Lund University

Master Thesis

Low-Frequency Noise in InGaAs Nanowire MOSFETs

Author:
Christian Mario Möhle

Supervisors:
Erik Lind
Claes Thelander

A thesis submitted in partial fulfilment of the requirements for the degree of Master of Science

in the

Nanoelectronics Group
Department of Electrical and Information Technology
Faculty of Engineering, LTH

May 2017

LUND UNIVERSITY

Abstract

Faculty of Engineering, LTH
Department of Electrical and Information Technology

Master of Science

Low-Frequency Noise in InGaAs Nanowire MOSFETs

by Christian Mario MÖHLE

Low-frequency (LF) noise ($1/f$ as well as random telegraph-signal (RTS) noise) measurements were performed on high-performance InGaAs nanowire (NW) metal-oxide-semiconductor field-effect transistors (MOSFETs).
$1/f$ noise measurements at room temperature (RT) show that the dominant noise mechanism is carrier number fluctuations. The minimum gate area normalized input gate voltage noise power spectral density (PSD) for these devices is as low as $S_{V_G} L_G W_G = 80 \mu m^2 \mu V^2 Hz^{-1}$, with a corresponding minimum trap density of $N_T = 9 \cdot 10^{18}\, cm^{-3} eV^{-1}$, where the calculation of N_T bases on elastic tunneling of the channel electrons to and from the trap states in the gate oxide. These values are among the lowest for III-V MOSFETs, demonstrating the feasibility of a high-quality, low trap density, high-κ gate oxide on InGaAs NW structures. $1/f$ noise measurements at low temperatures confirm that the dominant noise mechanism is carrier number fluctuations and also give some support for the elastic tunneling model.
RTS noise signals were found in approximately 2/3 of the considered devices, indicating that very few traps influence the electron transport through the channel. It turns out that the capture and emission time constant of a single active trap state change with the applied gate bias. Different models (elastic tunneling model and Shockley-Read-Hall (SRH) model) are discussed to explain this behavior. Another striking feature of the RTS noise measurements is the large spread in the RTS noise amplitudes, with maximum amplitudes of up to $1\, \mu A$. Simulations of single trap induced subband fluctuations can reproduce this large spread, showing that the magnitude of the RTS noise amplitude is especially sensitive to the position of the trap state along the NW circumference.
Temperature-dependent RTS noise measurements reveal a thermal activation of the time constants, which is inconsistent with elastic tunneling but can be explained by multi-phonon-assisted tunneling (inelastic tunneling).

Acknowledgements

First and foremost, I would like to thank you, Erik Lind, for enabling me to write my Master thesis in your research group and for being my first supervisor. You inducted me to the fascinating field of nanoelectronics and reinforced my interest in solid state physics and quantum mechanics, for which I am very grateful. Throughout the whole project, you inspired me with your ideas and helped me to interpret my measurement data. Thank you for always being approachable and for answering all my questions with a lot of patience! Likewise, I am indebted to my second supervisor, Claes Thelander.

The next person, I would like to thank is my practical coach, Markus Hellenbrand, who made me familiar with the low-frequency noise measurement setup and who always helped me out when technical problems appeared. On top of that, I very much appreciate how much time you spent with me, discussing my results and ever shining some more light on my them. It was very nice collaborating with you! Similarly, you, Cezar Zota, were also always there for me to answer my questions and to talk over my data. Thank you, in addition, for fabricating the transistors I measured on!

Apart from the above mentioned members of the nanoelectronics group, each and everyone else in this group deserves to be mentioned here for making me fell warm and welcome from the first day: Lars-Erik Wernersson, Mattias Borg, Johannes Svensson, Lars Ohlsson, Olli-Pekka Kilpi, Sebastian Heunisch, Adam Jönsson, Fredrik Lindelöw, Elvedin Memisevic and Stefan Andric. A special thank you also to my fellow Master students in the nanoelectronics group: Andreas Malmgren, Albin Linder, Daniel Svedbrand, George Gioulis, Lun Sang and Edvin Winqvist. Sharing the office with you was great!

Last but not least, I would like to express my deepest thankfulness to my family and friends. Anja and Heiko Möhle, there are no better parents than you on earth! Florian Möhle, you are the best brother anybody can ask for. An enormous thank you also to my grandparents, to my aunts and uncles and to my cousins. Among my friends, I want to set Kevin Kleinwort apart. There are no words needed to describe our friendship! Thank you for everything!

Contents

Abstract i

Acknowledgements ii

Contents iii

Abbreviations iv

1 Introduction 1

2 Theory 4
- 2.1 MOSFET Theory 4
  - 2.1.1 Planar, Drift and Diffusion MOSFET 4
  - 2.1.2 1D Ballistic MOSFET 10
- 2.2 Low-Frequancy Noise Theory 16
  - 2.2.1 General Noise Properties 17
  - 2.2.2 $1/f$ Noise 17
  - 2.2.3 RTS Noise 20

3 Device Fabrication 26

4 Experimental Setup 28

5 Results and Discussion 31
- 5.1 DC Performance 31
- 5.2 $1/f$ Noise at RT 31
- 5.3 $1/f$ Noise at T=11 K 35
- 5.4 RTS Noise at T=11 K 36
- 5.5 RTS Noise at Different Temperatures 41

6 Conclusions and Outlook 44

Bibliography 46

Abbreviations

LF	**L**ow-**F**requency
RTS	**R**andom-**T**elegraph-**S**ignal
NW	**N**ano**w**ire
MOSFET	**M**etal-**O**xide-**S**emiconductor **F**ield-**E**ffect **T**ransistor
CMOS	**C**omplementary **M**etal-**O**xide-**S**emiconductor
RT	**R**oom **T**emperature
PSD	**P**ower **S**pectral **D**ensity
SRH	**S**chockley-**R**ead-**H**all
IC	**I**ntegrated **C**ircuit
1D	One-**D**imensional
3D	Three-**D**imensional
EBL	**E**lectron-**B**eam **L**ithography
SMU	**S**ource **M**easure **U**nit

Chapter 1

Introduction

Transistors are the main building blocks of modern electronic devices, such as smartphones and laptops, and they are used to amplify or to switch electronic signals. Nowadays, the most common realization of a transistor is the MOSFET, in which a MOS capacitor, is placed in between a source and a drain. The source and drain consist of a semiconducting material contacted by a metal and in combination with the semiconductor of MOS capacitor, the structure forms two *pn*-junctions, which are connected back-to-back. The MOS capacitor metal and oxide play the role of the gate terminal that controls the current flowing in a narrow channel beneath the oxide between the source and the drain. Traditionally, polysilicon was used as the gate "metal", SiO_2 as the gate oxide and Si as the semiconducting channel material. Depending on the electrical polarity, MOSFETs can be categorized into nMOSFETs (current due to electrons) and pMOSFETs (current due to holes). Combining an nMOSFET and a pMOSFET results in a complementary metal-oxide-semiconductor (CMOS) inverter, needed for logic operations.

In 1965, Gordon Moore, co-founder of Intel, observed that the number of transistors per square inch in integrated circuits (ICs) had doubled every two years since the IC was invented and he predicted that this trend would continue for the forseeable future [1]. This prediction is known as "Moore's law" and has become a guideline for the semiconductor industry. In 1974, Robert H. Dennard proposed a transistor scaling law ("Dennard scaling"), based on keeping the electric field in the devices constant, to actually realize Moore's law [2]. By scaling down the transistor dimensions and the supply voltage (by a factor of $1/\kappa = 1/\sqrt{2} = 0.7$) and increasing the channel doping (by κ), the

electric field could be kept constant while the device area was halved in every scaling generation. Apart from enabling a doubling of transistors in ICs per generation (effectively leading to a cost reduction of the ICs), the Dennard scaling also improved the device performance in terms of speed and power consumption.

After several decades of successful application, Moore's law has faced an increasing amount of challenges. Technologywise, short-channel effects started to appear when scaling the gate length to very small dimensions. Among others, a leakage current through the channel became a severe problem, resulting in an increased power consumption. On top of that, the aggressive gate oxide scaling induced a gate leakage current. Processingwise, the transistor scaling has become more and more expensive since ultra small device dimensions require the use of new and more complex fabrication techniques (such as extreme ultraviolet lithography).

To overcome these obstacles, new technologies are needed. Replacing the polysilicon gate contact by a metal minimizes the gate resistance; utilizing a gate oxide with a higher dielectric constant than SiO_2 (e.g. HfO_2), effectively reduces the gate leakage current. Furthermore, building the transistors in a FinFET or nanowire geometry, in which the gate contacts the channel from almost all or all sides, lowers the short channel effects due to an improved electrostatic control of the channel [3]. This enables the use of highly scaled gate lengths.

Besides changing the gate contact, the gate oxide and the transistor structure, there are also considerations to replace the traditional channel material (Si) by a new material with enhanced transport properties. A higher mobility of the charge carriers leads to an increased on-current and thus to a faster switching behavior. Among these new materials, III-V compound semiconductors are particularly promising [4]. InGaAs is currently the most attractive candidate for future III-V based nMOSFETs while InGaSb is of great interest for pMOSFETs [5].

However, using III-V compound semiconductors as the channel material leads to a worse gate oxide quality (as compared with Si/SiO_2 devices) since native III-V oxides form a very poor interface to III-V channel materials. Depositing a foreign oxide on the III-V channel also results in the formation of a large amount of trap states at the channel-oxide interface (interface traps), as well as deeper in the oxide (border traps). These trap states affect the charge carrier transport through the channel, generally degrading the device performance and reliability quite significantly [6]. Thus, accurate and reliable

measurements of the oxide quality of III-V FETs are required for the device characterization and the process optimization. Conventional oxide characterization methods, such as C-V and charge pumping methods, cannot be used for ultra small devices without a body contact. Instead, LF noise measurements can be utilized to analyze the performance and reliability of highly scaled devices.

In this Master thesis, LF noise ($1/f$ as well as RTS noise) measurements are performed on high performance InGaAs NW nMOSFETs with varying gate lengths and gate widths. The goal of these measurements is to analyze the trap density in these devices and to gain insights about the trapping and detrapping mechanism of the electrons to and from the trap states.

The thesis is structured as follows: In chapter 2, the basic MOSFET theory for planar, drift and diffusion devices and for one-dimensional (1D), ballistic devices is presented. Also, the LF noise theory, including $1/f$ and RTS noise, is depicted. Chapter 3 describes the device fabrication and chapter 4 the experimental setup used for the noise measurements. In chapter 5, the measurement results are presented and discussed and chapter 6 summarizes the main outcomes of this thesis and gives an outlook on future work.

Chapter 2

Theory

2.1 MOSFET Theory

In this section, the basic MOSFET theory is introduced - first for planar MOSFETs, obeying the drift and diffusion transport mechanism through the channel, and later for 1D ballistic devices. The drift and diffusion theory is presented because it is the prerequisite to understand more advanced theories. It is applicable for long channel transistors, in which the mean free path of the electrons is much shorter than the channel length, or in other words, the transport is dominated by scattering events. The ballistic transport theory, on the other hand, applies for modern short channel devices with a mean free path of the electrons much longer than the channel length. The devices for the noise measurements operate in the quasi-ballistic regime (with an average transmission of 70%), so it is useful to explain the ballistic theory as well.
Confining the electrons to one dimension in NW structures (instead of to two dimensions in planar transistors) leads to additional quantum effects, considered in the ballistic transport theory section.
The drift and diffusion MOSFET theory is based on [7] (unless otherwise stated), whereas the ballistic MOSFET theory is built upon [8].

2.1.1 Planar, Drift and Diffusion MOSFET

This subsection deals with planar drift and diffusion MOSFETs. First of all, the ideal MOS capacitor electrostatics will be explained and afterwards, the operation of the ideal

MOSFET is examined.

The Ideal MOS Capacitor

In figure 2.1, a schematic of the MOS capacitor is depicted (perspective view in (a) and cross-sectional view in (b)), where d is the oxide thickness and V is the applied metal voltage. In a MOSFET, V will be the gate voltage, called V_{G} (or V_{GS}, when the source is used as the reference contact).

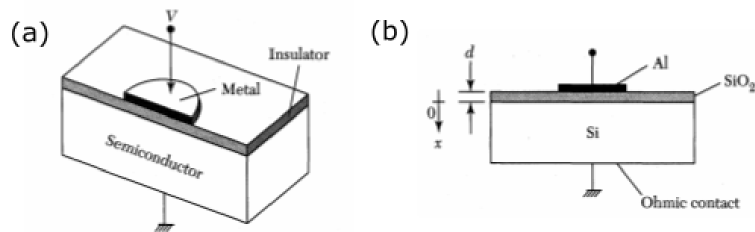

FIGURE 2.1: (a) Perspective view on a MOS capacitor. (b) Cross-sectional view on a MOS capacitor. Figures taken from [7].

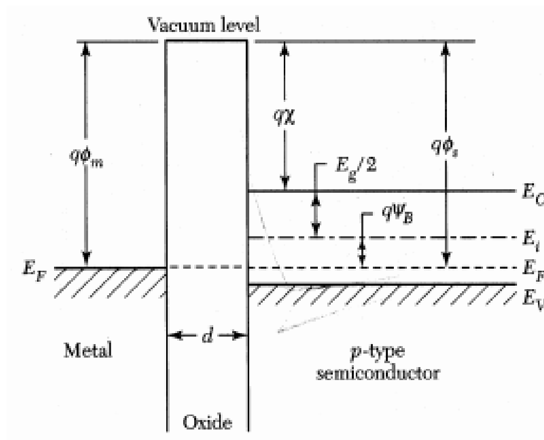

FIGURE 2.2: Band diagram of an ideal p-type MOS capacitor at $V = 0\,\mathrm{V}$. Figure taken from [7].

Figure 2.2 shows the energy band diagram of an ideal p-type MOS capacitor when $V = 0\,\mathrm{V}$. $q\phi_{\mathrm{m}}$ is defined as the gate metal work function (with q being the elementary charge) and $q\phi_{\mathrm{s}}$ as the semiconductor workfunction. $q\chi$ is the semiconductor electron affinity, $q\chi_{\mathrm{i}}$ the oxide electron affinity, $q\phi_{\mathrm{B}}$ the energy barrier between the gate metal and the oxide and $q\Psi_{\mathrm{B}}$ the energy difference between the semiconductor Fermi level (E_{F}) and the intrinsic semiconductor Fermi level (E_{i}).

An ideal MOS capacitor is characterized as follows: (1) When $V = 0\,\text{V}$, the gate metal Fermi level and the semiconductor Fermi level are aligned, implying flat bands ("flat-band condition") and $q\phi_\text{m} - q\phi_\text{s} = 0\,\text{J}$. (2) The only charges (net charges) that exist under any biasing conditions are the ones in the semiconductor (Q_s[1]) and the ones with opposite sign located at the metal-oxide interface in the metal (-Q_m), with $|Q_\text{s}| = |Q_\text{m}|$. (3) Charge carriers cannot penetrate through the oxide, meaning that there is no gate leakage current.

