics 117 (11) (2015) 112806. doi:10.1063/1.4913832.

[19] G. Greene-Diniz, K. J. Kuhn, P. K. Hurley, J. C. Greer, First principles modeling of defects in the $Al_2O_3/In_{0.53}Ga_{0.47}As$ system, Journal of Applied Physics 121 (7) (2017) 075703. doi:10.1063/1.4975033.

[20] F. P. Heiman, G. Warfield, The effects of oxide traps on the MOS capacitance, IEEE Transactions on Electron Devices 12 (4) (1965) 167–178. doi:10.1109/T-ED.1965.15475.

[21] T. Grasser, The capture/emission time map approach to the bias temperature instability, in: T. Grasser (Ed.), Bias Temperature Instability for Devices and Circuits, Springer New York, New York, NY, 2014, pp. 447–481. doi:10.1007/978-1-4614-7909-3_17.

[22] H. Chen, Y. Yuan, B. Yu, J. Ahn, P. C. McIntyre, P. M. Asbeck, M. J. W. Rodwell, Y. Taur, Interface-state modeling of Al_2O_3 –ingaas mos from depletion to inversion, IEEE Transactions on Electron Devices 59 (9) (2012) 2383–2389. doi:10.1109/TED.2012.2205255.

[23] Y. Yuan, L. Wang, B. Yu, B. Shin, J. Ahn, P. C. McIntyre, P. M. Asbeck, M. J. W. Rodwell, Y. Taur, A distributed model for border traps in Al_2O_3-InGaAs MOS devices, IEEE Electron Device Letters 32 (4) (2011) 485–487. doi:10.1109/LED.2011.2105241.

[24] I. Lundström, C. Svensson, Tunneling to traps in insulators 43 (12) (1972) 5045–5047. doi:10.1063/1.1661067.

[25] K. Jansson, E. Lind, L.-E. Wernersson, Ballistic modeling of inas nanowire transistors, Solid-State Electronics 115 (2016) 47 – 53. doi:10.1016/j.sse.2015.10.009.

[26] N. Khuchua, L. Khvedelidze, M. Tigishvili, N. Gorev, E. Privalov, I. Kodzhespirova, Deep-level effects in GaAs microelectronics: a review, Russian Microelectronics 32 (5) (2003) 257–274. doi:10.1023/A:1025528416032.

[27] P. S. Whitney, W. Lee, C. G. Fonstad, Capacitance transient analysis of molecular-beam epitaxial n-$In_{0.53}Ga_{0.47}As$ and n-$In_{0.52}Al_{0.48}$ as, Journal of Vacuum Science & Technology B: Microelectronics Processing and Phenomena 5 (3) (1987) 796–799. doi:10.1116/1.583753.

[28] T. Grasser, Stochastic charge trapping in oxides: From random telegraph noise to bias temperature instabilities, Microelectronics Reliability 52 (1) (2011) 39 – 70. doi:https://doi.org/10.1016/j.microrel.2011.09.002.

[29] G. Rzepa, J. Franco, A. Subirats, M. Jech, A. Chasin, A. Grill, M. Waltl, T. Knobloch, B. Stampfer, T. Chiarella, N. Horiguchi, L. . Ragnarsson, D. Linten, B. Kaczer, T. Grasser, Efficient physical defect model applied to PBTI in high-κ stacks, in: 2017 IEEE International Reliability Physics Symposium (IRPS), 2017, pp. XT–11.1–XT–11.6. doi:10.1109/IRPS.2017.7936425.

[30] T. Grasser, P. . Wagner, H. Reisinger, T. Aichinger, G. Pobegen, M. Nelhiebel, B. Kaczer, Analytic modeling of the bias temperature instability using capture/emission time maps, in: 2011 International Electron Devices Meeting, 2011, pp. 27.4.1–27.4.4. doi:10.1109/IEDM.2011.6131624.

[31] A. S. Babadi, E. Lind, L. E. Wernersson, Modeling of n-InAs metal oxide semiconductor capacitors with high-κ gate dielectric 116 (21) (2014) 214508–1–214508–6. doi:10.1063/1.4903520.

[32] G. Brammertz, A. Alian, D. H. Lin, M. Meuris, M. Caymax, W. . Wang, A combined interface and border trap model for high-mobility substrate metal–oxide–semiconductor devices applied to $In_{0.53}Ga_{0.47}As$ and InP capacitors, IEEE Transactions on Electron Devices 58 (11) (2011) 3890–3897. doi:10.1109/TED.2011.2165725.

[33] H. Reisinger, T. Grasser, W. Gustin, C. Schlünder, The statistical analysis of individual defects constituting NBTI and its implications for modeling DC- and AC-stress, in: 2010 IEEE International Reliability Physics Symposium, 2010, pp. 7–15. doi:10.1109/IRPS.2010.5488858.

[34] W. Liu, FET High-Frequency Properties, Wiley Interscience, 1999, Ch. 6, pp. 371–446.

[35] H. R. P. Roblin, Small- and large-signal AC models for the short-channel MODFET, Cambridge University Press, 2002, Ch. 12, pp. 384–411.

[36] S. Makovejev, S. Olsen, J. Raskin, RF extraction of self-heating effects in FinFETs, IEEE Transactions on Electron Devices 58 (10) (2011) 3335–3341. doi:10.1109/TED.2011.2162333.

[37] M. Isler, K. Schünemann, Impact-ionization effects on the high-frequency behavior of HFETs 52 (3) (2004) 858–863. doi:10.1109/TMTT.2004.823553.

[38] S. Johansson, M. Egard, S. G. Ghalamestani, B. M. Borg, M. Berg, L.-E. Wernersson, E. Lind, RF characterization of vertical InAs nanowire wrap-gate transistors integrated on Si substrates, IEEE Transactions on Microwave Theory and Techniques 59 (10) (2011) 2733–2738. doi:10.1109/TMTT.2011.2163076.

[39] H. R. P. Roblin, MODFET high-frequency performance, Cambridge University Press, 2002, Ch. 17, pp. 567–612.

[40] M. Hellenbrand, O.-P. Kilpi, J. Svensson, E. Lind, L.-E. Wernersson, Low-frequency noise in nanowire and planar III-V MOSFETs, Microelectronic Engineering 215 (2019) 110986. doi:https://doi.org/10.1016/j.mee.2019.110986.

