TECHNISCHE UNIVERSITÄT WIEN

DISSERTATION

Charge Trapping and Variability in CMOS Technologies at Cryogenic Temperatures

zur Erlangung des akademischen Grades

Doktor der technischen Wissenschaften

eingereicht an der Technischen Universität Wien

Fakultät für Elektrotechnik und Informationstechnik

von

JAKOB MICHL

Matrikelnummer 1302759

Betreut von UNIV.PROF. DIPL.-ING. DR.TECHN. TIBOR GRASSER

und ASSISTANT PROF. DIPL.-ING. DR.TECHN. MICHAEL WALTL

Wien, im August 2022

Acknowledgements

First of all, I would like to express my gratitude to **Professor Tibor Grasser**, for the opportunity to do research at the Institute for Microelectronics and for guiding me throughout my PhD study. I am truly grateful that he supported me whenever I needed help and that he motivated and inspired me by always asking the right questions. His passion and enthusiasm for research are contagious and encouraged me during the last years.

I would like to thank **Michael Waltl**, for co-supervising my PhD study and for securing the funding[1] necessary for my employment. He developed the measurement equipment I used for most of my characterization experiments and helped me whenever I encountered issues in the lab. His joy for experimental work made me excited about my experimental work and animated me to explore the boundaries of our setup.

Special thanks go to **Alexander Grill**, who supported me whenever needed, introduced me to all our simulation and measurement tools and had countless valuable discussions with me. He motivated me whenever I struggled during my studies and always strengthened my back. He advised me in innumerable decisions and became a true fried during my research visits at *imec*. Without him, this thesis would not have been possible.

I want to thank my colleagues **Markus Kampl**, **Theresia Knobloch**, **Christian Schleich**, **Bernhard Stampfer**, **Konstantinos Tselios**, and **Dominic Waldhör** for not only providing valuable input, but also for being very good friends.

I also want to thank reasearchers at *imec* with whom I collaborated with, especially **Ben Kaczer**, **Arnout Beckers**, **Dieter Claes**, and **Marko Simičić**. They made the time during my internships fly by.

Last but not least, I would like to thank my family and friends, especially my parents **Gerda** and **Herbert** and my sister **Kristina**. They are there for me when I need them and always have my back. Finally, I want to thank my girlfriend **Lisa**. Your support means the world to me.

[1]The financial support by the Austrian Federal Ministry for Digital and Economic Affairs, the National Foundation for Research, Technology and Development, and the Christian Doppler Research Association are gratefully acknowledged.

Abstract

CMOS technology operated at cryogenic temperatures is essential in various fields such as quantum computing (QC), where it serves as a classical control interface for qubits operating in the mK-regime, as host technology enabling a monolithic integration of qubits, or in high-performance computing (HPC) applications. For all these applications, the changing device properties towards cryogenic temperatures must be taken into account in circuit designs. Additionally, potential design optimizations which allow an operation using a very low supply voltage at cryogenic temperatures can be achieved. This is in particular of high interest for QC and HPC to reduce the energy consumption and increase operational frequency. However, the development of robust applications in this field is very challenging for reliability engineers, because their stable operation is very sensitive to drifts of the threshold voltage due to aging and to variability issues. Furthermore, in all applications related to quantum computing, noise originating from charge trapping is very critical, because it can decrease the fidelity of the qubits.

Over the past decades, a deep understanding of the role of defects in the oxide and at the interface between the oxide and substrate has been developed. Such defects can capture and emit a charge which leads to a change in the device electrostatics and are thus responsible for various reliability issues. The charge trapping kinetics can be approximated with the nonradiative multiphonon (NMP) model, which successfully describes the trapping kinetics even at cryogenic temperatures. Since solving the model in its full complexity is computationally expensive, an efficient model has been developed to allow the computation of charge transition rates of thousands of defects and has been implemented in the reliability simulator *Comphy*. This allows to calculate the cumulative response of many defects and enables a comparison of theoretical trap parameters with measurement data. To this end, bias temperature instability and random telegraph noise (RTN) measurements have been conducted between 4 K and room temperature on various technologies. The degradation behavior has been modeled, allowing the extraction of trap parameters and the identification of defect candidates responsible for altering the device electrostatics.