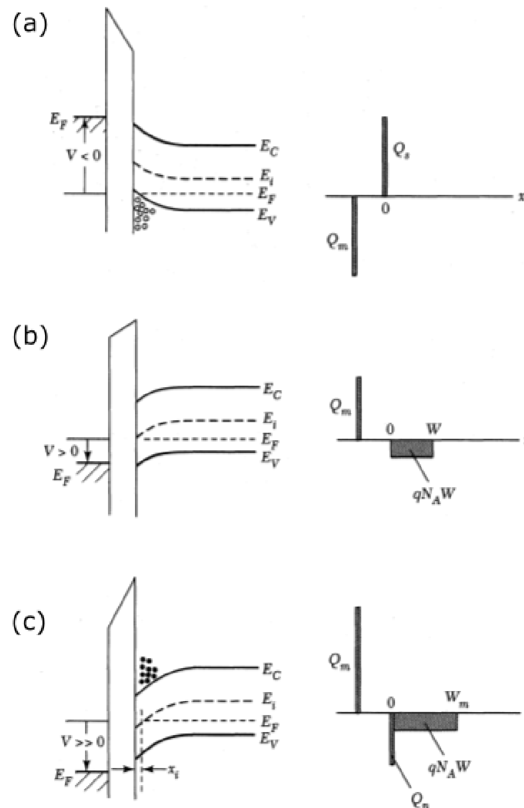

FIGURE 2.3: Band diagram of an ideal p-type MOS capacitor and corresponding charge distribution when (a) $V < 0\,\text{V}$, (b) $V > 0\,\text{V}$ and (c) $V \gg 0\,\text{V}$. Figures taken from [7].

When $V < 0\,\text{V}$, the metal Fermi level is raised by qV with respect to the semiconductor Fermi level, causing the bands in the semiconductor to bend upwards. As a consequence, holes will accumulate close to the semiconductor-oxide interface ("accumulation case"). The positive charge at the semiconductor-oxide interface attracts electrons in the metal to move to the metal-oxide interface, leading to a charge distribution as shown in figure 2.3 (a).

[1] Q_s and Q_m are the charges per unit area.

When a small positive metal voltage $V > 0\,\mathrm{V}$ is applied, the metal Fermi level is lowered and the bands in the semiconductor bend downwards. This causes holes to deplete the semiconductor-oxide interface region ("depletion case"), leaving behind a negative space charge per unit area of $Q_\mathrm{s} = -qN_\mathrm{A}W$, where N_A is the substrate doping level and W is the width of the surface depletion region. Q_s is balanced by a positive Q_m, consisting of accumulated holes (see figure 2.3 (b)).

Applying a larger positive metal bias, causes the bands to bend down even further, attracting minority carriers (electrons) to the semiconductor-oxide interface region. Eventually, the electron concentration at the surface will be larger than the hole concentration, resulting in an inverted surface ("inversion case"). In strong inversion (electron concentration near the semiconductor-oxide interface larger than the substrate doping level), most of the negative charges in the semiconductor consist of the electron charge Q_n in a narrow n-type inversion layer and the depletion-layer width reaches a maximum (W_m). The corresponding charge distribution is shown in figure 2.3 (c).

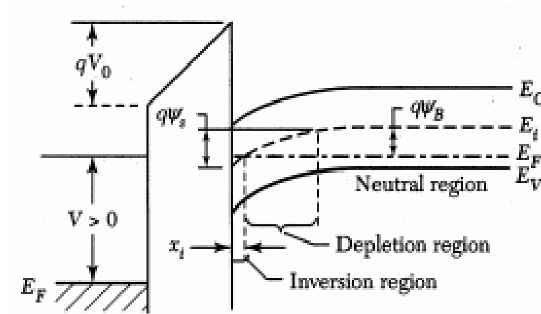

FIGURE 2.4: Band diagram of an ideal p-type MOS capacitor when $V > 0\,\mathrm{V}$, showing that V drops partly across the gate oxide (V_o) and partly across the semiconductor (Ψ_s). Figure taken from [7].

Figure 2.4 shows the band diagram of an ideal p-type MOS capacitor with a positive applied metal voltage in more detail. The electrostatic potential Ψ is defined as zero in the bulk of the semiconductor. The potential at the semiconductor surface is called the surface potential Ψ_s. The applied metal voltage drops partly across the oxide and partly across the semiconductor, so $V = V_\mathrm{ox} + \Psi_\mathrm{s}$. The total capacitance of the structure C is given by a series combination of the oxide capacitance C_ox and the semiconductor depletion-layer capacitance C_d, meaning $C = C_\mathrm{ox}C_\mathrm{d}/(C_\mathrm{ox}+C_\mathrm{d})$. Strong inversion occurs

when $\Psi_s = 2\Psi_B$, which appears when the metal voltage reaches a value of:

$$V_T = \frac{\sqrt{2\varepsilon_s q N_A (2\Psi_B)}}{C_{ox}} + 2\Psi_B, \tag{2.1}$$

where ε_s is the permittivity of the semiconductor and Ψ_B is the semiconductor bulk potential. V_T is called the threshold voltage.

The Ideal MOSFET

In an n-type MOSFET, the p-type MOS capacitor is embedded between two n-type regions (the source and the drain), forming two pn-junctions, which are connected back-to-back (see figure 2.5).

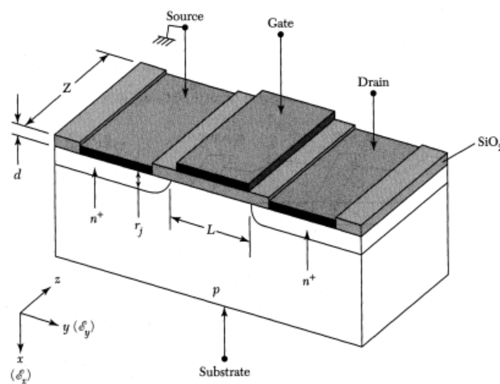

Figure 2.5: Perspective view on a conventional, planar MOSFET, consisting of a MOS capacitor embedded between two pn-junctions. Figure taken from [7].

Without applied gate voltage, no current can flow between the source and the drain (except for a small reverse leakage current) because the electrons in the n-type regions cannot penetrate through the p-type region. However, when a sufficiently large gate voltage is applied, the MOS structure becomes inverted and a surface inversion layer (or channel) is formed between the source and the drain. Consequently, a large current can flow through the MOSFET and the magnitude of the current can be modulated by the gate voltage.

Usually, the source contact is grounded, while a positive voltage is applied to the drain (drain voltage, V_{DS}), in which case electrons flow from the source to the drain and the resulting current (drain current, I_D) flows in the opposite direction.

Assuming the MOS structure is inverted, two different MOSFET operation regions can

be distinguished, depending on V_{DS}: the linear region ($V_{\mathrm{DS}} < V_{\mathrm{DS,sat}}$) and the saturation region ($V_{\mathrm{DS}} > V_{\mathrm{DS,sat}}$). When $V_{\mathrm{DS}} = V_{\mathrm{DS,sat}}$ ("pinch-off point"), the inversion layer thickness at the drain side decreases to zero and I_{D} saturates.

Expressions for I_{D} in the two regions can be derived under the following assumptions: (1) The gate structure is an ideal MOS capacitor. (2) Only drift current is considered. (3) The electron mobility in the channel is constant. (4) The doping in the channel is uniform. (5) The reverse-leakage current is negligible. (6) The transverse electric field created by V_{GS} is much larger than the longitudinal field created by V_{DS} ("gradual-channel approximation").

In the linear region, I_{D} is given by:

$$I_{\mathrm{D}} = \begin{cases} \frac{W_{\mathrm{G}}}{L_{\mathrm{G}}} \mu_{\mathrm{n}} C_{\mathrm{ox}} \left(V_{\mathrm{GS}} - V_{\mathrm{T}} - \frac{V_{\mathrm{DS}}}{2} \right) V_{\mathrm{DS}}, & V_{\mathrm{DS}} < (V_{\mathrm{GS}} - V_{\mathrm{T}}) \\ \frac{W_{\mathrm{G}}}{L_{\mathrm{G}}} \mu_{\mathrm{n}} C_{\mathrm{ox}} \left(V_{\mathrm{GS}} - V_{\mathrm{T}} \right) V_{\mathrm{DS}}, & V_{\mathrm{DS}} \ll (V_{\mathrm{GS}} - V_{\mathrm{T}}) \end{cases} \tag{2.2}$$

where L_{G} is the channel length (called L in figure 2.5), W_{G} is the channel width (called Z in figure 2.5) and μ_{n} is the electron mobility.

In the saturation region, the drain current is:

$$I_{\mathrm{D,sat}} = \left(\frac{W_{\mathrm{G}} C_{\mathrm{ox}} \mu_{\mathrm{n}}}{2 L_{\mathrm{G}}} \right) (V_{\mathrm{GS}} - V_{\mathrm{T}})^2. \tag{2.3}$$

In short channel MOSFETs (still obeying the drift and diffusion transport equations), the electron drift velocity saturates at v_{sat} and the drain current is instead given by [9]:

$$I_{\mathrm{D,sat}} = W_{\mathrm{G}} C_{\mathrm{ox}} v_{\mathrm{sat}} (V_{\mathrm{GS}} - V_{\mathrm{T}}). \tag{2.4}$$

In general, the drain current in the saturation region can be described by [9]:

$$I_{\mathrm{D,sat}} \sim (V_{\mathrm{GS}} - V_{\mathrm{T}})^{\alpha}, \quad 1 < \alpha < 2. \tag{2.5}$$

Independent of the operation region, it is useful to define two quantities, describing the transistor performance: the channel conductance

$$g_{\mathrm{D}} \equiv \frac{\partial I_{\mathrm{D}}}{\partial V_{\mathrm{DS}}}\Big|_{V_{\mathrm{GS}}=\mathrm{const.}} \tag{2.6}$$

and the transconductance

$$g_{\mathrm{m}} \equiv \frac{\partial I_{\mathrm{D}}}{\partial V_{\mathrm{GS}}}\Big|_{V_{\mathrm{DS}}=\mathrm{const.}}. \tag{2.7}$$

Ideally, $g_{\mathrm{D}} = 0\,\mathrm{A/V}$ in the saturation region and the normalized transconductance (by W_G or by the gate circumference in 1D structures) should be as high as possible.
In the subthreshold region ($V_{\mathrm{GS}} < V_{\mathrm{T}}$), I_{D} is dominated by diffusion instead of drift and is given by:

$$I_{\mathrm{D}} \sim e^{\frac{q(V_{\mathrm{GS}}-V_{\mathrm{T}})}{k_{\mathrm{B}}T}}, \tag{2.8}$$

where k_{B} is the Boltzmann constant and T is the temperature.
A particularly important parameter in this region is the subthreshold swing, which is defined as:

$$S_{\mathrm{s,th}} \equiv \ln(10)\frac{\mathrm{d}V_{\mathrm{GS}}}{\mathrm{d}[\ln{(I_{\mathrm{D}})}]} = \frac{\ln(10)k_{\mathrm{B}}T}{q} \tag{2.9}$$

and describes how sharply the transistor can be turned off by the gate voltage. The minimum value of $S_{\mathrm{s,th}}$ for both, drift and diffusion and ballistic MOSFETs is 60 mV/dec at RT. The closer $S_{\mathrm{s,th}}$ is to this minimum value, the better is the subthreshold device performance.

2.1.2 1D Ballistic MOSFET

Figure 2.6 summarizes the main results of the previous section. Figure 2.6 (a) shows a typical output characteristic ($I_{\mathrm{D}} - V_{\mathrm{DS}}$ curves) of a planar MOSFET at room temperature. In figure 2.6 (b), a transfer characteristic ($I_{\mathrm{D}} - V_{\mathrm{GS}}$ curve) is depicted for such a device. When V_{DS} is low (linear region), $I_{\mathrm{D}} \sim (V_{\mathrm{GS}} - V_{\mathrm{T}})$ (see equation 2.2); when V_{DS} is high (saturation region), $I_{\mathrm{D}} \sim (V_{\mathrm{GS}} - V_{\mathrm{T}})^{\alpha}$ (see equation 2.5), with $1 < \alpha < 2$

for planar drift and diffusion MOSFETs. For planar ballistic MOSFETs, $\alpha = 1$ (non-degenerate conditions) or $\alpha = 1.5$ (degenerate conditions). The transconductance is almost constant for low V_{DS}, whereas it has the form of $g_{\mathrm{m}} \sim (V_{\mathrm{GS}} - V_{\mathrm{T}})^{\alpha-1}$ for high V_{DS} (see figure 2.6 (c)).

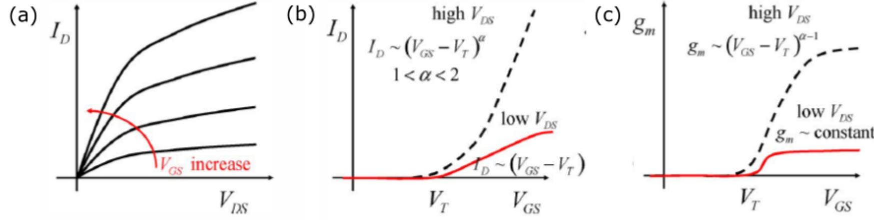

FIGURE 2.6: (a) Output characteristic of a planar MOSFET at RT. (b) Transfer characteristic of a planar MOSFET for low and high V_{DS} at RT. (c) $g_{\mathrm{m}} - V_{\mathrm{GS}}$ characteristic of a planar MOSFET for low and high V_{DS} at RT. Figures taken from [8].

The goal of this subsection is to analyze how these features change for a 1D ballistic device at $T = 0\,\mathrm{K}$, considering one subband first and two and more subbands afterwards.

1D Ballistic Transport - One Subband

The following 1D ballistic transport theory bases on the top-of-the-barrier model, in which the electron transport through the transistor is essentially determined by the physical processes at the source-channel energy barrier.

First, one parabolic subband is regarded:

$$E(k) = \varepsilon(0) + \frac{\hbar^2 k^2}{2m^*}, \tag{2.10}$$

where ε_0 is the potential at the top of the source-channel barrier (modified by the gate voltage), $\hbar$ is the reduced Planck constant ($\hbar = h/2\pi$), k is the electron wave vector (or wave number) and m^* is the effective electron mass.

Electrons with positive k-states are injected from the source (filled up to the Fermi level), and electrons with negative k-states from the drain (under low V_{DS}). Both electrodes are viewed as thermal equilibrium reservoirs.

The density of the electrons at the top of the barrier is controlled by MOS electrostatics

so that the charge in the semiconductor balances that in the gate.

The directed moments at the top of the barrier can be calculated to be:

$$n_L^+ = \frac{\sqrt{2m^*(E_\mathrm{F} - \varepsilon(0))}}{\pi\hbar}\Theta(E_\mathrm{F} - \varepsilon(0)) \tag{2.11}$$

$$n_\mathrm{L}^- = \frac{\sqrt{2m^*(E_\mathrm{F} - \varepsilon(0) - qV_\mathrm{DS})}}{\pi\hbar}\Theta(E_\mathrm{F} - \varepsilon(0) - qV_\mathrm{DS}) \tag{2.12}$$

$$I^+ = \frac{2q^2}{h}\frac{E_\mathrm{F} - \varepsilon(0)}{q}\Theta(E_\mathrm{F} - \varepsilon(0)) \tag{2.13}$$

$$I^- = \frac{2q^2}{h}\frac{E_\mathrm{F} - \varepsilon(0) - qV_\mathrm{DS}}{q}\Theta(E_\mathrm{F} - \varepsilon(0) - qV_\mathrm{DS}) \tag{2.14}$$

with $n_L = n_L^+ + n_L^-$ being the 1D electron density, E_F the Fermi level at the source side and $I_\mathrm{D} = I^+ - I^-$ the total drain current. The superscript "+" stands for electrons occupying the positive k-states and moving towards the drain. The superscript "-" stands for electrons occupying the negative k-states and moving towards the source.