[41] B. Kaczer, J. Franco, P. Weckx, P. Roussel, V. Putcha, E. Bury, M. Simicic, A. Chasin, D. Linten, B. Parvais, F. Catthoor, G. Rzepa, M. Waltl, T. Grasser, A brief overview of gate oxide defect properties and their relation to MOSFET instabilities and device and circuit time-dependent variability, Microelectronics Reliability 81 (2018) 186 – 194. doi:https://doi.org/10.1016/j.microrel.2017.11.022.

[42] E. Caruso, J. Lin, K. F. Burke, K. Cherkaoui, D. Esseni, F. Gity, S. Monaghan, P. Palestri, P. Hurley, L. Selmi, Profiling border-traps by TCAD analysis of multifrequency CV-curves in Al_2O_3/InGaAs stacks, in: 2018 Joint International EUROSOI Workshop and International Conference on Ultimate Integration on Silicon (EUROSOI-ULIS), 2018, pp. 1–4. doi:10.1109/ULIS.2018.8354757.

[43] G. Roll, E. Lind, M. Egard, S. Johansson, L. Ohlsson, L. E. Wernersson, Rf and DC analysis of stressed InGaAs mosfets, IEEE Electron Device Letters 35 (2) (2014) 181–183. doi:10.1109/LED.2013.2295526.

Paper VII

Paper VII

Reproduced, with permission, from

O.-P. Kilpi, M. Hellenbrand, J. Svensson, E. Lind, and L.-E. Wernersson, "Vertical nanowire III–V MOSFETs with improved high-frequency gain," *Electronics Letters*, Article in Press, Apr. 2020, doi: 10.1049/el.2020.0266.

This paper is a postprint of a paper submitted to and accepted for publication in *Electronics Letters* and is subject to Institution of Engineering and Technology Copyright. The copy of record is available at IET Digital Library.

Vertical nanowire III–V MOSFETs with improved high-frequency gain

O.-P. Kilpi✉, M. Hellenbrand, J. Svensson, E. Lind and L.-E. Wernersson

High-frequency performance of vertical InAs/InGaAs heterostructure nanowire MOSFETs on Si is demonstrated for the first time for a gate-last configuration. The device architecture allows highly asymmetric capacitances, which increases the power gain. A device with $L_g = 120$ nm demonstrates $f_T = 120$ GHz, $f_{max} = 130$ GHz and maximum stable gain (MSG) = 14.4 dB at 20 GHz. These metrics demonstrate the state-of-the-art performance of vertical nanowire MOSFETs.

Introduction: Moore's law has been the driving force for the development of electrical devices for the last few decades. Si-based MOSFETs have shown excellent scalability in digital applications and especially III–V MOSFETs have surpassed other transistor architectures in the transconductance g_m [1] and the on-current I_{on} [2, 3]. For high-frequency transistors, the inability to scale high electron mobility transistor (HEMTs) further due to the insufficient gate-barrier has stagnated the development of f_T at 700 GHz [4, 5]. Better scaling III–V MOSFETs could, therefore, surpass HEMTs also in high-frequency applications. Recently, several high-frequency planar MOSFETs have been presented with f_T and f_{max} about 400 GHz [6, 7]. In this Letter, the prospect of vertical III–V nanowires for high-frequency devices is evaluated.

III–V vertical vapour–liquid–solid (VLS) grown nanowires offer an interesting option as they are less limited by the lattice mismatch; therefore, they can be easily integrated on Si. The VLS growth further allows band-gap engineering along the electron transport direction, which enables improved intrinsic voltage gain (g_m/g_d) and higher breakdown voltage without substantially deteriorating other performance metrics [8, 9]. This kind of transistor architecture has a large potential for scaling due to the excellent electrostatics provided by the gate-all-around structure. Recently, DC operation of well scalable vertical nanowire MOSFETs has been presented [8]. In this Letter, the high-frequency performance of a similar structure is presented.

Device structure and fabrication: Schematics and SEM images of the finished device are shown in Fig. 1. The vertical nanowire MOSFET processes can generally be divided into gate-last and gate-first processes. The difference between the two processes is highlighted in Fig. 1*a*. Gate-first devices generate thin nanowire contacts and leave ungated regions at the top, which will increase the access resistance. Gate-last devices have addressed this problem by forming a recessed gate and adding metal on the nanowire sidewalls to reduce the access resistance, therefore decoupling the relation between the contact resistance and channel diameter. In this Letter, the RF performance of the gate-last MOSFETs is evaluated.

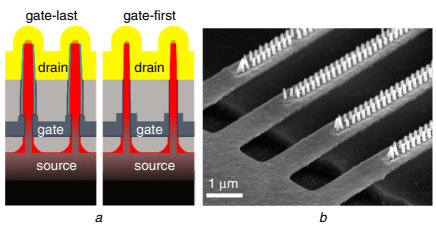


Fig. 1 *Illustrations of the vertical nanowire MOSFETs structure*
a Schematic illustration of finalised device
b SEM picture after gate-finger fabrication. Isolation mesa and air bridge allow reduction in the parasitic capacitance

The fabricated gate-last MOSFETs are based on $InAs/In_{0.4}Ga_{0.6}As$ heterostructure nanowires fabricated on a highly resistive Si{111} substrate. The nanowire consists of three parts: (i) a 100-nm-long highly doped InAs bottom segment; (ii) a 100-nm-long segment with unintentionally doped grading from InAs to InGaAs; and (iii) a 300-nm-long highly doped InGaAs top segment. The gate is connected to the InAs and the graded segment. A detailed DC evaluation of similar nanowire structures including gate-length scaling has been shown earlier [9]. The fabrication closely follows the process flow previously used up to the gate-metal deposition; here the process is adjusted by reducing the gate-drain and gate-source capacitance. After the 60-nm-thick W gate-metal is deposited, finger gate structures are patterned using deep ultraviolet and electron-beam lithography patterning. The finger gate process was presented in [10]. The patterned fingers are dry-etched and the isolation mesa is wet-etched, forming a structure as shown in Fig. 1*b*. In the figure, the gate pad is separated from the source by an air bridge between the source and gate pad. The device is finalised by depositing a second spacer, contact vias, and contact pads.