On large-area devices thousands of such traps can be electrically active simultaneously and the superposition of their responses can be measured. In contrast to that, on scaled devices only few traps are active at the same time, which allows to access properties of single defects. Time-dependent defect spectroscopy and RTN studies which are covered in this work show that there is charge trapping even in the limit

of cryogenic temperatures. Charge capture and emission rates become temperature independent towards 4 K which is a consequence of nuclear tunneling. This temperature independence of the trapping kinetics in the deep cryogenic regime has been modeled using the NMP model.

While single devices allow to study the physical device degradation mechanisms in detail, a knowledge of the distribution of device performance parameters is essential to qualify a full technology. For this, the characterization of *SmartArrays* with thousands of devices which can be addressed digitally has been performed. This allows the study of the variability of time-zero parameters and its dependence on the temperature. An increasing variability of important device parameters towards 4 K can be shown and explained with the occurrence of resonant tunneling which gets more prominent at cryogenic temperatures.

Kurzfassung

CMOS-Transistoren, die bei Tieftemperaturen betrieben werden, finden Anwendung in vielen Gebieten, wie der Quanteninformatik, wo die Steuerungselektronik für Qubits auf 4 K gekühlt wird, oder als Host-Technologie für eine monolithische Integration mit den Qubits. Auch für Hochleistungsrechner könnte ein Betrieb der CMOS-Logik bei Tieftemperaturen effizient sein. Alle diese Anwendungen basieren auf den temperaturabhängigen Transistoreigenschaften, welche eine Optimierung des Designs für den Betrieb bei Tieftemperaturen erlauben. Mit dem Betrieb der CMOS-Logik bei Tieftemperaturen entstehen neue Herausforderungen für Zuverlässigkeitsingenieur*innen. So sind Anwendungen im Zusammenhang mit Quantencomputern sehr empfindlich auf Rauschen, welches durch das Einfangen und Emittieren von Ladungen entsteht, und zu Dekohärenz der Qubits führen kann. Darüber hinaus lassen optimierte Designs wenig Spielraum für Variabilität zwischen den einzelnen Transistoren und für Schwellspannungsverschiebungen.

Die Rolle von Defekten im Oxid und an der Oberfläche zwischen Oxid und Substrat wurde in den letzten Jahrzehnten intensiv erforscht. Diese Defekte können Ladungsträger einfangen und emittieren, was zu einer Änderung in der Elektrostatik des Transistors führt und letztendlich die Zuverlässigkeit der Bauteile negativ beeinflusst. Diese Kinetik kann mit dem Nichtstrahlende-Multiphononen-Modell (*engl. nonradiative multiphonon*, NMP) beschrieben werden. Da das vollständige NMP-Modell sehr rechenintensiv ist, wird in dieser Arbeit eine effiziente Näherung präsentiert, welche in den Zuverlässigkeitssimulator *Comphy* implementiert wurde. Das ermöglicht die Berechnung des Einflusses von tausenden Defekten auf die Schwellspannung und einen Vergleich mit experimentellen Daten. Zu diesem Zweck wurde die Spannungs-Temperatur-Instabilität (*engl. bias-temperature-instability*) zwischen 4 K und Raumtemperatur mit verschiedenen CMOS-Technologien gemessen. Das Degradationsverhalten wurde anschließend modelliert und mögliche Defektkandidaten wurden identifiziert.