The potential at the top of the barrier as a function of V_GS and V_DS can be expressed as:

$$\varepsilon(0) = \frac{C_\mathrm{ox}}{C_\sum}(-qV_\mathrm{G}) + \frac{C_\mathrm{D}}{C_\sum}(-qV_\mathrm{D}) + \frac{C_\mathrm{S}}{C_\sum}(-qV_\mathrm{S}) + \frac{q^2}{C_\sum}(n_\mathrm{L} - n_0). \tag{2.15}$$

C_ox, C_S and C_D refer to the gate oxide, source and drain capacitance, respectively. $C_\sum = C_\mathrm{ox} + C_\mathrm{S} + C_\mathrm{D}$ is the sum of these capacitances and n_0 is the electron density without applied bias.

Assuming perfect gate control ($C_\mathrm{ox}/C_\sum = 1$, $C_\mathrm{S}/C_\sum = C_\mathrm{D}/C_\sum = 0$) and setting V_S as the reference electrode, equation 2.15 simplifies to:

$$\varepsilon(0) = -qV_\mathrm{GS} + \frac{q^2 n_\mathrm{L}}{C_\mathrm{ox}}. \tag{2.16}$$

The drain current as a function of V_GS and V_DS can be calculated by solving equation 2.11, 2.12 and 2.16 self-consistently. It is, however, possible to find simple analytic expressions for I_D in the three regions: (1) $\varepsilon(0) > E_\mathrm{F}$, (2) $E_\mathrm{F} - qV_\mathrm{DS} < \varepsilon(0) \leq E_\mathrm{F}$, and (3) $\varepsilon(0) \leq E_\mathrm{F} - qV_\mathrm{DS}$.

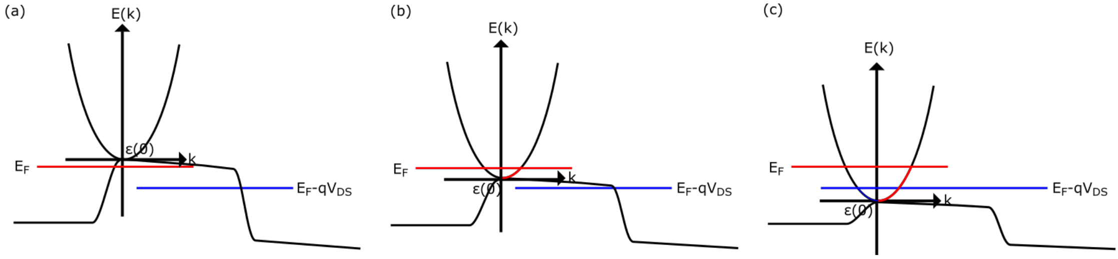

FIGURE 2.7: Band diagram of a 1D ballistic MOSFET with one subband for (a) $\varepsilon(0) > E_\mathrm{F}$ (region (1)), (b) $E_\mathrm{F} - qV_\mathrm{DS} < \varepsilon(0) \leq E_\mathrm{F}$ (region (2)) and (c) $\varepsilon(0) \leq E_\mathrm{F} - qV_\mathrm{DS}$ (region (3)).

In region (1), no states are occupied and thus $I_\mathrm{D} = 0$ (see figure 2.7 (a)).

When V_GS increases to $V_\mathrm{T} \equiv -E_\mathrm{F}/q$, $\varepsilon(0)$ is aligned with E_F and I_D starts to flow. In region (2), I_D is given as (see also figure 2.7 (b)):

$$I_\mathrm{D} = \frac{2q^2}{h} \frac{C_\mathrm{ox}}{C_\mathrm{ox} + C_\mathrm{q}} (V_\mathrm{GS} - V_\mathrm{T}), \tag{2.17}$$

with C_q being the quantum capacitance, taking into account the effects of confining the electrons to 1D:

$$C_\mathrm{q} = \frac{\sqrt{2m^*} q^2 / \pi\hbar}{-(\sqrt{2m^*} q^2 / h C_\mathrm{ox}) + \sqrt{2m^* q^4 / h^2 C_\mathrm{ox}^2 + q(V_\mathrm{GS} - V_\mathrm{T})}}. \tag{2.18}$$

In region (3), I_D becomes independent of V_GS (see figure 2.7 (c)):

$$I_\mathrm{D} = \frac{2q^2}{h} V_\mathrm{DS}. \tag{2.19}$$

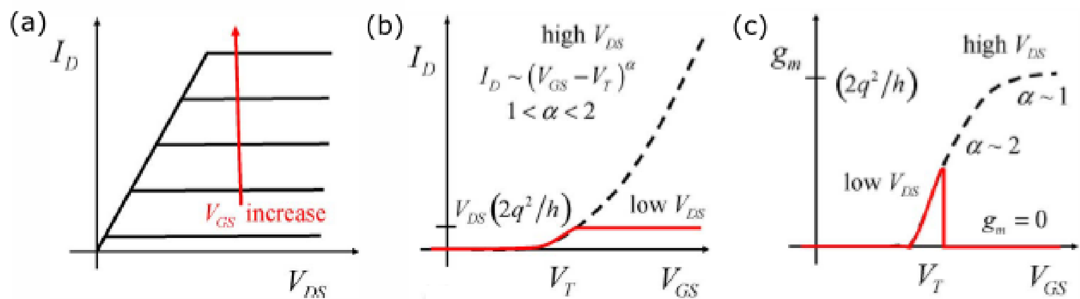

FIGURE 2.8: (a) Output characteristic of a 1D ballistic MOSFET with one subband at $T = 0\,\mathrm{K}$. (b) Transfer characteristic of a 1D ballistic MOSFET with one subband for low and high V_DS at $T = 0\,\mathrm{K}$. (c) $g_\mathrm{m} - V_\mathrm{GS}$ characteristic of a 1D ballistic MOSFET with one subband for low and high V_DS at $T = 0\,\mathrm{K}$. Figures taken from [8].

Figure 2.8 (a) shows the output characteristic of a 1D ballistic MOSFET. For low V_{DS} (linear region), I_{D} is independent of V_{GS} as shown in equation 2.19. For high V_{DS} (saturation region), I_{D} is saturated and independent of V_{DS} as given in equation 2.17. Figure 2.8 (b) illustrate the transfer characteristic of a 1D ballistic MOSFET at low and high V_{DS}. When V_{DS} is low (see figure 2.9 (a)), I_{D} starts to increase with V_{GS} according to equation 2.17, but then saturates and becomes independent of V_{GS}, as shown in equation 2.19. Consequently, the transfer curve has a graded step around V_{T} and g_{m} has the shape of a triangular spike at V_{T}. As V_{DS} becomes smaller, the transfer curve changes into an abrupt step function and g_{m} has a sharp peak at V_{T}. This behavior is different from the planar MOSFET, where g_{m} is constant for low V_{DS}.
For high V_{DS} (see figure 2.9 (b)), the transfer curve does not saturate but keeps increasing according to equation 2.17. When $C_{\mathrm{ox}} \ll C_{\mathrm{q}}$, the equation 2.17 reduces to:

$$I_{\mathrm{D}} = \frac{h}{4m^* q} C_{\mathrm{ox}}^2 (V_{\mathrm{GS}} - V_{\mathrm{T}})^2, \tag{2.20}$$

so $\alpha = 2$, which is different from a planar ballistic MOSFET for which $\alpha = 1.5$ in the degenerate case. When $C_{\mathrm{q}} \ll C_{\mathrm{ox}}$ (occurring when V_{GS} is high), equation 2.17 is:

$$I_{\mathrm{D}} = \frac{2q^2}{h} (V_{\mathrm{GS}} - V_{\mathrm{T}}). \tag{2.21}$$

Thus, for a 1D ballistic MOSFET, α varies from 1 to 2 in the degenerate case, depending on C_{ox} and V_{GS}. The transconductance when V_{DS} is high is given as:

$$g_{\mathrm{m}} = \begin{cases} \frac{h}{2m^* q} C_{\mathrm{ox}}^2 (V_{\mathrm{GS}} - V_{\mathrm{T}}), & C_{\mathrm{ox}} \ll C_{\mathrm{q}} \\ \frac{2q^2}{h}, & C_{\mathrm{q}} \ll C_{\mathrm{ox}} \end{cases} \tag{2.22}$$

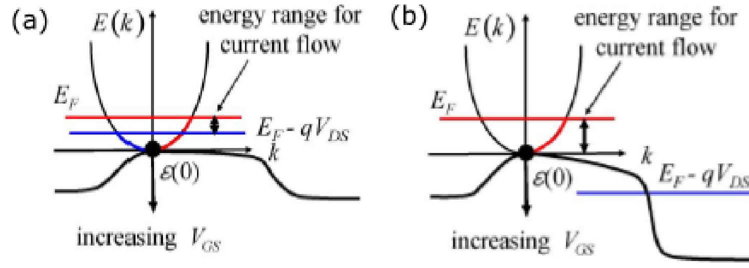

FIGURE 2.9: Band diagram of a 1D ballistic MOSFET with one subband for (a) low and (b) high V_{DS}. Figures taken from [8].

1D Ballistic Transport - Two and More Subbands

In this section, the *I*-*V* characteristics of a 1D ballistic MOSFET with two or more occupied subbands at $T = 0\,\mathrm{K}$ will be discussed. In figure 2.10 (a), the band diagram with two subbands is depicted for low V_{DS} and the same thing is depicted in figure 2.10 (b) for high V_{DS}.

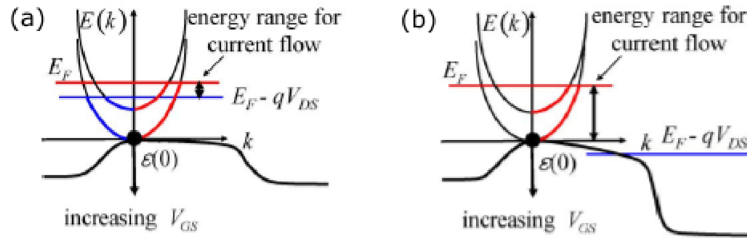

FIGURE 2.10: Band diagram of a 1D ballistic MOSFET with two subbands for (a) low and (b) high V_{DS}. Figures taken from [8].Figure taken from [8].

When V_{DS} is low (linear region), the energy range for the current flow is constant, but I_{D} increases in steps because the number of conducting channels (M) increases with V_{GS}:

$$I_{\mathrm{D}} = M \left(\frac{2q^2}{h} \right) V_{\mathrm{DS}}. \tag{2.23}$$

The transconductance peaks whenever E_{F} is aligned with a subband minimum. For M conducting modes, the transconductance as a function of V_{GS} is a series of M delta peaks.

For high V_{DS} (saturation region), the energy range for the current flow increases with

V_{GS} and I_{D} also increases because the number of conducting channels increases. I_{D} is given as:

$$I_{\mathrm{D}} \sim \sum_{m=1}^{M} (V_{\mathrm{GS}} - V_{\mathrm{T}})^{\alpha_{\mathrm{m}}}, \quad 1 < \alpha_{\mathrm{m}} < 2. \tag{2.24}$$

In general, g_{m} vs. I_{DS} has a complicated shape because α depends on V_{GS}.
Figure 2.11 (a) shows the output characteristic for a 1D ballistic MOSFET with two occupied subbands, (b) the transfer characteristic for low and high V_{DS} and (c) g_{m} vs. V_{GS} for low and high V_{DS}.

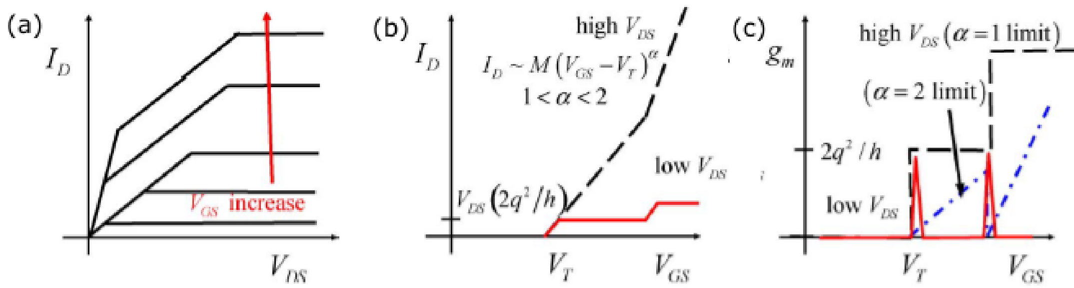

FIGURE 2.11: (a) Output characteristic of a 1D ballistic MOSFET with two subbands at $T = 0\,\mathrm{K}$. (b) Transfer characteristic of a 1D ballistic MOSFET with two subbands for low and high V_{DS} at $T = 0\,\mathrm{K}$. (c) $g_{\mathrm{m}} - V_{\mathrm{GS}}$ characteristic of a 1D ballistic MOSFET with two subbands for low and high V_{DS} at $T = 0\,\mathrm{K}$. Figures taken from [8].

2.2 Low-Frequancy Noise Theory

In this thesis, the term "noise" refers to random fluctuations of the drain current, caused by random fluctuations in the physical transport process through the channel of the transistors. The reason for the fluctuations in the transport process is trap states at the channel-oxide interface (interface traps) and deeper in the oxide (border traps), which can capture and emit electrons from and to the channel.
This section starts off by introducing some general noise properties and later on focuses on two particularly important types of noise: $1/f$ noise and RTS noise. $1/f$ noise is expected to be dominant in devices with a relatively large gate area containing many trap states. As the gate area decreases due to transistor scaling, the amount of trap states in the gate oxide also decreases. Eventually, the charge transport through the channel is limited by only one (or very few) active traps, resulting in RTS noise.

2.2.1 General Noise Properties

A randomly fluctuating current can be expressed in the time domain as [10]:

$$I(t) = \bar{I} + i_{\mathrm{n}}(t), \tag{2.25}$$

where $\bar{I}$ is the average current and $i_{\mathrm{n}}(t)$ is the fluctuating part. The value of $i_{\mathrm{n}}(t)$ is random at every time, so $i_{\mathrm{n}}(t)$ can be viewed as a random variable $X(t)$. This random variable is described by a squared quantity, the noise "power", which is defined as [10]:

$$\overline{X^2(t)} = \lim_{T\to\infty} \frac{1}{T} \int_0^T X^2(t)\mathrm{d}t = \int_0^\infty S_X(f)\mathrm{d}f. \tag{2.26}$$

S is called the noise PSD. Depending on S, the noise can be categorized into white noise (S independent of f), $1/f$ noise ($S \sim 1/f$) and RTS noise ($S \sim 1/f^2$).

2.2.2 $1/f$ Noise

$1/f$ noise is the common name for current fluctuations with a PSD proportional to $1/f^\beta$, with β close to 1, and it can be observed in MOSFETs in the low-frequency part of the spectrum (10^0 to 10^4 Hz). There are essentially two mechanisms (or a combination of both) that can cause $1/f$ noise: number fluctuations and mobility fluctuations [10].
In the number fluctuation model, $1/f$ noise is caused by many active, non-interacting traps, each trapping and detrapping electrons from and to the channel and contributing with a Lorentzian type of PSD ($S \sim 1/f^2$). Figure 2.12 shows the superposition of four Lorentzians, giving rise to a total spectrum that exhibits a $1/f$ dependence.

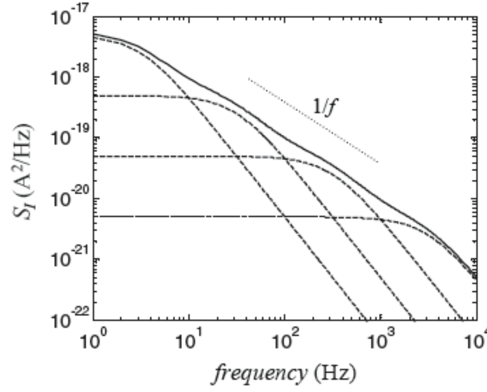

FIGURE 2.12: Superposition of four Lorentzians, resulting in total $1/f$ spectrum. Figure taken from [10].