Results: In Fig. 2, transfer and output characteristics of a device with 180 nanowires are shown. The channel diameter is 30 nm and $L_g = 120$ nm. The device saturates well, which leads to a good intrinsic voltage gain $g_m/g_d = 12$. The device $R_{on} = 780\ \Omega\mu m$ and $g_m = 1.1$ mS/µm correlate well with the state-of-the-art vertical III–V MOSFETs with the corresponding L_g, although here, the doped source causes a degraded subthreshold performance (subthreshold swing = 380 mV/dec) due to a dopant memory effect during growth.

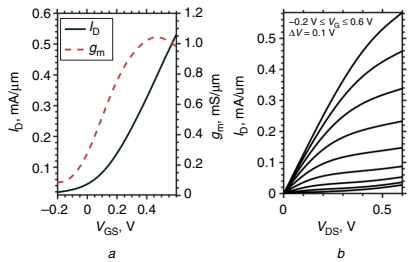


Fig. 2 *DC characterisation of the vertical nanowire MOSFET with 180 nanowires, average diameter 30 nm and $L_g = 120$ nm*
a Transfer characteristic measured at $V_{DD} = 0.5$ V
b Output characteristic

In Fig. 3, the high-frequency performance of the MOSFET and the corresponding small-signal model are presented. *S*-parameters were measured from 10 MHz to 67 GHz using an Agilent E8361A vector network analyser. Calibration was performed by using off-chip two-port line reflect reflect match calibration and on-chip open/short de-embedding. The implemented small-signal model results in a good fit to the experimental data, as shown in Fig. 3*a*, where measured and modelled forward current gain h_{21}, the unilateral power gain U, and the maximum stable/available gain MSG/MAG are presented. Based on the model, $f_t = 122$ GHz and $f_{max} = 131$ GHz can be extrapolated. The device achieves a high gain of MSG = 14.5 dB at 20 GHz, which is comparable to planar RF nanowire MOSFETs with $f_{max} = 400$ GHz [11]. The lower f_{max} in the vertical MOSFET is attributed mainly to the larger gate resistance R_g. This conclusion can be drawn from the approximate equations, which relate the small-signal parameters from Fig. 3*a* to the cut-off frequency f_T, the maximum oscillation frequency f_{max}, and the maximum stable gain (MSG)

$$f_{max} \simeq \sqrt{f_T/\left(8\pi R_g C_{gd}\right)}, \tag{1}$$

$$f_T \simeq g_{m,i}/\left[2\pi\left(C_{gd} + C_{gs,i} + C_{gs,p}\right)\right], \tag{2}$$

$$|\mathrm{MSG}| \simeq g_{m,i}/\left(\omega C_{gd}\right). \tag{3}$$

In the on-state of the transistor, the parasitic gate-drain capacitance dominates over the intrinsic one, so that the model in Fig. 3*b* does not differentiate between the two and C_{gd} mainly consists of the parasitic contribution. Owing to the asymmetric gate-last process of the vertical MOSFETs, a reduction of this parasitic C_{gd} to about 8 fF was possible without deteriorating other parameters. It is likely that the main restriction for $f_{max}$is the gate resistance R_g. Reduction of R_g would require further process development.

For the gate-source capacitance, both the parasitic $C_{gs,p}$ and the intrinsic $C_{gs,i}$ can be determined from the small-signal model. Equation (2)

indicates that $C_{gs,p} = 15$ fF is a major limiting factor for f_T. $C_{gs,p}$ mainly originates from the plate-capacitor-like structure and fringing capacitance of the gate fingers, compare Fig. 1*b*. Owing to the highly doped shell around the bottom of the nanowires (cf. Fig. 1*a*) it is possible to reduce $C_{gs,p}$ by increasing the distance of the gate fingers from the substrate without largely increasing the access resistance. The intrinsic $C_{gs,i}$ can be scaled by reducing the gate length. Gate-length down to 25 nm has already been demonstrated for vertical nanowire MOSFETs with a 100-nm-thick bottom spacer [8]. Thus, by straightforward dimension scaling, f_T and f_{max} can be drastically increased in further transistor generations.

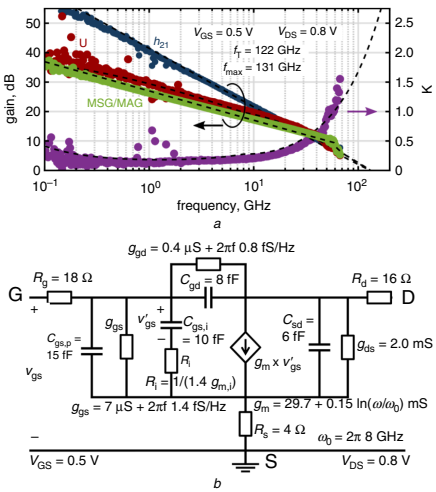


Fig. 3 *RF characterisation of the vertical nanowire MOSFET with 180 nanowires, average diameter 30 nm and L_g = 120*

a Measured current gain h_{21}, maximum stable/available gain MSG/MAG, and stability factor *K*. Dashed line describe fitted small-signal model
b Small-signal model with the fitted values for the measurement presented in (a)

Table 1 benchmarks the performance of this work versus other III–V RF MOSFETs on Si [6] and versus 28 nm silicon on insulator (SOI) RF MOSFETs [7] with comparable gate-lengths. Although the device structure of the asymmetric MOSFETs presented here is not fully optimised yet, its RF performance is comparable to that of the other technologies. As discussed, further improvement is mostly a matter of scaling the structure to reduce the capacitances.

Table 1: Comparison of RF MOSFETs on Si with similar gate-lengths

	L_g, nm	g_m, mS/µm	f_T, GHz	f_{max}, GHz
this work	120	1.1	125	130
SOI 28 nm [7]	90	0.9	110	103
SOI 28 nm [7]	150	0.75	70	80
III–V on Si [6]	100	1.3	140	170
III–V on Si [6]	150	1	100	120

In Fig. 4*a*, C_{gs} and C_{gd} are shown as a function of V_{GS} at different drive voltages ($V_{dd} = 0.2$, 0.5 and 0.8 V). For C_{gd}, it is straightforward to extract the parasitic capacitance from the off-state, $C_{gd,p} \simeq 8$ fF. Furthermore, the figure confirms that for high V_{dd}, C_{gd} is barely affected by an intrinsic contribution. C_{gs} does not saturate for the lowest V_{GS} measured, which is likely because of the high channel doping due to the memory effect as mentioned previously. However, based on the trend in Fig. 4*a*, $C_{gs,p} \simeq 15$ fF and $C_{gs,i} \simeq 10$ fF can be estimated for the small-signal model in Fig. 3*b*. Fig. 4*b* presents g_m derived from *S*-parameter measurements at 50 MHz for the same bias points as the capacitances in Fig. 4*a*. The extracted g_m corresponds well with the DC measurements in Fig. 2 and only minor improvement in g_m is observed when increasing V_{dd} over 0.5 V.