Während in Transistoren mit großen Geometrien tausende von Defekten gleichzeitig aktiv sind, erlauben skalierte Transistoren die Untersuchung von Eigenschaften einzelner Defekte. In dieser Arbeit werden dazu Messungen mit der zeitabhängigen Defektspektroskopie (*engl. time-dependent defect spectroscopy*) und von Telegraphenrauschen (*engl. random telegraph noise*) präsentiert. Diese Messungen zeigen, dass Defekte selbst bei Tieftemperaturen noch Ladungsträger einfangen und emittieren. Die zugehörigen

Raten werden temperaturunabhängig, was durch Tunneln zwischen atomistischen Konfigurationen (*engl. nuclear tunneling*) erklärt werden kann.

Einzelne Bauteile bieten die Möglichkeit, physikalische Defektmechanismen zu untersuchen. Um eine Technologie vollständig charakterisieren zu können, ist auch eine Untersuchung der Variabilität der Bauteileigenschaften notwendig. Dazu wurden *SmartArrays* verwendet, die es erlauben, tausende von Transistoren digital anzusteuern. Das ermöglicht die Untersuchung der Verteilung von Transistoreigenschaften und deren Temperaturabhängigkeit. Die Variabilität wichtiger Parameter nimmt bei 4 K zu. Das kann mit Hilfe von resonantem Tunneln erklärt werden, welches bei Tieftemperaturen auftritt.

Contents

List of Figures

List of Abbreviations

aRTN	anomalous random telegraph noise
ASIC	application-specific integrated circuit
BTI	bias temperature instability
CCR	closed cycle refrigirator MOS
CMOS	complementary MOS
CSA	charge sheet approximation
CV	capacitance-voltage
DFT	density functional theory
DIBL	drain induced barrier lowering
DUT	device under test
eMSM	extended MSM
EOT	equivalent oxide thickness
ESiD	effective single defect decomposition
FET	field-effect transistor
HCD	hot carrier degradation
HKMG	high-κ metal gate
HPC	high-performance computing
imec	Interuniversity Microelectronics Centre
IRDS	International Roadmap for Devices and Systems
ITRS	International Technology Roadmap of Semiconductors
IV	current-voltage
JJ	Josephson junction
MOS	metal-oxide-semiconductor
MOSFET	MOS field-effect-transistor
MSM	measure stress measure
NMP	non-radiative multi-phonon
NNLS	non-negative least square
OTF	on-the-fly
PDK	process design kit
PEC	potential energy curve
PES	potential energy surface
PNO	plasma-nitrided oxides

PSD	power spectral density
QC	quantum computing
QD	quantum dot
qubit	quantum bit
RT	room temperature
RTN	random telegraph noise
RTS	random telegraph signal
SEM	scanning electron microscope
SEPC	strong electron-phonon coupling
SET	single electron transistor
SPM	saddlepoint method
SRH	Shockley–Read–Hall
TCAD	technology computer-aided design
TDDS	time-dependent defect spectroscopy
VLSI	very large-scale integration
WEPC	weak electron-phonon coupling
WKB	Wentzel–Kramers–Brillouin

Chapter 1

Introduction

The *Digital Revolution* goes hand in hand with the rise of the modern semiconductor industry and was enabled by the invention of metal-oxide-semiconductor field-effective transistors (MOSFETs) as the backbones of digital circuits. This would not have been possible without understanding and exploiting effects of quantum mechanics. The development of quantum physics and the resulting scientific and technological breakthroughs in the 20th century are therefore often called the *First Quantum Revolution*. Currently, we are in the middle of the *Second Quantum Revolution*, where devices actively operate with quantum states with the ultimate goal of developing a universal quantum computer [1]. In the following chapter, the role of MOSFETs operated at cryogenic temperatures (cryo-CMOS) and the related challenges will be discussed, setting a special focus on quantum computing (QC) and high-performance computing (HPC). In this context, especially the role of reliability issues in cryo-CMOS devices will be highlighted. The chapter will close by presenting the scope and outline of this thesis.