In 1957, A. L. WcWhorter developed a model (based on number fluctuations), in which the trapping and detrapping events are described by elastic tunneling transitions, and following this approach, a $1/f$ noise spectrum is obtained for a trap density that is uniform in energy and in distance from the channel-oxide interface [10].
For a trap density that is non-uniform in depth, the $1/f^\beta$ noise exponent (β) deviates from one. $\beta > 1$ is expected from a trap density that is lower close to the channel-oxide interface than in the interior of the gate oxide and $\beta < 1$ for the opposite case [11].
It was later found that a trapped charge carrier also affects the channel mobility through Coulomb interaction. The so-called correlated mobility fluctuations give a correction to the number fluctuation noise model [10].
Taking number fluctuations and correlated mobility fluctuations into account, the drain current noise PSD normalized by the drain current squared ($S_{I_\mathrm{D}}/I_\mathrm{D}^2$) and the input gate voltage noise PSD ($S_{V_\mathrm{G}} = S_{I_\mathrm{D}}/g_\mathrm{m}^2$) can be expressed as (valid in strong inversion) [12]:

$$\frac{S_{I_\mathrm{D}}}{I_\mathrm{D}^2} = \left(1 + \frac{\alpha\mu_\mathrm{n}C_\mathrm{ox}I_\mathrm{D}}{g_\mathrm{m}}\right)^2 \left(\frac{g_\mathrm{m}}{I_\mathrm{D}}\right)^2 S_{V_\mathrm{fb}} \tag{2.27}$$

$$S_{V_\mathrm{G}} = S_{V_\mathrm{fb}}\left(1 + \alpha\mu_0 C_\mathrm{ox}(V_\mathrm{GS} - V_\mathrm{T})\right)^2, \tag{2.28}$$

where α is a scattering coefficient and μ_0 is the low-field electron mobility. The terms containing α and μ are used to describe the correlated mobility fluctuations.
S_{V_fb} is the flat band voltage noise PSD and depends in essence on the physical trapping

and detrapping mechanism. For elastic tunneling processes according to the McWhorter model, $S_{V_{\mathrm{fb}}}$ is given by [12]:

$$S_{V_{\mathrm{fb}}} = \frac{q^2 k_{\mathrm{B}} T \lambda N_{\mathrm{T}}}{W_{\mathrm{G}} L_{\mathrm{G}} C_{\mathrm{ox}}^2 f^{\beta}}, \tag{2.29}$$

with N_{T} being the trap density at the semiconductor Fermi level. λ is the tunneling attenuation length, which can be derived as [10]:

$$\lambda = \left(\frac{4\pi}{h} \sqrt{2m^* \Phi_{\mathrm{B}}} \right)^{-1}, \tag{2.30}$$

where Φ_{B} is the energy barrier for the elastic tunneling transitions.

A theory that relies purely on mobility fluctuations of the charge carriers in the channel is the Hooge model. According to this model, $S_{I_{\mathrm{D}}}/I_{\mathrm{D}}^2$ can be described as [12]:

$$\frac{S_{I_{\mathrm{D}}}}{I_{\mathrm{D}}^2} = \frac{q \alpha_{\mathrm{H}} V_{\mathrm{DS}} \langle \mu_{\mathrm{n}} \rangle}{f L_{\mathrm{G}}^2 I_{\mathrm{D}}}, \tag{2.31}$$

where $\alpha_{\mathrm{H}} \approx 10^{-4} - 10^{-6}$ is the Hooge parameter and $\langle \mu_{\mathrm{n}} \rangle$ is the mean electron mobility. As apparent from equation 2.27 and 2.31, $S_{I_{\mathrm{D}}}/I_{\mathrm{D}}^2 \sim (g_{\mathrm{m}}/I_{\mathrm{D}})^2$ when the $1/f$ noise is due to number fluctuations, whereas $S_{I_{\mathrm{D}}}/I_{\mathrm{D}}^2 \sim 1/I_{\mathrm{D}}$ when mobility fluctuations are the dominant noise mechanism.

The $1/f$ noise in MOSFETs is generally generated in the channel and in the source/drain access series resistances. The total drain current noise PSD ($S_{I_{\mathrm{D,tot}}}$) is given by [10]:

$$S_{I_{\mathrm{D,tot}}} = \frac{S_{I_{\mathrm{D,ch}}} + g_{\mathrm{D}}^2 R_{\mathrm{SD}}^2 S_{I_{R_{\mathrm{SD}}}}}{\left(1 + g_{\mathrm{D}} \left(R_{\mathrm{SD}} + R_{\mathrm{L}}\right)\right)^2}, \tag{2.32}$$

where $S_{I_{\mathrm{D,ch}}}$ is the drain current noise PSD in the channel, $S_{I_{R_{\mathrm{SD}}}}$ is the drain current noise PSD in the source/drain resistance, R_{SD} is the total source/drain resistance ($R_{\mathrm{S}} = R_{\mathrm{D}} = R_{\mathrm{SD}}/2$) and R_{L} is the load resistance between the source and the drain.

As opposed to $S_{I_{R_{\mathrm{SD}}}}$, $S_{I_{\mathrm{D,ch}}}$ depends on the gate area ($A = W_{\mathrm{G}} L_{\mathrm{G}}$) [10]:

$$S_{I_{\mathrm{D,ch}}} \sim \frac{I_{\mathrm{D}}^2}{f W_{\mathrm{G}} L_{\mathrm{G}} (V_{\mathrm{GS}} - V_{\mathrm{T}})} \tag{2.33}$$

$$S_{I_{R_{\mathrm{SD}}}} \sim \frac{I_{\mathrm{D}}^2}{f}. \tag{2.34}$$

If the $1/f$ noise originating from the channel is dominant, the total drain current noise PSD increases with decreasing gate area. If the $1/f$ noise originating from the source/drain resistance is dominant, the total drain current PSD is independent of the gate area.

2.2.3 RTS Noise

Apart from a performance boost and a cost reduction, transistor scaling also leads to a decrease in the number of trap states in the gate oxide. In highly scaled transistors, the charge transport through the channel can be affected by only one (or very few) active traps. In the case of one active trap, the drain current switches between two discrete levels ("two-level RTS noise"), as illustrated in figure 2.13 [13].

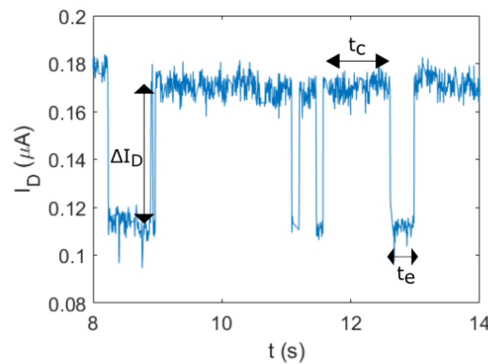

FIGURE 2.13: Two-level RTS noise in one of the considered transistors.

The switching process is characterized by three parameters: the characteristic time that the current spends in the high current level (called capture time constant (τ_{c})), the characteristic time that the current spends in the low current level (called emission time constant (τ_{e})) and the RTS noise amplitude (ΔI_{D}).

The switching of the current between two levels is usually interpreted in terms of trapping

and detrapping of a single electron (in an nMOSFET) by a trap state at the channel-oxide interface or in the oxide at tunneling distance from the inversion layer [13].
The first part of this subsection focuses on the theoretical description of the time constants and in particular on how the time constants change with gate voltage (or drain current). The second part highlights the RTS noise amplitudes, showing how they will be affected by the device size and operation conditions.

2.2.3.1 RTS Noise Time Constants

For a single, non-interacting trap, the times that the current spends in the high and in the low current level (t_{c} and t_{e}, respectively) are exponentially distributed [13]:

$$P_{\mathrm{c,e}}(t_{\mathrm{c,e}}) = \frac{1}{\tau_{\mathrm{c,e}}} \exp\left(-\frac{t_{\mathrm{c,e}}}{\tau_{\mathrm{c,e}}}\right) \tag{2.35}$$

and each of the two distributions is characterized by the corresponding time constant (τ_{c} or τ_{e}). $P_{\mathrm{c,e}}\mathrm{d}t_{\mathrm{c,e}}$ is the probability that the high/low current level will not make a transition during time $t_{\mathrm{c,e}}$, but will make a transition during time $t_{\mathrm{c,e}} + \mathrm{d}t_{\mathrm{c,e}}$.

Shockley-Read-Hall Framework and Beyond

It is usually observed that the time constants of trap states change with the applied voltages (especially the gate voltage). Most commonly, the SRH theory is used to describe this observation [13].
The original SRH theory (see [14] for more details) explains the recombination of electrons and holes in semiconductors via a trap state that is located in the band gap of the semiconductor at energy E_{T}. This trap state may capture or emit an electron from or to the bottom of the conduction band (E_{C}) and upon the electron capture or emission process, the trap state changes its charging status from neutral to negative (for electron capture) or from negative to neutral (for electron emission), assuming that the trap state was uncharged initially. The trap state may also capture or emit an electron from or to the top of the valence band (E_{V}), which can be viewed as a hole emission or a hole capture process (see figure 2.14). The excess/missing energy for the transitions is released/supplied by the emission/absorption of photons or phonons.

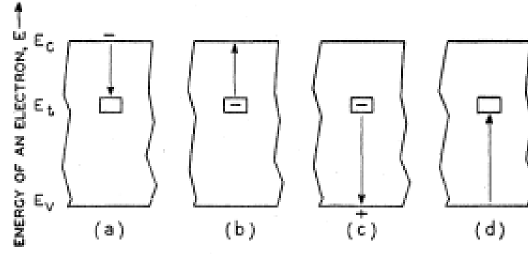

FIGURE 2.14: SRH processes: (a) electron capture, (b) electron emission, (c) hole capture, (d) hole emission. Figure taken from [14].

As such, the original SRH theory only considers energetic transitions. Therefore, this theory is only applicable for interface traps with a trap energy that is located in the semiconductor band gap. For border traps, the original SRH theory has been modified to account for the physical distance between the initial energy state and the trap state, using an additional elastic tunneling term. It is, however, questionable if this modified theory is strictly valid [15] (especially when E_T is not within the semiconductor bandgap).
Considering a elastic tunneling and a SRH process, the capture time constant is given by [16, 17]:

$$\tau_{\mathrm{c}} = \frac{1}{n\sigma_{\mathrm{n}}v_{\mathrm{th}}}, \tag{2.36}$$

with n being the electron concentration, σ_{n} the electron capture cross section and v_{th} the thermal velocity of the electrons. The electron capture cross section contains a tunneling coefficient (Θ) to account for the elastic tunneling process into the oxide ($\sigma_{\mathrm{n}} = \sigma_0\Theta$). Apart from the above-mentioned assumptions, the expression for τ_{c} is based on on an electron gas in equilibrium with the lattice, with negligible carrier heating in the lateral electric field and it neglects inversion layer quantization.
The ratio of the capture and emission time constant, called the mark space ratio, can be derived from the principle of detailed balance (equal number of excitations and de-excitations) [16]:

$$\frac{\tau_{\mathrm{c}}}{\tau_{\mathrm{e}}} = g \cdot \exp\left(\frac{E_{\mathrm{T}} - E_{\mathrm{F}}}{k_{\mathrm{B}}T}\right) = \frac{1 - f_{\mathrm{T}}}{f_{\mathrm{T}}}, \tag{2.37}$$

where f_{T} is the trap occupancy, given by the Fermi-Dirac distribution. The trap degeneracy (g) is usually not known and assumed to be unity [13].

Using $n = N_{\mathrm{C}} \cdot \exp[-(E_{\mathrm{C}} - E_{\mathrm{F}})/k_{\mathrm{B}}T]$, with N_{C} being the density of states in the conduction band, τ_{e} can be expressed as [13]:

$$\tau_e = \frac{1}{N_{\mathrm{C}} v_{\mathrm{th}} \sigma_{\mathrm{n}}} \exp\left(\frac{E_{\mathrm{C}} - E_{\mathrm{T}}}{k_{\mathrm{B}}T}\right). \tag{2.38}$$

From the (modified) SRH theory it is expected that τ_{c} decreases with increasing V_{GS} (or I_{D}), since n increases with increasing V_{GS}. In a first order approximation (electron velocity does not change with V_{GS}), $n \sim I_{\mathrm{D}}$. τ_{e}, on the other hand, increases with V_{GS}. This is the typical behavior of an acceptor type of trap in nMOSFETs, changing its charging status from neutral to negative upon electron capture. For an acceptor trap, the mark space ratio decreases exponentially with increasing V_{GS} assuming E_{T} decreases linearly when V_{GS} is increased. There are also traps with a different polarity, called donor traps for nMOSFETs, changing their charging status from positive to neutral upon electron capture. For these traps, the mark space ratio increases with increasing V_{GS} [13].

It is often found that both time constants are thermally activated (time constants change with temperature). To account for the energy barrier (E_{B}) that the electrons have to overcome for the capture process, the SRH theory can be further modified by multiplying σ_{n} with the additional term $R = \exp(-E_{\mathrm{B}}/k_{\mathrm{B}}T)$ and equation 2.36 can be rewritten as [17]:

$$\tau_{\mathrm{c}} = \frac{\exp\left(\frac{E_{\mathrm{B}}}{k_{\mathrm{B}}T}\right)}{n \sigma_{\mathrm{n}} v_{\mathrm{th}}}, \tag{2.39}$$

again assuming a non-quantized inversion layer.
E_{B} can be interpreted as a strong lattice relaxation in the gate oxide during the electron capture process. The lattice relaxation occurs by the emission of multiple phonons, dissipating part of the excess energy of the trapped carrier. This non-elastic, multi-phonon assisted tunneling takes into account electron-phonon interactions, causing a rearrangement of the atoms in the defect structure due to the electrostatic coupling between the nuclei and the trapped electron. Such a structural rearrangement, which requires the atoms to leave their equilibrium positions, is associated with the system overcoming the energy barrier E_{B} [18].

In the multi-phonon assisted tunneling picture, the emission time constant is given by a change in enthalpy (ΔH_{CT}) and in entropy ($T\Delta S$) [17]:

$$\tau_e = \frac{\exp\left(-\frac{\Delta S}{k_{\mathrm{B}}}\right)}{N_{\mathrm{C}} v_{\mathrm{th}} \sigma_{\mathrm{n0}}} \exp\left(\frac{\Delta H_{\mathrm{CT}} + E_{\mathrm{B}}}{k_{\mathrm{B}} T}\right). \tag{2.40}$$

Elastic Tunneling Transitions

It is expected that SRH processes (in combination with elastic tunneling) and/or multi-phonon assisted tunneling processes dominate around RT. At lower temperatures, however, elastic tunneling at a constant energy around the semiconductor Fermi level becomes more significant [13]. A theory that relies purely on elastic tunneling to and from interface or border trap states is the above-mentioned McWhorter model for $1/f$ noise. In the elastic tunneling model, the time constant τ is connected to the trap depth in the oxide (z_{T}) according to [10]:

$$\tau = \tau_0(E) \cdot \exp\left(\frac{z_{\mathrm{T}}}{\lambda}\right), \tag{2.41}$$

where $\tau_0(E)$ is the time constant for a trap located at the channel-oxide interface (usually assumed to be energy independent and having a value of 10^{-10} s).