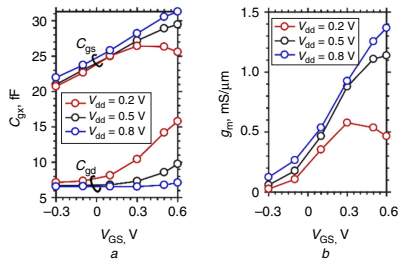


Fig. 4 *C_{gs} and C_{gd} bias dependence*

a Transistor with 180 nanowires, average diameter 30 nm and $L_g = 120$ nm. g_m bias dependence
b Measured at 50 MHz. g_m corresponds well with the earlier presented DC measurements

In the small-signal model in Fig. 3, the intrinsic transconductance ($g_{m,i}$) is the constant part of the total g_m. At frequencies below a certain ω_0, g_m is reduced due to the presence of border traps in the gate oxide [12], which gives rise to the logarithmic term in the expression for g_m. ω_0 is the frequency at which border traps cannot respond to the AC signal anymore. The intrinsic resistance $R_i = 1/(1.4\ g_{m,i})$ in series with $C_{gs,i}$ takes into account the delay, which charge carriers experience when moving inside the channel. Physically, a corresponding intrinsic resistance should be placed in series with the intrinsic gate-drain capacitance, but since C_{gd} is dominated by the parasitic component, as described earlier, the resistance is disregarded.

Besides the g_m-f dispersion, border traps also give rise to the frequency dependence of the conductances g_{gd} and g_{gs} in the small-signal model, which is based on modelling the border traps by distributed RC networks as elaborated in [13]. The constant part of g_{gd} and g_{gs} models DC gate leakage. All components in the model can be determined from the real and imaginary part of the admittance parameters at different frequencies and all are necessary for accurate modelling of the RF response of the transistors. Fig. 3*a* demonstrates excellent agreement between measured and modelled values.

Conclusion: Vertical nanowire gate-last high-frequency MOSFETs on Si have been demonstrated. The devices show performance comparable to or higher than other Si and III–V MOSFETs on Si with similar gate-length. The devices are fabricated by using a scalable gate-last process, which has been shown to be scalable down to 25 nm.

Acknowledgments: This work was supported in part by the Swedish Research Council, the Knut and Alice Wallenberg Foundation, the Swedish Foundation for Strategic Research, and the European Union H2020 program INSIGHT (grant no. 688784).

© The Institution of Engineering and Technology 2020
Submitted: *27 March 2020*
doi: 10.1049/el.2020.0266
One or more of the Figures in this Letter are available in colour online.

O.-P. Kilpi, M. Hellenbrand, J. Svensson, E. Lind and L.-E. Wernersson (*Department of Electrical and Information Technology, Lund University, Box 118 S-22100, Lund, Sweden*)

✉ E-mail: olli-pekka.kilpi@eit.lth.se

References

1 Lin, J., Cai, X., Wu, Y., *et al.*: 'Record maximum transconductance of 3.45 mS/µm for III–V FETs', *IEEE Electron Device Lett.*, 2016, **37**, (4), pp. 381–384
2 Zota, C.B., Lindelow, F., Wernersson, L., *et al.*: 'InGaAs tri-gate MOSFETs with record on-current'. 2016 IEEE Int. Electron Devices Meeting (IEDM), San Francisco, CA, USA, 2016, pp. 3.2. 1–3.2. 4
3 Chang, S., Li, X., Oxland, R., *et al.*: 'InAs N-MOSFETs with record performance of $I_{on} = 600$ µA/µm at $I_{off} = 100$ nA/µm ($V_d = 0.5$ V)'. IEDM Technical Digest, Washington, DC, USA, 2013, p. 16.1

4 Del Alamo, J.A.: 'Nanometre-scale electronics with III–V compound semiconductors', *Nature*, 2011, **479**, (7373), p. 317
5 Lind, E.: 'High frequency III–V nanowire MOSFETs', *Semicond. Sci. Technol.*, 2016, **31**, (9), p. 093005
6 Zota, C., Convertino, C., Deshpande, V., *et al.*: 'InGaAs-on-insulator MOSFETs featuring scaled logic devices and record RF performance'. 2018 IEEE Symp. on VLSI Technology, Honolulu, HI, USA, 2018, pp. 165–166
7 Kazemi Esfeh, B., Kilchytska, V., Barral, V., *et al.*: 'Assessment of 28 nm UTBB FD-SOI technology platform for RF applications: figures of merit and effect of parasitic elements', *Solid-State Electron.*, 2016, **117**, pp. 130–137
8 Kilpi, O., Svensson, J., and Wernersson, L.: 'Sub-100-nm gate-length scaling of vertical InAs/InGaAs nanowire MOSFETs on Si'. 2017 IEEE Int. Electron Devices Meeting (IEDM), San Francisco, CA, USA, 2017, pp. 17.3. 1–17.3. 4
9 Kilpi, O.-P., Svensson, J., Wu, J., *et al.*: 'Vertical InAs/InGaAs heterostructure metal–oxide–semiconductor field-effect transistors on Si', *Nano Lett.*, 2017, **17**, (10), pp. 6006–6010
10 Johansson, S., Memisevic, E., Wernersson, L., *et al.*: 'High-frequency gate-all-around vertical InAs nanowire MOSFETs on Si substrates', *IEEE Electron Device Lett.*, 2014, **35**, (5), pp. 518–520
11 Zota, C.B., Lindelöw, F., Wernersson, L.-E., *et al.*: 'High-frequency InGaAs tri-gate MOSFETs with f_{max} of 400 GHz', *Electron. Lett.*, 2016, **52**, (22), pp. 1869–1871
12 Johansson, S., Berg, M., Persson, K., *et al.*: 'A high-frequency transconductance method for characterization of high-k border traps in III–V MOSFETs', *IEEE Trans. Electron Devices*, 2012, **60**, (2), pp. 776–781
13 Yuan, Y., Wang, L., Yu, B., *et al.*: 'A distributed model for border traps in Al_2O_3–InGaAs MOS devices', *IEEE Electron Device Lett..*, 2011, **32**, (4), pp. 485–487

Paper VIII

Paper VIII

Reproduced, with permission, from

M. Hellenbrand, E. Memisevic, J. Svensson, A. Krishnaraja, E. Lind, and L.-E. Wernersson, "Capacitance Measurements in Vertical III–V Nanowire TFETs," *IEEE Electron Device Letters*, vol. 39, no. 7, pp. 943–946, Jul. 2018, doi: 10.1109/LED.2018.2833168.