1.1 Field-Effect Transistor

The *Digital Revolution* would not have been possible without the field-effect transistor (FET), whose invention started in the early days of the 20th century. Around 1925, the Austrian physicist Julius Edgar Lilienfeld started to develop the idea and theoretical description on how to use the field effect to modulate the conductivity in a semiconductor triode structure resulting in a patent application in 1928 [2]. Only seven years later, in 1935, the German Physicist Oskar Heil patented the first insulated-gate FET [3]. However, it is unknown whether Lilienfeld or Heil were able to realize a working transistor as proof of concept [4]. From 1948 on developments accelerated when the first junction-gate FET patent was filed by Bell Labs, as suggested by Brattain, Bardeen and Shockley [5]. However, it was not until the 1960s when first MOSFETs where fabricated by Kahng [6] and Atalla [7]. While the first structures suffered from large variations due to a high density of interface defects [8], manufacturing processes improved rapidly and commercially fabricated transistors became available in 1964 by Fairchild and by

RCA [4]. Already 1963 the researchers Sah and Wanlass from Fairchild suggested a new type of MOSFET logic combining both nMOS and pMOS, called complementary MOS (CMOS) [9], and filed a patent which was granted in 1967 [10]. This new logic allowed logic circuits operating at significantly lower power consumption and became the de facto standard fabrication process for very-large-scale integration (VLSI) chips in the 1970s. Since then, many improvements of the processing steps, material systems involved and circuit technology enabled to reduce the feature size from the µm-range to a few nm nowadays. For this, the lithography process was improved to enable the fabrication of gate lengths smaller than wavelength of the used UV light. Furthermore, along many other optimizations, new gate and insulator materials such as high-κ dielectrics were introduced and the geometry was improved to non-planar transistors (e.g. FinFETs, GAA FETs).

All these improvements in the fabrication process for manufacturing smaller feature sizes, are summarized in what is today well-known as *Moore's law*. In 1965 Gordon Moore, who is a co-founder of Fairchild and Intel, proposed that the number of transistors per chip will double every year [11]. In 1975 he revised this estimation for proposing a redoubling of transistors per chip approximately every two years. This forecast has been followed by the industry for almost 50 years now. However, the scaling of transistors has a natural limit set by the atomic dimensions. The International Technology Roadmap of Semiconductors (ITRS) [12], which sets development standards and outlines future developments, published its final roadmap in 2016, because the classical scaling approach would reach its ultimate limits in the 2020s [13]. Thus, in 2016 the IEEE International Roadmap for Devices and Systems (IRDS) [14] was founded which is the successor of the ITRS and which has a broader focus on future developments. One of the newly defined key topics is *Cryogenic Electronics and Quantum Information*, which discusses the role of quantum computing as a new emerging field.

1.2 Cryo-CMOS Applications

As *Moore's law* reaches its limits with respect to scaling the transistors on the chips, it is necessary to develop *Beyond CMOS* technologies [14]. *Beyond CMOS* as it is defined by the IRDS spans a wide range of research fields, such as carbon nanotube FETs, graphene based devices, molecular electronics, spintronics, optical computing, or superconducting computing [14]. The most prominent application of the latter, is superconducting quantum computing. Various approaches for building a quantum computer (including the superconducting approach) need cryogenic environments, which have made cryo-CMOS an emerging field in the past years. However, cryo-CMOS has also multiple other potential applications, such as high-performance computing [15–17] or space technologies [18]. Quantum computing (QC) and high-performance computing (HPC) will be introduced in more detail in the following, because these are central cryo-CMOS applications in the *Beyond CMOS* program of TU Vienna's research partner *imec*.