2.2.3.2 RTS Noise Amplitude Behavior

Apart from the time constants, the RTS noise amplitude can be evaluated. It is, however, difficult to understand its dependence on the bias conditions and also the large spread in values that has been observed for traps in different transistors and even for different traps in the same transistor [13].

Assuming a uniform channel and a uniform change in current, the relative RTS noise amplitude can be expressed as [19, 20]:

$$\frac{\Delta I_{\mathrm{D}}}{I_{\mathrm{D}}} = -\frac{1}{WL}\left(\frac{1}{\frac{k_{\mathrm{B}}T}{q(C_{\mathrm{ox}}+C_{\mathrm{d}})} + n} \pm \alpha\mu_{\mathrm{n}}\right), \tag{2.42}$$

with C_{D} being the depletion capacitance and n_{S} the surface electron density. The first term accounts for electron number fluctuations while the second term describes mobility fluctuations upon the electron capture process (see section 2.2.2 for further information about mobility fluctuations). This equation was derived for planar MOSFETs and cannot explain relative RTS noise amplitudes larger than 1 % [13].

A theory that can describe large RTS noise amplitudes in planar MOSFETs (with III-V channels) is the "hole-in-the-inversion-layer" model, which bases on local perturbations of the channel. According to this model, a trapped electron creates a "hole" in the inversion layer through Coulomb repulsion and the size of the hole determines the RTS noise amplitude. A smaller gate voltage leads to a larger hole size and therefore to a bigger RTS noise amplitude. From this theory, it is expected that the use of high-mobility materials in MOSFETs will increase the RTS noise amplitude compared with silicon devices [21].

Three-dimensional (3D) atomistic simulations of silicon-on-insulator FinFETs, which consider acceptor interface traps that cause number fluctuations, also show quite large relative RTS noise amplitudes of up to 20 %. It is found that the relative RTS noise amplitude depends on the trap location. A trap in the middle region between the source and the drain, which is located near the bottom of the fin has the strongest impact. Furthermore, it is observed that the RTS noise amplitude increases with decreasing gate voltage (above V_{T}) and that scaling of the oxide thickness reduces the RTS noise amplitude [22].

Chapter 3

Device Fabrication

In this section, the fabrication of the InGaAs NW MOSFETs that are used for the LF noise measuremennts will be outlined. The device fabrication was performed by others and is described in more detail in [23].

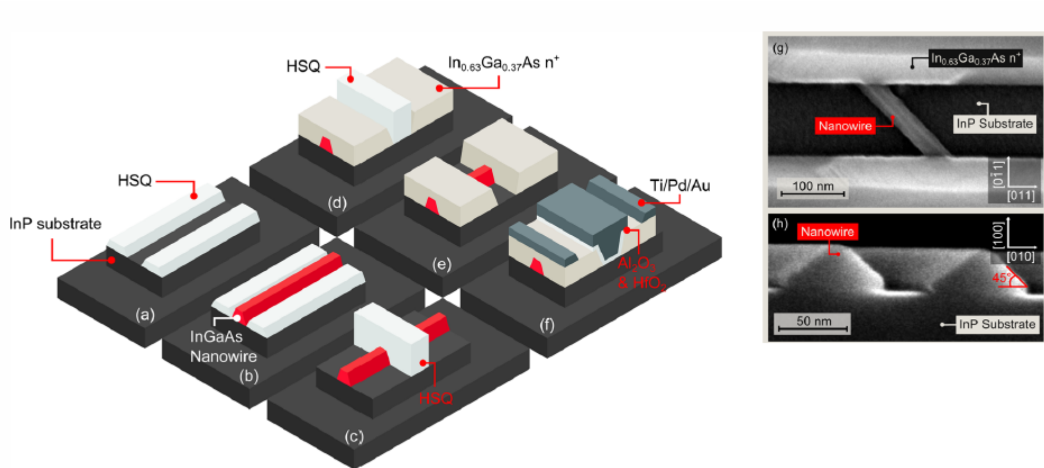

FIGURE 3.1: (a) Two HSQ strips are patterned on a Fe doped InP substrate using EBL. (b) $In_{0.63}Ga_{0.37}As$ is regrown by MOCVD, resulting in the formation of a nanowire between the strips. (c) An HSQ line is patterned across the nanowire to define the gate length. (d) Highly doped $In_{0.63}Ga_{0.37}As$ is regrown by MOCVD, forming the source and drain regions. (e) The HSQ line is stripped by buffered oxide etch. (f) Final device with the deposited gate oxide (Al_2O_3/HfO_2), the gate metal (Ti/Pd/Au) and the source and drain contacts (Ti/Pd/Au). (g) SEM image of the device corresponding to stage (e). (h) Cross-sectional SEM image of the nanowires showing a facet angles of 45°. Figures taken from [23].

The process starts off by spin-coating HSQ on a semi-insulating (100) Fe doped InP substrate. Next, the HSQ layer is patterned by electron-beam lithography (EBL) and the non-exposed area is dissolved in TMAH. The remaining HSQ structures are two parallel strips with a narrow spacing between them (see figure 3.1 (a)).
Since HSQ transforms into SiO_x when cured by electron beam exposure, the strips can be used as a growth mask during the subsequent metal-organic chemical vapor deposition (MOCVD) growth of 15 nm $In_{0.63}Ga_{0.37}As$. Between the two strips, a NW with the composition of $In_{0.85}Ga_{0.15}As$ (determined by optical characterization) is formed. The NW is parallel to the surface and oriented along the [001] direction (see figure 3.1 (b)). The facet angles are 45°, as shown in the cross-sectional scanning electron microscope (SEM) image in figure 3.1 (h).
After removing the strips, the NW height and width are reduced by several cycles of digital etching through ozone oxidation and diluted HCl etching. The next step is to pattern an HSQ line across the nanowire that serves to define the gate length of the transistor (see figure 3.1 (c).
Then, a second MOCVD regrowth step of 40 nm $In_{0.63}Ga_{0.37}As$ n^{++} ($N_D = 5{\cdot}10^{19}\,cm^{-3}$) is performed, as shown in figure 3.1 (d). The two highly doped regions form the source and the drain of the device.
After stripping the HSQ line (depicted in figure 3.1 (e)), the NW is passivated by sulfur treatment. Figure 3.1 (h) shows an SEM image of the device in stage (e). Subsequently, the gate oxide (10 cycles of aluminum oxide (Al_2O_3) and 60 cycles of HfO_2) is deposited by atomic layer deposition (ALD), followed by the evaporation and deposition of Ti/Pd/Au as the gate metal by lift-off. Afterwards, the gate oxide is etched in the source and drain region using the gate metal as a mask. The source and drain contacts are then formed by the deposition of Ti/Pd/Au using lift-off. Gold pad metallization by lift-off completes the process. The final device is shown in figure 3.1 (f).

Chapter 4

Experimental Setup

This section describes the experimental setup for the $1/f$ and RTS noise measurements. For both types of measurements, a Lake Shore Cryotronics CRX-4K probe station was used to contact the transistors. The temperature in the probe station was controlled by a liquid helium operated compressor whereas the air pressure in the probe station was regulated by a vacuum turbo pump.
When performing the 1/f noise measurements, a low-noise current preamplifier (model SR570 from Stanford Research Systems) was utilized to supply the contacted transistor with a constant drain voltage of $V_{\mathrm{DS}} = 50\,\mathrm{mV}$ and to amplify the drain current signal (determined by the sensitivity of the preamplifier (S)). The low-noise current preamplifier works as a transimpedance amplifier, converting the input drain current signal into an output drain voltage signal. The output of the current preamplifier was connected to the input of a lock-in amplifier (model SR830 from Stanford Research Systems), which was used to analyze the drain voltage signal. The lock-in amplifier, connected to a laptop with the corresponding measurement software, extracts the AC component of the drain voltage signal (the fluctuating part of the signal due to noise) and measures its frequency dependence ($v_{\mathrm{n}}(f)$) using a narrow band pass filter. The sought-after drain current noise PSD of the transistor (as a function of frequency) is given by: $S_{I_{\mathrm{D}}}(f) = S \cdot v_{\mathrm{n}}^2(f)$. A Keysight B2912A source measure unit (SMU), also connected to the laptop and controlled by the measurement software, was utilized to apply the gate voltages and to monitor the source current during the measurements.

Two types of $1/f$ noise measurements were performed: frequency sweeps and 10 Hz measurements. For the first type of $1/f$ noise measurement, the drain voltage noise was measured at different lock-in amplifier frequencies (logarithmically swept from $f = 10^1$ Hz to $f = 10^4$ Hz). At each frequency, v_n was measured 55 times and the average value of these 55 measurements constituted the data point for the given frequency. This procedure was repeated for different gate biases. For the second type of measurement, v_n was measured at a fixed lock-in amplifier frequency of $f = 10\,\mathrm{Hz}$. 16 datapoints (meaning 16 repetitions of the measurement) were recorded at each gate voltage.

For the RTS noise measurements, only the SMU was utilized to apply constant drain and gate voltages and to measure the drain current over time.

External Noise From the $1/f$ Noise Measurement Setup

This section lists the external noise contributions arising from the $1/f$ noise measurement setup and compares them with the internal drain current noise PSD of the transistors, which is in the order of $S_{I_\mathrm{D}} \approx 10^{-16} - 10^{-24}\,\mathrm{A}^2/\mathrm{Hz}$, depending on V_GS and f. The goal is to show that the setup described above can be used for the $1/f$ noise measurements.

First of all, external noise arises from the electrical resistance (R) of the cables and the probe tip. The total electrical resistance has an estimated magnitude of $R = 1 - 10\,\Omega$, resulting in thermal noise (belonging to white noise) PSD of $S_{I,R} = 4k_\mathrm{B}T/R \approx 10^{-20} - 10^{-21}\,\mathrm{A}^2/\mathrm{Hz}$ at RT. Clearly, the thermal noise PSD of the resistance is in the same order as the lower values of S_{I_D}, making them untrustworthy. However, for gate voltages not more than 0.2 V below the threshold voltage, S_{I_D} is well above the thermal noise limit for all considered frequencies.

The current preamplifier causes $1/f$ noise at low frequencies (until $f \approx 10^2$ Hz) and white noise at higher frequencies. The white noise depends on the selected sensitivity, where a low sensitivity means a high amplification and vice versa. The higher the sensitivity, the higher the white noise. For the frequency sweeps, the sensitivity was changed from $S = 1 \cdot 10^{-7}\,\mathrm{A/V}$ to $S = 2 \cdot 10^{-6}\,\mathrm{A/V}$ (high bandwidth setting), or from $S = 1 \cdot 10^{-7}\,\mathrm{A/V}$ to $S = 2 \cdot 10^{-6}\,\mathrm{A/V}$ (low noise setting), when increasing the gate voltage, leading to a white noise PSD of the current preamplifier of $S_{I,\mathrm{CPA}} \approx 3.6 \cdot 10^{-25} - 3.6 \cdot 10^{-23}\,\mathrm{A}^2/\mathrm{Hz}$. These values are well below the drain current noise PSD of the transistors, given that

the gate voltage is not too small. The $1/f$ noise PSD of the current preamplifier at low frequencies is approximately one or two orders of magnitude higher than the white noise, but is still below S_{I_D}. For the 10 Hz measurements, even lower sensitivities could be chosen, resulting in a lower white noise PSD and in a lower $1/f$ noise PSD. At $f = 10\,\text{Hz}$, the $1/f$ noise PSD is the dominant contribution from the current preamplifier, but it is far below the drain current noise PSD of the transistors. The current preamplifier also sets the drain voltage of the transistors with some inaccuracy, which introduces more external noise. However, this contribution is extremely small compared with the $1/f$ and white noise PSD of the preamplifier.

The noise of lock-in amplifier is specified with $6\,\text{nV}/\sqrt{\text{Hz}}$, which is at least three orders of magnitude below the measured drain voltage noise of transistors for all V_GS and f. Consequently, this noise is negligible.

Finally, the SMU introduces external noise due to an inaccuracy when applying the gate voltages. As stated above, the resulting noise is extremely small compared with all the other external noise contributions and can therefore be neglected.

Limitations of the RTS Noise Measurement Setup

The RTS noise measurements are mainly limited by the time resolution of the SMU. The smallest time resolution is $10\,\mu\text{s}$, meaning that only RTS noise with time constants larger than this value can be resolved. On the other hand, RTS noise with (very) large time constants can theoretically be measured. In practice however, the RTS noise is often not stable enough to record a sufficient amount of transitions between the high and the low drain current level and therefore the time constants cannot be extracted.

Chapter 5

Results and Discussion

5.1 DC Performance

The devices used for the noise measurements show excellent DC performance, with a peak transconductance of $g_{\mathrm{m}} = 2.9\,\mathrm{mS}/\mu\mathrm{m}$, a minimum subthreshold swing of $S = 77\,\mathrm{mV/decade}$ and an on-current of $I_{\mathrm{D,on}} = 565\,\mu\mathrm{A}/\mu\mathrm{m}$ (at $I_{\mathrm{D,\,off}} = 100\,\mathrm{nA}/\mu\mathrm{m}$), all at $V_{\mathrm{DS}} = 0.5\,\mathrm{V}$. $I_{\mathrm{D,on}}$ is the highest reported on-current for any transistor [24]. Furthermore, it was shown that these transistors operate in the quasi-ballistic regime with a transmission of about 70 %, which was obtained from quantized conductance measurements at $T = 10\,\mathrm{K}$, and was shown to be valid also at RT [23].

In the following sections, the results of the low-frequency noise measurements on transistors with varying gate lengths $L_{\mathrm{G}} = 50 - 85\,\mathrm{nm}$ and gate widths $W_{\mathrm{G}} = 27 - 100\,\mathrm{nm}$ will be presented and discussed.

5.2 $1/f$ Noise at RT

5.2.1 10 Hz Measurements

To determine whether the $1/f$ noise originates from the channel or from the source/drain resistance, the impact of the gate area scaling on the drain current noise PSD can be studied. Figure 5.1 (a) shows the normalized drain current noise PSD ($S_{I_{\mathrm{D}}}/I_{\mathrm{D}}^2$) at a fixed frequency of 10 Hz plotted against the drain current for three different devices

with gate areas of $A = 1350\,\text{nm}^2$, $6300\,\text{nm}^2$ and $8500\,\text{nm}^2$. It is clearly visible that the normalized drain current noise PSD is inversely proportional to the gate area (at a fixed drain current), meaning that the $1/f$ noise stems from the channel and not from the source/drain resistance. In the latter case, the normalized drain current noise PSD would be independent of A (see equation 2.32, 2.33 and 2.34).

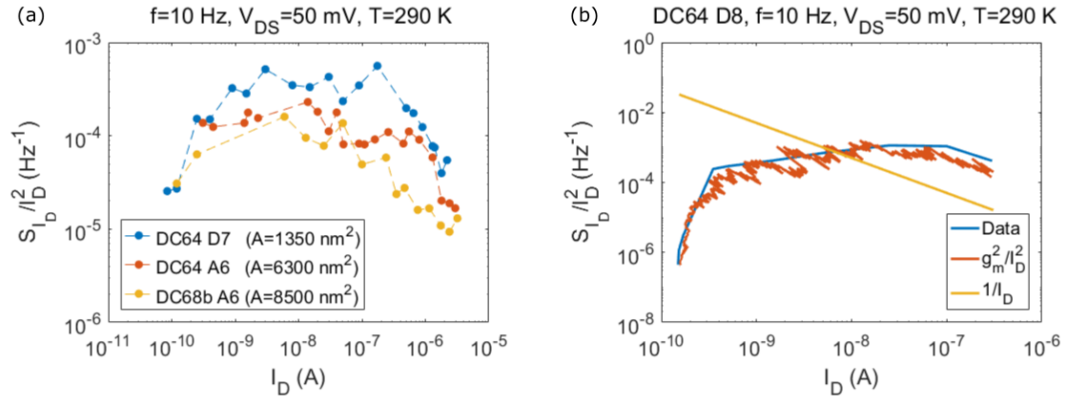

FIGURE 5.1: (a) Impact of the gate area scaling on S_{I_D}/I_D^2, showing that the $1/f$ noise originates from the channel. (b) In all devices, $S_{I_D}/I_D^2 \sim (g_m/I_D)^2$ (depicted here for one single device with $L_G = 50\,\text{nm}$ and $W_G = 27\,\text{nm}$), indicating that the $1/f$ noise is due to number fluctuations.