© 2018 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission.

Capacitance Measurements in Vertical III–V Nanowire TFETs

Markus Hellenbrand, Elvedin Memisevic, Johannes Svensson, Abinaya Krishnaraja, Erik Lind, and Lars-Erik Wernersson

***Abstract*— By measuring scattering parameters over a wide range of bias points, we study the intrinsic gate capacitance as well as the charge partitioning of vertical nanowire tunnel field-effect transistors (TFETs). The gate-to-drain capacitance C_{gs} is found to largely dominate the ON-state of TFETs, whereas the gate-to-source capacitance C_{gs} is sufficiently small to be completely dominated by the parasitic components. This indicates that the tunnel junction on the source side almost completely decouples the channel charge from the small-signal variation in the source, while the absence of a tunnel junction on the drain side allows the channel charge to follow the drain small-signal variation much more directly.**

***Index Terms*— Vertical nanowires, III-V, TFET, small-signal model, intrinsic capacitance, RF, C_{gd}, C_{gs}.**

I. Introduction

In RECENT years, the performance of tunnel field-effect transistors (TFETs) has made significant progress and devices with not only inverse subthreshold slopes S below 60 mV/decade but also with on-currents approaching technically relevant current levels have been realized [1]–[4]. With the intention of realizing TFET-based circuit implementations, not only the investigation of the DC properties of TFETs is of interest, but more and more so the investigation of their high-frequency characteristics. Here, we experimentally investigate the capacitance and charge partitioning of the gate-to-drain capacitance C_{gd} and the gate-to-source capacitance C_{gs}. Both were extensively simulated in the past [5]–[9] and the unanimous conclusion is that in TFETs – in contrast to metal-oxide-semiconductor field-effect transistors (MOSFETs) without a tunnel junction – C_{gd} assumes much larger values than C_{gs}, which constitutes a large Miller capacitance adversely affecting circuit parameters such as stability, switching energy, and propagation delay [8], [10]. Here, we use two-port scattering parameter (s-parameter) measurements and small-signal modeling of the measured values for a large range of bias points to determine experimentally and with high accuracy the intrinsic transistor capacitances in vertical III-V nanowire TFETs. The only other experimental results on TFET capacitances up to date, to our knowledge, are s-parameter measurements for fewer bias points, still augmented by simulations [11], and two-terminal capacitance-voltage measurements on Si TFETs [12].

Manuscript received April 23, 2018; accepted April 30, 2018. Date of publication May 4, 2018; date of current version June 26, 2018. This work was supported in part by the Swedish Foundation for Strategic Research, in part by the Swedish Research Council, and in part by the European Union Seventh Framework Program E2SWITCH under Grant 619509. The review of this letter was arranged by Editor M. Passlack. *(Corresponding author: Markus Hellenbrand.)*

The authors are with the Department of Electrical and Information Technology, Lund University, SE-221 00 Lund, Sweden (e-mail: markus.hellenbrand@eit.lth.se).

Color versions of one or more of the figures in this letter are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/LED.2018.2833168

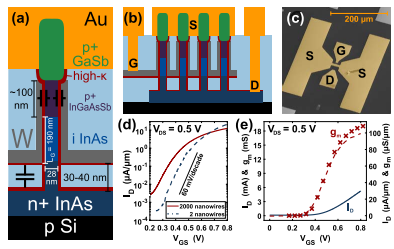


Fig. 1. (a) Schematic representation of the measured TFETs (not to scale), zoomed in on a single nanowire to clarify the details of the structure. The schematic capacitors indicate the largest parasitic contributions. (b) Schematic representation of the RF array layout. (c) SEM image of the top metal pads (colored in yellow) to contact the transistors in the probe station. (d) Logarithmic transfer curve of an RF transistor and a reference transistor with two nanowires. Despite an array of 2000 nanowires, minimum slopes of 75 mV/decade were achieved. (e) DC transfer curve with the extrinsic RF g_m (crosses) added on top of the DC g_m. Excellent agreement between the two supports the accuracy of the small-signal parameter extraction. Normalization to the total circumference of the InAs channel.

II. Device Structure

The fabrication of the sample was based on the processing scheme that is used for our devices, which consistently reach values of S below 60 mV/decade. The processing is reported in detail in [1] and here we only focus on the resulting device structure, schematically depicted in Fig. 1(a) and (b). Vertical nanowires were grown by metal organic vapor phase epitaxy from Au seed particles on a highly n-doped InAs layer integrated on a high resistivity Si substrate. The tunnel junction was realized by a p+ $In_{0.32}Ga_{0.68}As_{0.72}Sb_{0.28}$/i-InAs interface. For optimized top and bottom contacts, p+ GaSb and n+ InAs segments followed before and after, as depicted in Fig. 1(a). After growth, the InAs segments were digitally etched to final diameters of 28 nm. The gate stack was formed by a high-κ gate oxide (five cycles Al_2O_3 and 36 cycles HfO_2, effective oxide thickness $\approx$ 1.4 nm) and 40 nm W gate metal, which was separated from the bottom InAs layer by the gate oxide and an evaporated spacer of 40 nm SiO_x. The spacer between

0741-3106 © 2018 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission.
See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

gate and top metal was formed by a photoresist, and gate and drain were contacted by via holes (Fig. 1(b)).

To compensate for the low currents in TFETs and to enable measurable RF gain, each RF transistor comprised an array of up to 2000 nanowires, which were arranged on mesas (Fig. 1(b)) to reduce parasitic capacitances. Due to processing and growth variations, not all of the nanowires within a single transistor exhibit the same threshold voltage V_T, which deteriorates the inverse subthreshold slope of the overall device [13]. Still, even for devices with up to 2000 nanowires, S as low as 75 mV/decade at $V_{DS} = 0.5$ V could be achieved (Fig. 1(d)), although for most of the RF transistors, S was between 140 and 200 mV/decade. Reference transistors on the same sample with only two nanowires achieved values of S below 60 mV/decade (Fig. 1(d)). Besides the representative transfer curve for the subthreshold characteristics in Fig. 1(d), Fig. 1(e) also presents an example of the linear transfer characteristics, which have the extrinsic transconductance values, determined by the small-signal model, added on top of the DC curve. Good agreement between the two supports the quality of the small-signal fits and typical peak transconductance values for the measured devices were about 113 µS/µm at $V_{DS} = 0.5$ V.

III. Measurement Technique and Modeling

Prior to measurement, the setup was calibrated off chip by a load-reflect-reflect-match algorithm and open and short on-chip de-embedding structures were used to remove the influence of the top metal pads shown in Fig. 1(c). s-parameters were measured for several transistors from 10 MHz to 15 GHz at a constant drain-source voltage $V_{DS} = 0.5$ V for gate-source voltages V_{GS} ranging from the off-state to the on-state, and at a constant $V_{GS} \approx V_T$ for V_{DS} from 0 to 0.5 V. R_s and R_d, derived from transistor off-state measurements, were subtracted from the de-embedded s-parameters (inset Fig. 2(a)) and analytical expressions of the admittance parameters (y-parameters) of the model in Fig. 2(a) were fitted to the measured data. Here, R_g, R_d, and R_s are the gate, drain, and source resistances, respectively, $g_{m,i}$ is the intrinsic transconductance and g_{ds} the output conductance. The C_{ij} are the capacitances between the respective terminals i and j and, in Fig. 2, contain both parasitic and intrinsic components. g_{gd} takes into account losses from gate oxide defects, causing the deviation from the slope of -20 dB/decade for the unilateral power gain U in Fig. 2(c). g_{gd} could also include gate leakage, which, however, was found to be negligible, just as a respective element g_{gs} in parallel with C_{gs}, or an effect of the gate oxide defects on $g_{m,i}$. Excellent agreement between measured and modeled data in Fig. 2 proves the accurate determination of the small-signal parameters; representative values are provided in Fig. 2(a).

IV. Results and Analysis

The analysis here focuses on the gate capacitances $C_{gd} = -\partial Q_G/\partial V_d$ and $C_{gs} = -\partial Q_G/\partial V_s$, which describe the change of charge Q_G on the gate with respect to small-signal excitations of the drain and the source, respectively, and thus, due to charge neutrality requirements, the change of charge in the channel associated with either of the two contacts. The constant levels of C_{gd} for low V_{GS} in the inset

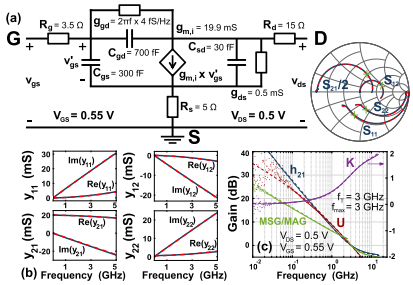


Fig. 2. (a) Small-signal model with representative extracted parameters. The Smith chart shows the measured, de-embedded (solid, blue) and the modelled (broken, red) s-parameters. Green crosses indicate f_T. (b) Measured (solid, blue) and modeled (broken, red) y-parameters, modeled values as in (a). (c) Forward current gain h_{21}, unitary power gain U, maximum stable and maximum available gain MSG and MAG, and stability factor K. Dots are measured data, broken lines are modelled with the values in (a). For all figures, the modelled curves agree well with the measured ones.

of Fig. 3(a) and for high V_{DS} in the inset of Fig. 3(b) can be identified as parasitic components of the total C_{gd}, which originate in the plate-capacitor-like structure between the gate metal pad and the InAs bottom layer as indicated in Fig. 1(a). The two different levels for C_{gd} originate in two different pad sizes of 23×20 µm and 23×26 µm, and calculating the plate capacitances for both yields values of 500 fF and 650 fF, which is very close to the measured parasitics in Fig. 3. For C_{gs}, the main parasitic contribution originates in the overlap of the gate metal with the InGaAsSb source segment of the nanowire, as indicated in Fig. 1(a). Analytical calculation of only this geometrical cylinder capacitance yields values between 570 fF and 700 fF, depending on the exact length of the overlap. This is clearly larger than the constant C_{gs} levels in the insets in Fig. 3. In a more accurate calculation, these values would be lowered by the semiconductor capacitance of the source, the determination of which is difficult, however, and goes beyond the scope of this study, so that the approximate geometrical values should suffice to explain the origin of the parasitic C_{gs}.

Given the physical structure of the transistors, schematically depicted in Fig. 1, the parasitic capacitance components can be assumed to be in parallel with the intrinsic components, so that the parasitic contributions can just be subtracted from the total values. The result of this is presented in Fig. 3, where the change of the intrinsic C_{gd} with V_{GS} and V_{DS} becomes apparent, while the intrinsic C_{gs} remains close to zero. The latter means that in these devices the intrinsic C_{gs}, even far into the on-state, is much smaller than the parasitics, which indicates that at the measured frequencies, the tunnel junction between source and channel almost completely decouples the channel charge from the source small-signal voltage. This means that the tunnel transmission T_{BTB} is low, which can be verified by the estimation $T_{BTB} \approx \exp\left(-4\sqrt{2m^*}E_G^{3/2}/(3q\hbar\xi)\right)$ [14], where q is the elemental charge and $\hbar$ the reduced Planck constant. Inserting parameters in accordance with [1] – the effective tunneling mass $m^* \approx 0.04m_0$, the bandgap $E_G \approx 0.44$ eV, and the electric field at the junction $\xi \approx 1$ MV/cm – results in $T_{BTB} \approx 0.02$ in the on-state. As a self-consistent further verification, this value for T_{BTB} can be used to