The next step is to analyze the dominant noise mechanism. If S_{I_D}/I_D^2 is proportional to $(g_m/I_D)^2$, the $1/f$ noise is due to number fluctuations. If, on the other hand, S_{I_D}/I_D^2 is proportional to $1/I_D$, mobility fluctuations are the dominant noise mechanism (see equation 2.27, 2.28 and 2.31).

In all of the measured transistors (independent of L_G and W_G), S_{I_D}/I_D^2 is proportional to $(g_m/I_D)^2$, so the $1/f$ noise is due to number fluctuations. This is depicted in figure 5.1 (b) for one single transistor with $L_G = 50\,\text{nm}$ and $W_G = 27\,\text{nm}$. Interestingly, the opposite trend was observed in a recent paper about low-frequency noise in near-ballistic InGaAs NW MOSFETs (similar devices), where they found that mobility fluctuations were the dominant noise mechanism [25].

To compare the noise performance of these devices with other devices, the gate area normalized input gate voltage noise PSD ($S_{V_G}L_GW_G$) is calculated, where $S_{V_G} = S_{I_D}/g_m^2$, as defined in the theory section. In figure 5.2, $S_{V_G}L_GW_G$ is plotted against the overdrive voltage ($V_{ov} = V_{GS} - V_T$) for four different transistors.

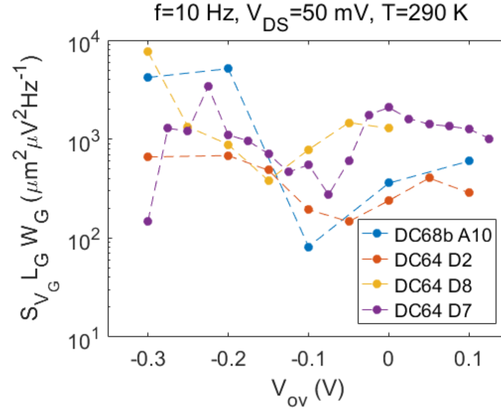

FIGURE 5.2: $S_{V_G} L_G W_G$ plotted against V_{ov} for transistor DC68b A10 ($L_G = 70\,$nm, $W_G = 78\,$nm), DC64 D2 ($L_G = 50\,$nm, $W_G = 27\,$nm), DC64 D8 ($L_G = 50\,$nm, $W_G = 27\,$nm) and DC64 D7 ($L_G = 50\,$nm, $W_G = 27\,$nm). The low values indicate the feasibility of a high-quality Al_2O_3/HfO_2 gate oxide on InGaAs.

The values of $S_{V_G} L_G W_G$ are mostly between 100 and 1000 μm$^2\mu$V^{2}Hz^{-1} (around $V_{ov} = 0\,$V), with a minimum value of 80 μm$^2\mu$V^{2}Hz^{-1}. Assuming that the trapping and de-trapping mechanism is elastic tunneling of electrons to and from the trap states, the corresponding minimum trap density can be calculated with the aid of equation 2.28 and 2.29: $N_T = 9 \cdot 10^{18}\,\text{cm}^{-3}\text{eV}^{-1}$. For this calculation, a tunneling attenuation length of $\lambda = 0.13\,$nm (using $m^*_{Al_2O_3} = 0.23 \cdot m_e$ [27] and $\Phi_B = 2.4\,$eV [28]) and an oxide capacitance per unit area of $C_{ox} = 0.029\,\text{F/m}^2$ (using EOT = 1.2 nm) are utilized.
These values are among the lowest for III-V NW MOSFETs (slightly worse than the ones reported for near-ballistic InGaAs NW MOSFETs [25] and comparable to the ones for vertical InAs NW MOSFETs [26]), demonstrating the feasibility of a high-quality Al_2O_3/HfO_2 gate oxide on InGaAs NW structures. On top of that, the low values suggest that the electron transport through the channel is limited by relatively few traps. At some gate voltages (when only one of these few traps is active), RTS noise should be visible.

5.2.2 Frequency Sweeps

When measuring the drain current noise PSD as a function of frequency, the $1/f^\beta$ noise exponent (β) can be studied. In each transistor, S_{I_D} is measured as a function of f for a number of different overdrive voltages, ranging from $V_{ov} = -0.2\,$V to $V_{ov} = 0.3\,$V. β is obtained from the slope of a linear fit to the S_{I_D}-f curves in a log-log plot.
It is observed that in every transistor, β varies between 0.7 and 1.5 when sweeping

the overdrive voltage in the above mentioned range. This is depicted in figure 5.3 for transistor DC68b A6 ($L_G = 85\,\text{nm}, W_G = 100\,\text{nm}$) for three different overdrive voltages.

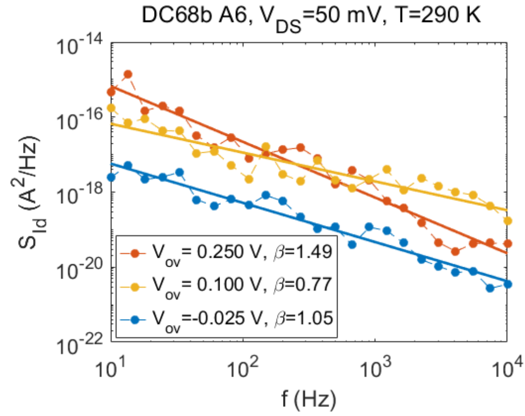

FIGURE 5.3: S_{I_D} plotted against f for three different overdrive voltages in transistor DC68b A6. The $1/f^\beta$ noise exponent (β) varies between 0.77 and 1.49 when sweeping the gate voltage. Assuming elastic tunneling, the varying β can be explained by different trap densities at different gate voltages.

In the elastic tunneling picture, a noise exponent different from one is obtained for a trap density that is non-uniform in depth. $\beta > 1$ results from a trap density that is lower close to the channel-oxide interface than in the interior of the gate oxide and $\beta < 1$ for the opposite case. For a trap density that is uniform in depth, $\beta = 1$ [11].

For each frequency (overdrive voltage fixed), the corresponding trap density (equation 2.28 and 2.29) and the corresponding trap depth (equation 2.41) can be calculated (assuming elastic tunneling). Figure 5.4 shows the trap density plotted against the trap depth for the same transistor and the same three overdrive voltages as in figure 5.3.

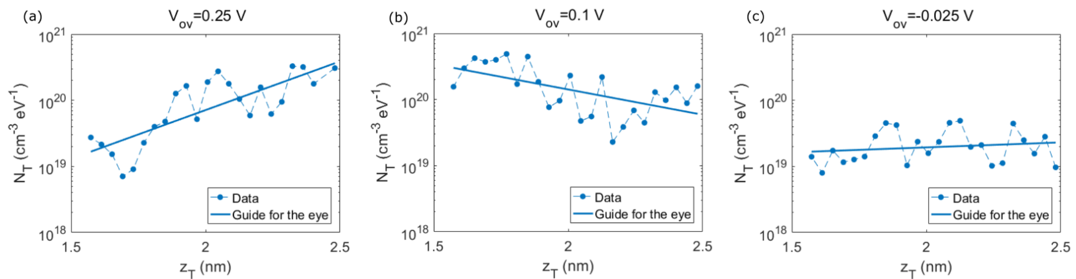

FIGURE 5.4: Trap densities for transistor DC68b A6 ($L_G = 85\,\text{nm}, W_G = 100\,\text{nm}$) at (a) $V_{ov} = 0.25\,\text{V}$, (b) $V_{ov} = 0.1\,\text{V}$ and (c) $V_{ov} = -0.025\,\text{V}$.

Indeed, the varying $1/f^\beta$ noise exponent can be traced back to different trap densities at different gate voltages.

5.3 $1/f$ Noise at T=11 K

To verify the main results/assumptions of the previous section, a few $1/f$ noise measurements were repeated at $T = 11\,\mathrm{K}$.
Again, it is found that in all of the transistors, the normalized drain current noise PSD at a fixed frequency of 10 Hz is proportional to $(g_\mathrm{m}/I_\mathrm{D})^2$ when plotted against the drain current, indicating that the $1/f$ noise is due to number fluctuations. This is exemplified in figure 5.5 (a) for transistor DC64 D2 ($L_\mathrm{G} = 50\,\mathrm{nm}$, $W_\mathrm{G} = 27\,\mathrm{nm}$).

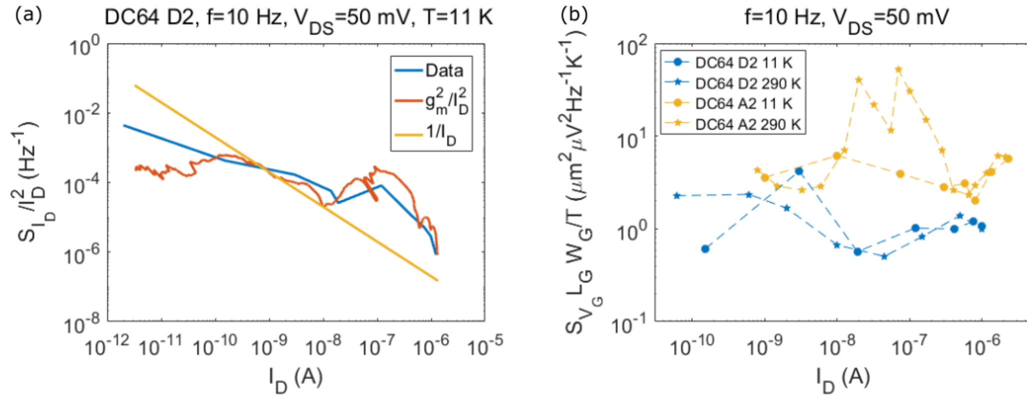

FIGURE 5.5: (a) Normalized drain current noise PSD plotted against the drain current for transistor DC64 D2, showing that $S_{I_\mathrm{D}}/I_\mathrm{D}^2 \sim (g_\mathrm{m}/I_\mathrm{D})^2$, which means that the $1/f$ noise at $T = 11\,\mathrm{K}$ is due to number fluctuations.(b) $S_{V_\mathrm{G}} L_\mathrm{G} W_\mathrm{G}$ scales with the temperature, supporting the elastic tunneling model.

If the trapping and detrapping mechanism is based on elastic tunneling as assumed before, $S_{V_\mathrm{G}} L_\mathrm{G} W_\mathrm{G}$ will decrease linearly with decreasing temperature (see equation 2.28 and 2.29). In figure 5.5 (b), $S_{V_\mathrm{G}} L_\mathrm{G} W_\mathrm{G}$ normalized by T is plotted against the drain current for two transistors (DC64 D2 ($L_\mathrm{G} = 50\,\mathrm{nm}$, $W_\mathrm{G} = 27\,\mathrm{nm}$) and DC64 A2 ($L_\mathrm{G} = 50\,\mathrm{nm}$, $W_\mathrm{G} = 90\,\mathrm{nm}$)) and two temperatures ($T = 290\,\mathrm{K}$ and $T = 11\,\mathrm{K}$). The two temperature curves for both transistors coincide reasonably well, giving some support for the elastic tunneling model.
From the frequency sweeps at $T = 11\,\mathrm{K}$, no further insights can be gained. Again, β varies between 0.7 and 1.5 when sweeping the overdrive voltage from $V_\mathrm{ov} = -0.2\,\mathrm{V}$ to $V_\mathrm{ov} = 0.3\,\mathrm{V}$.

5.4 RTS Noise at T=11 K

To analyze the impact of single trap states on the electron transport through the channel, RTS noise measurements at $T = 11\,\mathrm{K}$ were performed. RTS noise signals could be observed in $\sim 2/3$ of the considered devices, demonstrating that very few traps are responsible for the noise in these devices.

Figure 5.6 (a) shows an excerpt of the RTS noise signal for transistor DC68b D4 ($L_{\mathrm{G}} = 70\,\mathrm{nm}$, $W_{\mathrm{G}} = 31\,\mathrm{nm}$), which was biased with $V_{\mathrm{DS}} = 0.069\,\mathrm{V}$ and $V_{\mathrm{GS}} = 0.39\,\mathrm{V}$. At this bias point, the drain current switches between two discrete levels, meaning only one active trap. The time intervals in which the current is in the high current level are the capture times (t_{c}), and the time intervals in which the current is in the low current level are the emission times (t_{e}), as indicated in the figure.

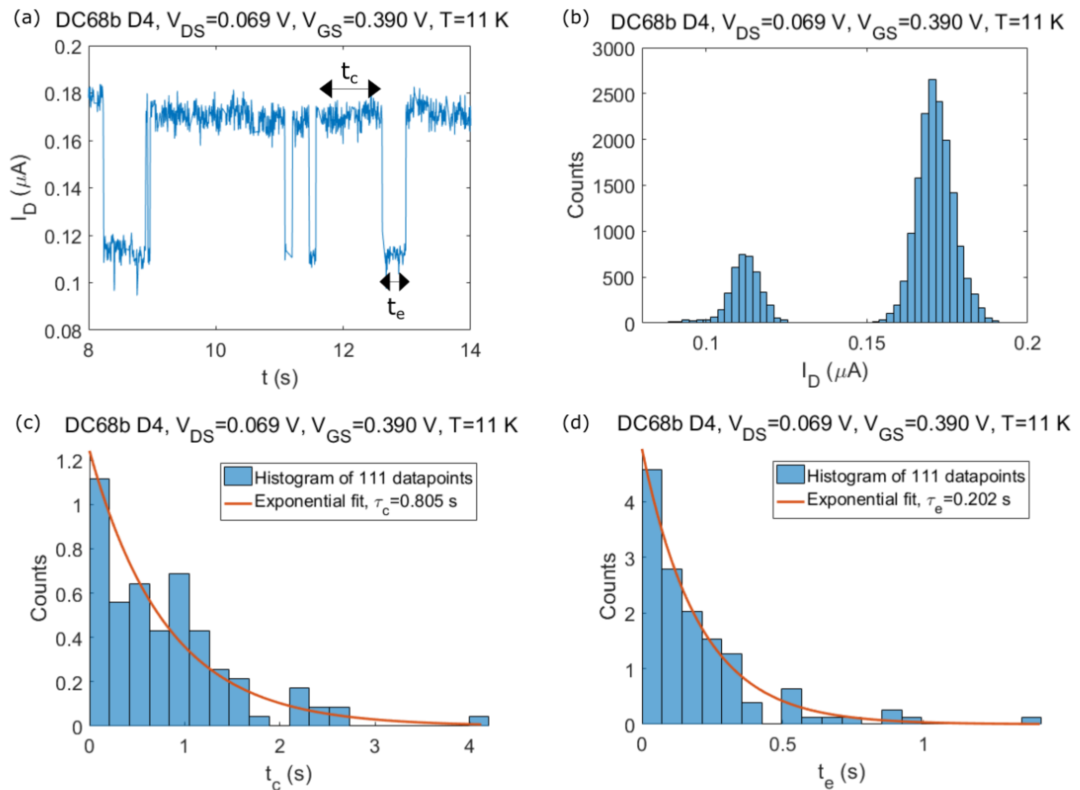

FIGURE 5.6: (a) Excerpt of the two-level RTS noise signal for transistor DC68b D4, biased with $V_{\mathrm{DS}} = 0.069\,\mathrm{V}$ and $V_{\mathrm{GS}} = 0.39\,\mathrm{V}$. (b) Histogram of the complete RTS noise signal. (c) Normalized histogram of the capture times, resulting in $\tau_{\mathrm{c}} = 0.805\,\mathrm{s}$. (d) Normalized histogram of the emission times, giving $\tau_{\mathrm{e}} = 0.202\,\mathrm{s}$.