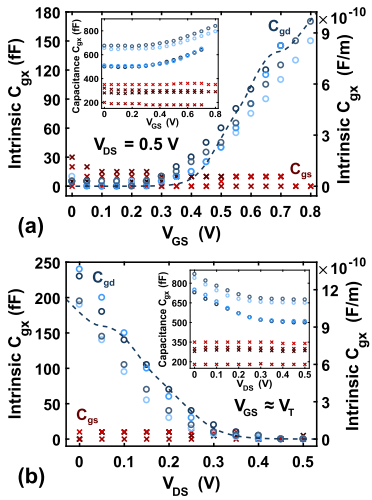


Fig. 3. Extracted capacitances for four different devices. C_{gd} (blue circles) and C_{gs} (red crosses) as functions of V_{GS} for $V_{DS} = 0.5$ V (a) and of V_{DS} for $V_{GS} \approx V_T$ (b). The insets show the extracted capacitances with the parasitics included. The intrinsic C_{gd} emerges from the parasitic floor for high V_{GS}/low V_{DS}, whereas the intrinsic C_{gs} is so small that it is dominated by parasitics over the whole measurement range. The broken lines in both figures are analytical calculations of C_{gd}.

calculate the on-current and with the expression for the 1D current from [14] the result is very close to the measured values in Fig. 1(d)/(e).

Due to the low tunnel transmission, the majority of the channel charge is supplied from the drain side at all times, so that the change of C_{gd} with V_{GS}, as shown in Fig. 3(a), can be explained as follows. For a constant $V_{DS} = 0.5$ V and a low V_{GS}, the thermionic barrier from the drain into the channel is high, only few electrons can enter the channel and thus C_{gd} is low. Lowering the channel potential energy by increasing V_{GS} reduces this barrier, more and more electrons can enter the channel from the drain side and C_{gd} increases. The rate at which C_{gd} increases with a change in V_{GS} depends on the electrostatic control of the gate over the channel. For a constant V_{GS} in Fig. 3(b), a variation in V_{DS} has the same effect: A low V_{DS} constitutes a low drain-to-channel barrier, a large amount of electrons can enter the channel and C_{gd} is large. Increasing V_{DS} increases the barrier height, fewer electrons can enter the channel and C_{gd} decreases. This description can be expressed mathematically, starting with the channel charge Q_D supplied from the drain side [15]:

$$Q_D = qL_G \sum_n \int_{E_0}^{\infty} \frac{f_D(E)}{1 + T_{BTB}(E)} D_{1D}(E - E_n) dE. \quad (1)$$

Here, L_G is the gate length, E_0 the bottom of the channel conduction band, f_D the Fermi-Dirac distribution in the drain contact, $T_{BTB} \leq 0.02$ the tunnel transmission as noted before, D_{1D} the 1D density of states in the channel, and E_n the bottom of the respective sub-band n. Since the low tunnel transmission decouples the influence of the source from the channel, we are effectively left with a two-terminal device, so that moving E_0 by changing V_{GS}, and moving the Fermi level E_F by changing V_{DS}, has the same effect on $Q_D(E_F - E_0)$. Thus, we can obtain an expression

$$V_x = \frac{(E_F - E_0)}{q} + \frac{Q_D(E_F - E_0)}{C_{ox}} + \frac{(E_F - E_0)}{q} \frac{q D_{it}}{C_{ox}}, \quad (2)$$

which relates the applied voltage V_x on either terminal, gate or drain, to the energy. In (2), C_{ox} is the gate oxide capacitance and D_{it} the interface defect density. With (1) and (2), using the effective mass approximation for D_{1D}, numerical calculation of $C_{gd} = \partial Q_D / \partial V_d$ becomes straightforward and good agreement is achieved between measured and calculated values, as can be seen in Fig. 3. The only fitting parameters were the threshold voltage V_T and the interface defect density $D_{it} = 9 \times 10^{12}$ eV^{-1}cm^{-2}, which is close to measured D_{it} values for the same gate stack [16], especially when taking into account the large range of energies that is probed in the measurements here. The bumps in the analytical calculations originate in the sub-band spacing and cannot be resolved in the measurement due to statistical effects.

Equation (1) can also be expressed for the channel charge Q_S supplied from the source side by exchanging $1/(1+T_{BTB})$ and f_D with $T_{BTB}/(1+T_{BTB})$ and f_S, and with $T_{BTB} \leq 0.02$ as noted earlier, this yields values for C_{gs} not larger than 2.5 fF. This supports the previous interpretation of the intrinsic C_{gs} being much smaller than the parasitics from the gate-source overlap. For RF optimization, this overlap can be removed completely by optimized processing, so that only a small fringing contribution between the gate metal and the source should remain as parasitic C_{gs}. Once the parasitic overlap is removed, due to their different proportionalities to T_{BTB}, measuring the ratio of the intrinsic C_{gs} and the intrinsic C_{gd} should make it possible to determine the tunnel transmission experimentally. The parasitic part of C_{gd} can be reduced by increasing the bottom spacer thickness and by introducing finger gate contacts, which have been shown to reduce the bottom parasitic capacitance by up to 77% [17]. Furthermore, decreasing L_G and increasing T_{BTB} (cp. (1)) can reduce also the intrinsic C_{gd}. Inserting a total C_{gd} of 70 fF (40 fF parasitics from a 120-nm-thick bottom spacer and finger gates), C_{gs} of 10 fF, and $g_{gd,\omega}$ of 0.4 fF/Hz (by decreasing the amount of gate oxide defects) into the small-signal model in Fig. 2 for otherwise unchanged values, yields an increase of f_T and f_{max} by a factor of ten. Further optimization can be achieved by a more narrow spacing of the nanowires and by improving the tunnel junction, which will lead to a higher transconductance-to-capacitance ratio and should enable an increase by another factor of two.

V. Conclusion

From our measurements it is clear that in the on-state, C_{gd} largely dominates over C_{gs}. This confirms simulation results and stresses the need to increase the tunnel transmission. Furthermore, the measured asymmetric charge partitioning provides insight into the transport mechanisms of TFETs and the outlined scaling points the way to higher TFET RF performance and experimental determination of the tunnel transmission.

REFERENCES

[1] E. Memišević, M. Hellenbrand, E. Lind, A. R. Persson, S. Sant, A. Schenk, J. Svensson, R. Wallenberg, and L.-E. Wernersson, "Individual defects in InAs/InGaAsSb/GaSb nanowire tunnel field-effect transistors operating below 60 mV/decade," *Nano Lett.*, vol. 17, no. 7, pp. 4373–4380, Jun. 2017, doi: 10.1021/acs.nanolett.7b01455.

[2] D. H. Ahn, S. M. Ji, M. Takenaka, and S. Takagi, "Performance improvement of $In_xGa_{1-x}As$ tunnel FETs with quantum well and EOT scaling," in *Proc. IEEE Symp. VLSI Technol.*, Jun. 2016, pp. 1–2, doi: 10.1109/VLSIT.2016.7573443.

[3] A. Alian, Y. Mols, C. C. M. Bordallo, D. Verreck, A. Verhulst, A. Vandooren, R. Rooyackers, P. G. D. Agopian, J. A. Martino, A. Thean, D. Lin, D. Mocuta, and N. Collaert, "InGaAs tunnel FET with sub-nanometer EOT and sub-60 mV/dec sub-threshold swing at room temperature," *Appl. Phys. Lett.*, vol. 109, no. 24, p. 243502-1–243502-4, Dec. 2016, doi: 10.1063/1.4971830.

[4] X. Zhao, A. Vardi, and J. A. del Alamo, "Sub-thermal subthreshold characteristics in top-down InGaAs/InAs heterojunction vertical nanowire tunnel FETs," *IEEE Electron Device Lett.*, vol. 38, no. 7, pp. 855–858, Jul. 2017, doi: 10.1109/LED.2017.2702612.

[5] S. Cho, J. S. Lee, K. R. Kim, B.-G. Park, J. S. Harris, Jr., and I. M. Kang, "Analyses on small-signal parameters and radio-frequency modeling of gate-all-around tunneling field-effect transistors," *IEEE Trans. Electron Devices*, vol. 58, no. 12, pp. 4164–4171, Dec. 2011, doi: 10.1109/TED.2011.2167335.

[6] J. Wang, C. Wu, Q. Huang, C. Wang, and R. Huang, "A closed-form capacitance model for tunnel FETs with explicit surface potential solutions," *J. Appl. Phys.*, vol. 116, no. 9, p. 094501-1–094501-8, Sep. 2014, doi: 10.1063/1.4894624.

[7] Y. Yang, X. Tong, L.-T. Yang, P.-F. Guo, L. Fan, and Y.-C. Yeo, "Tunneling field-effect transistor: Capacitance components and modeling," *IEEE Electron Device Lett.*, vol. 31, no. 7, pp. 752–754, Jul. 2010, doi: 10.1109/LED.2010.2047240.

[8] S. Mookerjea, R. Krishnan, S. Datta, and V. Narayanan, "On enhanced Miller capacitance effect in interband tunnel transistors," *IEEE Electron Device Lett.*, vol. 30, no. 10, pp. 1102–1104, Oct. 2009, doi: 10.1109/LED.2009.2028907.

[9] D. Esseni, M. Guglielmini, B. Kapidani, T. Rollo, and M. Alioto, "Tunnel FETs for ultralow voltage digital VLSI circuits: Part I—Device-circuit interaction and evaluation at device level," *IEEE Trans. Very Large Scale Integr. (VLSI) Syst.*, vol. 22, no. 12, pp. 2488–2498, Dec. 2014, doi: 10.1109/TVLSI.2013.2293135.

[10] N. Dagtekin and A. M. Ionescu, "Impact of super-linear onset, off-region due to uni-directional conductance and dominant C_{GD} on performance of TFET-based circuits," *IEEE J. Electron Devices Soc.*, vol. 3, no. 3, pp. 233–239, May 2015, doi: 10.1109/JEDS.2014.2377576.

[11] R. Bijesh, H. Liu, H. Madan, D. Mohata, W. Li, N. V. Nguyen, D. Gundlach, C. A. Richter, J. Maier, K. Wang, T. Clarke, J. M. Fastenau, D. Loubychev, W. K. Liu, V. Narayanan, and S. Datta, "Demonstration of $In_{0.9}Ga_{0.1}As/GaAs_{0.18}Sb_{0.82}$ near broken-gap tunnel FET with I_{ON}=740 $\mu A/\mu m$, G_M=700 $\mu S/\mu m$ and gigahertz switching performance at $V_{DS} = 0.5$ V," in *IEDM Tech. Dig.*, Dec. 2013, pp. 28.2.1–28.2.4, doi: 10.1109/IEDM.2013.6724708.

[12] C. Liu, S. Glass, G. V. Luong, K. Narimani, Q. Han, A. T. Tiedemann, A. Fox, W. Yu, X. Wang, S. Mantl, and Q.-T. Zhao, "Experimental investigation of $C - V$ characteristics of Si tunnel FETs," *IEEE Electron Device Lett.*, vol. 38, no. 6, pp. 818–821, Jun. 2017, doi: 10.1109/LED.2017.2695193.

[13] E. Memisevic, J. Svensson, E. Lind, and L.-E. Wernersson, "InAs/InGaAsSb/GaSb nanowire tunnel field-effect transistors," *IEEE Trans. Electron Devices*, vol. 64, no. 11, pp. 4746–4751, Nov. 2017, doi: 10.1109/TED.2017.2750763.

[14] A. C. Seabaugh and Q. Zhang, "Low-voltage tunnel transistors for beyond CMOS logic," *Proc. IEEE*, vol. 98, no. 12, pp. 2095–2110, Dec. 2010, doi: 10.1109/JPROC.2010.2070470.

[15] M. Lundstrom and J. Guo, *Nanoscale Transistors*. New York, NY, USA: Springer, 2006.

[16] J. Wu, A. S. Babadi, D. Jacobsson, J. Colvin, S. Yngmann, R. Timm, E. Lind, and L.-E. Wernersson, "Low trap density in InAs/high-k nanowire gate stacks with optimized growth and doping conditions," *Nano Lett.*, vol. 16, no. 4, pp. 2418–2425, Mar. 2016, doi: 10.1021/acs.nanolett.5b05253.

[17] S. Johansson, E. Memisevic, L.-E. Wernersson, and E. Lind, "High-frequency gate-all-around vertical InAs nanowire MOSFETs on Si substrates," *IEEE Electron Device Lett.*, vol. 35, no. 5, pp. 518–520, May 2014, doi: 10.1109/LED.2014.2310119.