A histogram of the complete RTS noise signal is depicted in figure 5.6 (b), showing

two distinctive peaks at two different drain currents, which again implies two-level RTS noise. The drain currents in the high and in the low current level follow a Gaussian distribution and the difference in the mean values can be associated with the RTS noise amplitude (ΔI_D). The higher peak at the higher drain current suggests that the current is predominantly in the high current level, leading to a larger τ_c as compared with τ_e. Figure 5.6 (c) and (d) show a normalized histogram of the capture and emission times, respectively. The capture and emission times are exponentially distributed, characterized by the time constants (τ_c or τ_e), according to equation 2.35. At this bias point, $\tau_c = 0.805\,\text{s}$ and $\tau_e = 0.202\,\text{s}$.

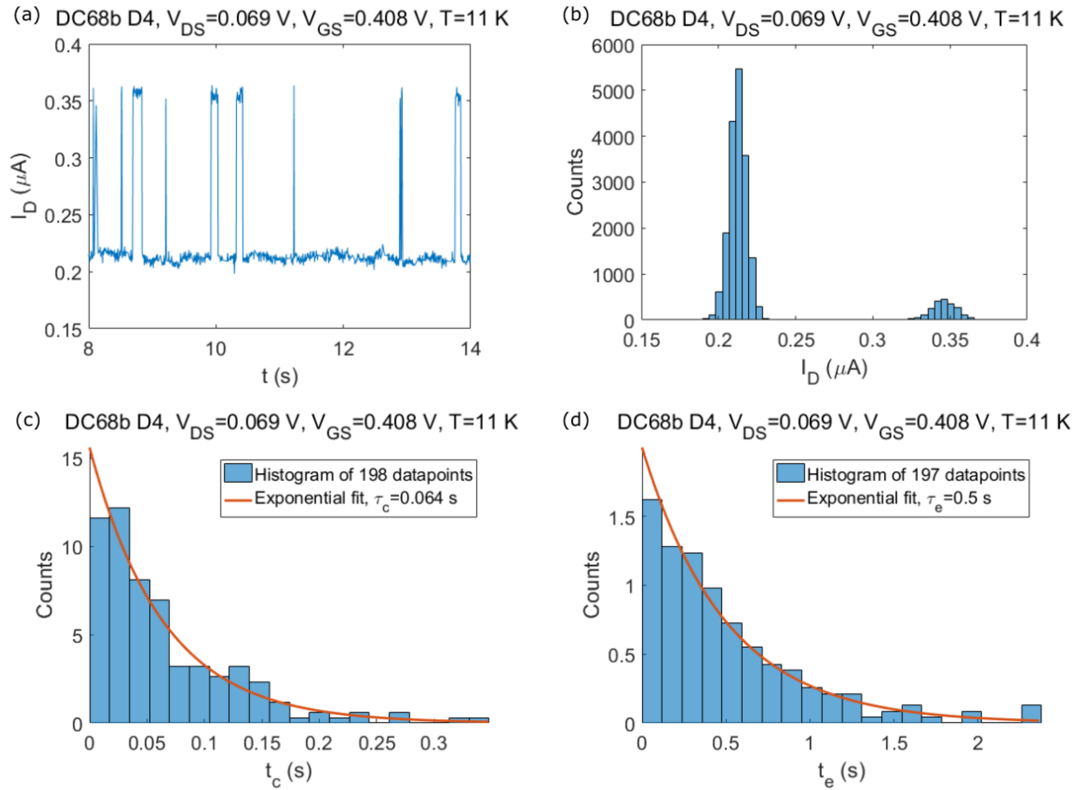

FIGURE 5.7: (a) Excerpt of the two-level RTS noise signal for transistor DC68b D4, biased with $V_{DS} = 0.069\,\text{V}$ and $V_{GS} = 0.408\,\text{V}$. (b) Histogram of the complete RTS noise signal. (c) Normalized histogram of the capture times, resulting in $\tau_c = 0.064\,\text{s}$. (d) Normalized histogram of the emission times, giving $\tau_e = 0.5\,\text{s}$.

To investigate how the gate bias affects τ_c and τ_e (for the same trap state), the gate bias was successively increased from $V_{GS} = 0.39\,\text{V}$ to $V_{GS} = 0.408\,\text{V}$, while the drain voltage was kept constant at $V_{DS} = 0.069\,\text{V}$. Figure 5.7 shows the same plots as in figure 5.6, but for $V_{GS} = 0.408\,\text{V}$. This time, the peak at the lower drain current is higher, meaning

that the drain current is mainly in the low current level and $\tau_e > \tau_c$. From the histogram analysis of the capture and emission times, the time constants can be determined to be: $\tau_c = 0.064\,\mathrm{s}$ and $\tau_e = 0.5\,\mathrm{s}$.

In figure 5.8 (a), τ_c and τ_e are plotted against the mean upper level of the drain current corresponding to each applied gate voltage. Obviously, τ_c decreases with the drain current, whereas the opposite trend is observed for τ_e.

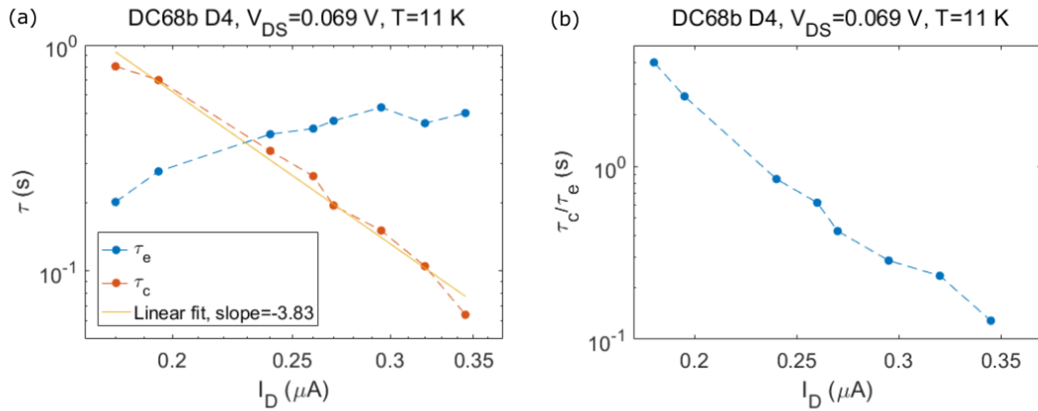

FIGURE 5.8: (a) Capture and emission time constants for the trap in transistor DC68b D4 plotted against the mean upper level of the drain current. τ_c decreases with increasing drain current whereas the opposite trend is observed for τ_e. (b) The mark space ratio decreases exponentially with increasing drain current.

In the purely elastic tunneling model (no SRH processes), the change of the time constants with the drain current can be qualitatively explained by the position of the trap energy level in the gate oxide relative to the semiconductor Fermi level. When changing the gate bias, the trap energy level is shifted up or down, but the semiconductor Fermi level remains constant. When $\tau_c > \tau_e$, the trap energy level is located above the Fermi level, which prevents electrons at the semiconductor Fermi level from tunneling elastically into the trap state. The number of electrons with a high enough energy for the elastic tunneling transition is low at $T = 11\,\mathrm{K}$, giving rise to a large τ_c. On the other hand, once an electron is captured, the emission process out of the trap state is easily possible because $E_T > E_F$. This explains the small τ_e.

Increasing the gate voltage leads to a lowering of the trap energy level until E_T eventually drops below E_F and $\tau_c < \tau_e$. When $E_T < E_F$, there are enough electrons available for the elastic tunneling transition into the trap state, causing τ_c to be small. However, tunneling out of the trap state requires an empty state at the same energy level in the

semiconductor, which is usually not given when $E_{\mathrm{T}} < E_{\mathrm{F}}$, explaining the large τ_{e}.
At the bias point where $\tau_{\mathrm{c}} = \tau_{\mathrm{e}} \approx 0.38\,\mathrm{s}$, the trap energy level is aligned with the semiconductor Fermi level and the trap depth in the oxide can be calculated using equation 2.41 with $\tau_0 = 10^{-10}\,\mathrm{s}$: $z_{\mathrm{T}} = 2.9\,\mathrm{nm}$.
This explanation only holds for degenerate conditions, where the semiconductor Fermi level is (far) above the bottom of the conduction band. In this case, this is a somewhat crude approximation because the applied gate voltages are just below the threshold voltage. Furthermore, the quantization of the energy states in the semiconductor is neglected.
The most commonly used approach in the literature to explain the variation of the time constants with the drain current is the SRH theory, which considers elastic tunneling and energetic transitions. This theory predicts a decrease of τ_{c} with I_D, and an increase of τ_{e} with I_D (see equation 2.36 and 2.38) for an acceptor type of trap, which is in agreement with the observations. Also, an exponential decrease of the mark space ratio with the drain current is expected, assuming that the trap energy level decreases linearly with increasing drain current (see figure 5.8 (b)). Following the SRH model, $\tau_{\mathrm{c}} \sim 1/I_{\mathrm{D}}$ (in a first order approximation), meaning that the measured values of τ_{c} should lie on a straight line with a slope of -1 when τ_{c} is plotted against I_{D}. Quite contrary to this, the slope of a linear fit to τ_{c} is -3.83. Other traps in other transistors show even larger negative exponents of up to -17. One explanation for this deviation might be the neglected quantization effects. Some researchers also mention a Coulomb blockade effect. Details about the Coulomb blockade effect can be found in [13].

The RTS noise amplitude as a function of the drain current is depicted in figure 5.9, displaying that ΔI_{D} increases approximately linearly with I_{D} (from $\Delta I_{\mathrm{D}} = 0.06\,\mu\mathrm{A}$ to $\Delta I_{\mathrm{D}} = 0.13\,\mu\mathrm{A}$). The relative RTS noise amplitude varies between $\Delta I_{\mathrm{D}}/I_{\mathrm{D}} = 33\,\%$ and $\Delta I_{\mathrm{D}}/I_{\mathrm{D}} = 44\,\%$.

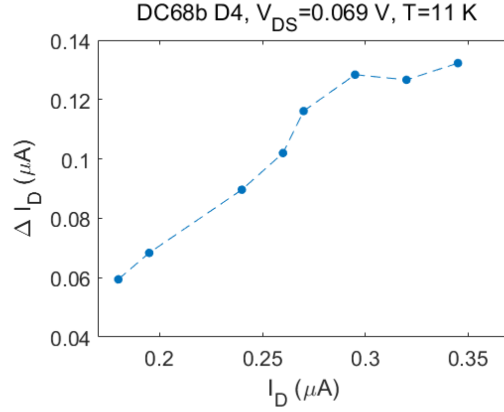

FIGURE 5.9: RTS noise amplitude for a trap in transistor DC68b D4 as a function of the drain current. ΔI_{D} increases from $0.06\,\mu$A to $0.13\,\mu$A, the relative drain current amplitude varies between $\Delta I_{\mathrm{D}}/I_{\mathrm{D}} = 33\,\%$ and $\Delta I_{\mathrm{D}}/I_{\mathrm{D}} = 44\,\%$.

In a small number of transistors, RTS noise amplitudes of up to $1\,\mu$A were observed. Figure 5.10 (a) shows the output characteristic of such a transistor (DC68b C2 with $L_{\mathrm{G}} = 55\,$nm and $W_{\mathrm{G}} = 32\,$nm) and in figure 5.10 (b), an excerpt of the RTS noise signal with applied biases of $V_{\mathrm{DS}} = 0.1\,$V and $V_{\mathrm{GS}} = 0.45\,$V is shown. At this bias point, the time constants are $\tau_{\mathrm{c}} = 0.046\,$s and $\tau_{\mathrm{e}} = 0.267\,$s, and the RTS noise amplitude is $\Delta I_{\mathrm{D}} = 0.8\,\mu$A. The relative RTS noise amplitude amounts to $\Delta I_{\mathrm{D}}/I_{\mathrm{D}} = 67\,\%$.

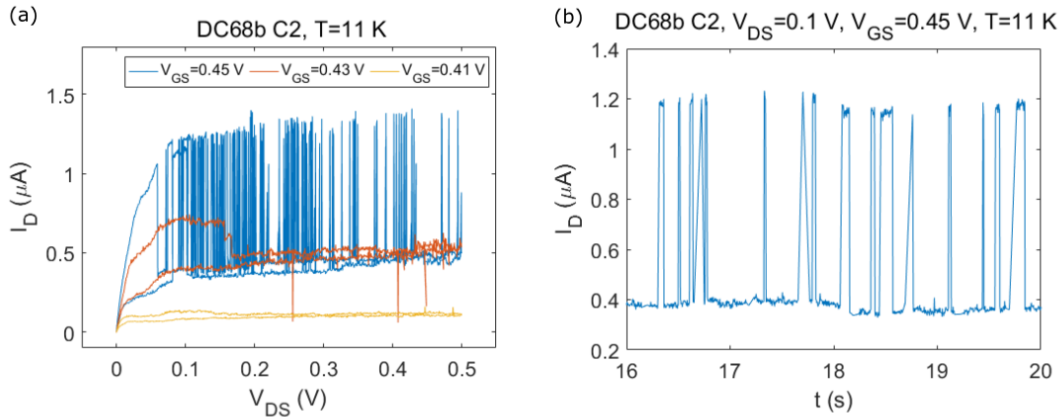

FIGURE 5.10: (a) Output characteristic of transistor DC68b C2, showing huge RTS noise amplitudes. (b) Excerpt of the RTS noise signal at $V_{\mathrm{DS}} = 0.1\,$V and $V_{\mathrm{GS}} = 0.45\,$V.

The different magnitudes of the RTS noise amplitudes can be explained by single trap induced subband fluctuations. Simulations of the impact of a trap located at a depth of $z_{\mathrm{T}} = 2.9\,$nm show RTS noise amplitudes of $\Delta I_{\mathrm{D}} = 0.05 - 0.7\,\mu$A, depending on the exact position of the trap along the NW circumference. These numbers are in good agreement

with the measured RTS noise amplitudes, as shown in figure 5.9 and 5.10 (b). Note that the RTS noise amplitudes depend on the exact physical position of the trap with respect to the channel but not on the trapping and detrapping mechanism.

5.5 RTS Noise at Different Temperatures

To further analyze the trapping and detrapping mechanism of electrons from and to the trap states, RTS noise measurements of the same trap state were performed at different temperatures. In the purely elastic tunneling model, no thermal activation of the time constants is expected, whereas a thermal activation arises when multi-phonon processes are considered.

Figure 5.11 (a) and (b) show the capture and emission time constant of a single trap state in transistor DC64 A9 ($L_{\mathrm{G}} = 60\,\mathrm{nm}$, $W_{\mathrm{G}} = 62\,\mathrm{nm}$) plotted against the drain current for three different temperatures ($T = 11\,\mathrm{K}$, $T = 18.9\,\mathrm{K}$ and $T = 48\,\mathrm{K}$). The capture time constant shows a thermal activation while the emission time constant does not. Usually, it is found that both time constants are thermally activated (if any of them is). Also note that the emission time constant decreases slightly with increasing drain current, as opposed to the commonly observed trend (slight increase of τ_{e} with increasing drain current).

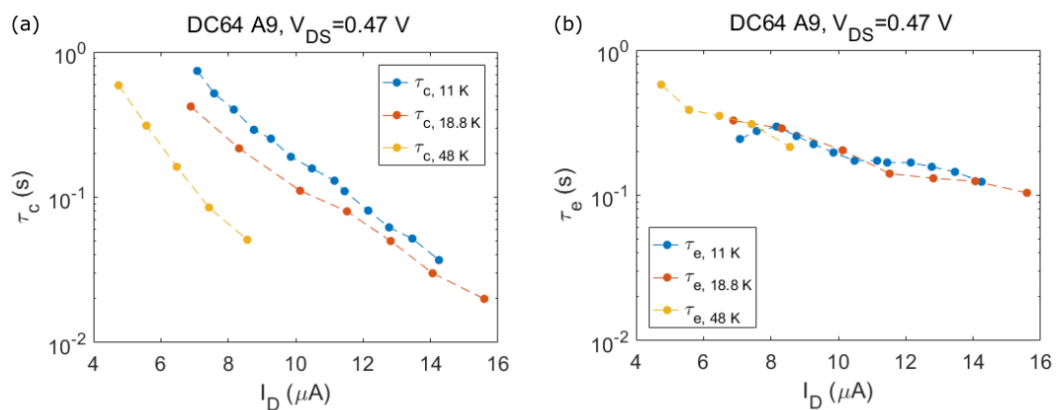

FIGURE 5.11: (a) Capture time constant of a single trap state in transistor DC64 A9 plotted against the drain current for three different temperatures. A thermal activation is visible. (b) Emission time constant of the same trap state as in (a), showing no thermal activation.

From a semi-logarithmic plot of the capture time constant at a fixed drain current against the inverse temperature, the energy barrier for the electron capture process (E_{B}) can be

calculated (see figure 5.12). The slope of the linear fit multiplied by k_B corresponds to E_B, following equation 2.39. For this trap state, $E_\mathrm{B} \approx 2.27\,\mathrm{meV}$.

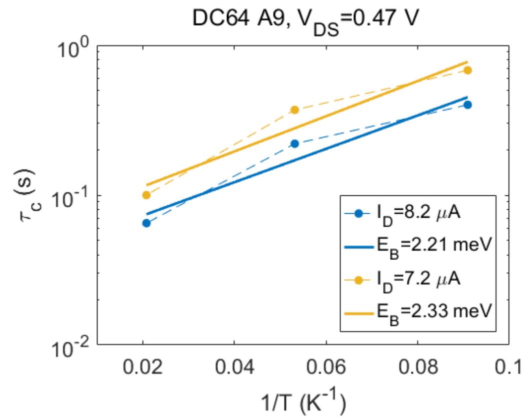

FIGURE 5.12: Capture time constant of the trap state in transistor DC64 A9 plotted against the inverse temperature at two fixed drain currents. The extracted capture energy barriers are $E_\mathrm{B} = 2.21\,\mathrm{meV}$ and $E_\mathrm{B} = 2.33\,\mathrm{meV}$.

Figure 5.13 (a) and (b) show the capture and emission time constant of another single trap state in another transistor (DC64 D2 with $L_\mathrm{G} = 50\,\mathrm{nm}$ and $W_\mathrm{G} = 27\,\mathrm{nm}$), which are plotted against the drain current for temperatures of $T = 197.5$, $T = 207.4\,\mathrm{K}$ and $T = 217.4\,\mathrm{K}$ (significantly higher temperatures than for the transistor before). This time, the emission time constant is thermally activated while capture time constant is not.

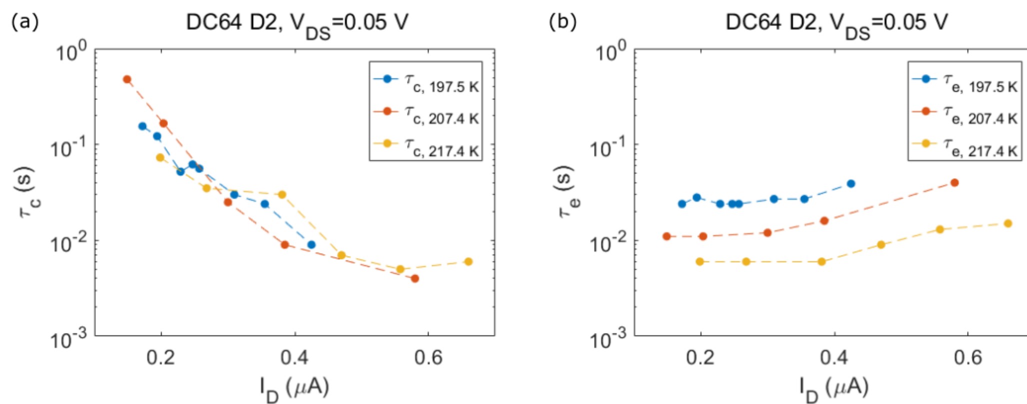

FIGURE 5.13: (a) Capture time constant of a single trap state in transistor DC64 D2 plotted against the drain current for three different temperatures, showing no thermal activation. (b) Emission time constant of the same trap state as in (a), showing thermal activation.

The energy barrier for the electron emission process is $\Delta E \approx 236.32\,\mathrm{meV}$ (see figure 5.14).

It is not so surprising that different values for E_B (capture barrier) and ΔE (emission barrier) are found. First of all, the temperature was much lower in the first measurement series, meaning that less thermal energy could be supplied from the system to overcome energy barriers (explaining the low E_B). Also keep in mind that E_B and ΔE can quite generally have different values and also note that E_B and ΔE were measured for different traps.

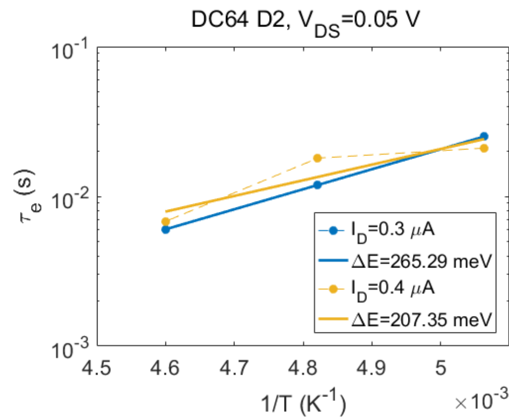

FIGURE 5.14: Emission time constant of the trap state in transistor DC64 D2 plotted against the inverse temperature at two fixed drain currents. The extracted emission energy barriers are $\Delta E = 265.29\,\text{meV}$ and $\Delta E = 207.35\,\text{meV}$.

The fact that a thermal activation of the capture or emission time constant could be observed calls in question the elastic tunneling model. However, from the multi-phonon model, it is expected that both time constants are thermally activated, as opposed to the findings here. To clearly identify the trapping and detrapping mechanism, further temperature-dependent RTS noise measurements are needed.

I will present the results of my Master thesis at the 20th conference on “Insulating Films on Semiconductors” (Infos) in June in Potsdam (Germany), and the results will also be published under the title “1/f and RTS Noise in InGaAs Nanowire MOSFETs” in the journal “Microelectronic Engineering” (accepted paper).

Chapter 6

Conclusions and Outlook

Low-frequency noise measurements were performed on high-performance InGaAs NW MOSFETs.

$1/f$ noise measurements at RT show that the low-frequency noise stems from the channel rather than from the source/drain resistance and that the dominant noise mechanism is carrier number fluctuations. The minimum gate area normalized input gate voltage noise PSD for these devices is as low as $S_{V_G} L_G W_G = 80 \mu\text{m}^2 \mu\text{V}^2\text{Hz}^{-1}$, with a corresponding minimum trap density of $N_T = 9 \cdot 10^{18}\,\text{cm}^{-3}\text{eV}^{-1}$, where the calculation of N_T bases on elastic tunneling of the electrons to and from the trap states in the gate oxide. These values are among the lowest for III-V MOSFETs, demonstrating the feasibility of a high-quality, low trap density, high-κ gate oxide on InGaAs NW structures. Furthermore, a gate voltage dependence on the $1/f^{\beta}$ noise exponent (β) is observed, which can be traced back to a spatially non-uniform trap density in the elastic tunneling picture.

$1/f$ noise measurements at low temperatures confirm that the dominant noise mechanism is carrier number fluctuations and also give some support for the elastic tunneling model.

In approximately 2/3 of the considered devices, RTS noise signals were found. This shows that very few traps limit the electron transport through the channel. It is observed that the capture and emission time constant of a single active trap state change with the applied gate bias. Two models (elastic tunneling model and SRH model) are discussed to explain this behavior. Another striking feature of the RTS noise measurements is the large spread in the RTS noise amplitudes, with maximum amplitudes of up

to 1 μA (relative RTS noise amplitudes of up to 67 %). Simulations of the impact of single trap induced subband fluctuations can reproduce this large spread in the RTS noise amplitudes and they show that the magnitude of the RTS noise amplitude is especially sensitive to the position of the trap state along the NW circumference. Temperature-dependent RTS noise measurements reveal a thermal activation of the time constants (either the capture time constant or the emission time constant is thermally activated), which cannot be explained by the elastic tunneling model. The multi-phonon theory, on the other hand, predicts such a thermal activation, but for both time constants. More temperature-dependent RTS noise measurements are needed to clearly identify the trapping and detrapping mechanism.

Bibliography

[1] G. E. Moore, "Cramming more components onto integrated circuits," *Electronics*, vol. 38, April 1965.

[2] R. H. Dennard, F. H. Gaensslein, 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, pp. 256–268, October 1974.

[3] 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, pp. 668–677, 2014.

[4] J. A. D. Alamo, "Nanometre-scale electronics with III-V compound semiconductors," *Nature*, vol. 479, pp. 317–323, November 2011.

[5] J. A. D. Alamo, D. A. Antoniadis, J. Lin, W. Lu, A. Vardi, and X. Zhao, "Nanometer-Scale III-V MOSFETs," *IEEE Journal of the Electron Devices Society*, vol. 4, pp. 205–214, September 2016.

[6] X. Sun and T. P. Ma, "Electrical Characterization of Gate Traps in FETs With Ge and III-V Channels," *IEEE Transactions on Device and Materials Reliability*, vol. 13, pp. 463–479, December 2013.

[7] S. Sze and M.-K. Lee, *Semiconductor Devices*. Wiley, 3 ed., 2013.

[8] R. Kim and M. S. Lundstrom, "Characteristic Features of 1-D Ballistic Transport in Nanowire MOSFETs," *IEEE Transactions on Nanotechnology*, vol. 7, pp. 787–794, November 2008.

[9] M. Lundstrom and J. Guo, *Nanoscale Transistors*. Springer, 2006.

[10] M. von Haartman and M. Östling, *Low-Frequency Noise in Advanced MOS Devices.* Springer, 2007.

[11] R. Jayaraman and C. G. Sodini, "A 1/f noise technique to extract the oxide trap density near the conduction band edge of silicon," *IEEE Transactions on Electron Devices*, vol. 36, pp. 1773–1782, September 1989.

[12] G. Ghibaudo and T. Boutchacha, "Electrical noise and RTS fluctuations in advanced CMOS devices," *Microelectronics Reliability*, vol. 42, pp. 573–582, April-May 2002.

[13] E. Simeon, *Random Telegraph Signals in Metal-Oxide-Semiconductor Devices.* 2015.

[14] W. Shockley and J. W. T. Read, "Statistics of the Recombination of Holes and Electrons," *Physical Review*, vol. 87, pp. 835–842, Sptember 1952.

[15] T. Grasser, "Stochastic charge trapping in oxides: From random telegraph noise to bias temperature instabilities," *Microelectronics Reliability*, vol. 52, pp. 39–70, 2012.

[16] K. S. Ralls, W. J. Skocpol, L. D. Jackel, R. E. Howard, L. A. Fetter, R. W. Epworth, and D. M. Tennant, "Discrete Resistance Switching in Submicrometer Silicon Inversion Layers: Individual Interface Traps and Low-Frequancy ($1/f$?) Noise," *Physical Review Letters*, vol. 52, pp. 228–231, January 1984.

[17] M. J. Kirton and M. J. Uren, "Capture and emission kinetics of individual Si:SiO2 interface states," *Appl. Phys. Lett.*, vol. 48, pp. 1270–1272, May 1986.

[18] D. Veksler, G. Bersuker, S. Rumyantsev, M. Shur, H. Park, C. Young, K. Y. Lim, W. Taylor, and R. Jammy, "Understanding noise measurements in MOSFETs: the role of traps structural relaxation," *Reliability Physics Symposium (IRPS), IEEE International*, pp. 73–79, May 2010.

[19] Z. Shi, J.-P. Mieville, and M. Dutoit, "Random telegraph signals in deep submicron n-MOSFETs," *IEEE Transactions on Electron Devices*, vol. 41, pp. 1161–1168, July 1994.

[20] K. K. Hung, P. K. Ko, C. Hu, and Y. C. Cheng, "Random Telegraph N of Deep-Submicrometer MOSFETs," *IEEE Electron Device Letters*, vol. 11, pp. 90–92, February 1990.

[21] K. P. Cheung, J. P. Campbell, and A. Oates, "High mobility channel from the prospective of random telegraph noise," *Solid-State Device Research Conference (ESSDERC)*, pp. 335–338, October 2011.

[22] M.-L. Fan, V. P.-H. Hu, Y.-N. Chen, P. Su, and C. T. Chuang, "Analysis of Single-Trap-Induced Random Telegraph Noise on FinFET Devices, 6T SRAM Cell, and Logic Circuits," *IEEE Transactions on Electron Devices*, vol. 59, pp. 2227–2234, August 2012.

[23] C. Zota, L.-E. Wernersson, and E. Lind, "Quantized-Conduction and High Mobility in Selectively Grown $In_xGa_{1-x}As$ Nanowires," *ACS Nano*, vol. 9, pp. 9892–9897, 2015.

[24] C. Zota, L.-E. Wernersson, and E. Lind, "High-Performance Lateral Nanowire InGaAs MOSFETs With Improved On-Current," *IEEE Electron Device Letters*, vol. 37, pp. 1264–1267, October 2016.

[25] M. Si, N. Conrad, S. Shin, J. Gu, J. Zhang, M. A. Alam, and P. D. Ye, "Low-Frequency Noise and Random Telegraph Noise on Near-Ballistic III-V MOSFETs," *IEEE Transactions on Electron Devices*, vol. 62, pp. 3508–3515, November 2015.

[26] 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, p. 033508, July 2013.

[27] N. Li, E. S. Harmon, J. Hyland, D. B. Salzman, T. P. Ma, Y. Xuan, and P. D. Ye, "Properties of InAs metal-oxide-semiconductor structures with atomic-layer-depositioned Al_2O_3 Dielectric," *Applied Physics Letters*, vol. 92, p. 143507, April 2008.

[28] N. V. Nguyen, M. Xu, O. A. Kirillov, P. D. Ye, C. Wang, K. Cheung, and J. S. Suehle, "Band offsets of $Al_2O_3/In_xGa_{1-x}As$ (x=0.53 and 0.75) and the effects of postdeposition annealing," *Applied Physics Letters*, vol. 96, p. 052107, February 2010.