1.2.1 Quantum Computing

With the *First Quantum Revolution* in the early 20th century, a range of technologies which determine our daily lives nowadays emerged. Electronics, satellites, lasers, medical imagery, etc. would not have been realizable without a deep understanding of quantum mechanics. Now, for the *Second Quantum Revolution*, global players invest billions for research and development of new technologies. In its 14th 5-Year-Plan for 2021 to 2025, China declared quantum technologies as a key technology [19]. Within the *Quantum Flagship* project started in 2018 [20], the European Union invests at least 1 Billion Euros in the research on quantum technologies with the ultimate goal of developing a quantum computer (QC) that is able to outperform classical computers in certain tasks. But not only governments invest Billions into QC, also private technology companies see a potential market and already demonstrated a QC on small scales, e.g. Sycamore from Google [21], IBM Q [22] or Quantum Inspire from Intel/QuTech [23]. An overview of the largest QC projects is given in Fig. 1.1.

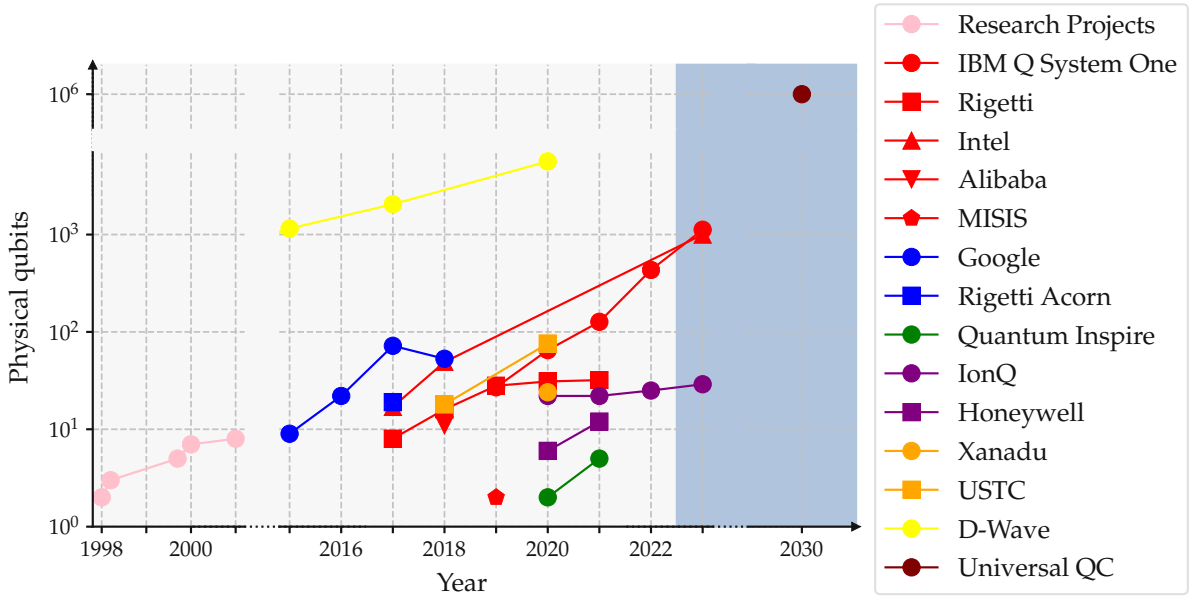

Figure 1.1. First proof of concepts for QC were shown around the turn of the millennium [24–27]. Different approaches are under development for universal quantum computers (UQC): Superconducting qubits: IBM [28], Rigetti [29, 30], Intel [31], Alibaba [32], MISIS [33], Transmon qubits: Goolge [21, 34, 35], Rigetti (Project Acorn) [36], Electron spin qubits: QuTech/TU Delft [23, 37, 38], Trapped Ions: IonQ [39, 40], Honeywell [41–43], Photonics: Xanadu [44, 45], USTC [46, 47]. Quantum annealing systems, which are not UQC, are commercially available since 2011 [48–51]. The goal announced by Intel and Google is a UQC with 1 million physical qubits by 2030.

While classical computers can solve problems from the complexity class BPP (bounded-error probabilistic polynomial time) with efficient probabilistic algorithms, QCs can access the complexity class BQP (bounded-error quantum polynomial time) and can thus solve certain problems very efficiently, which are not accessible for classi-

cal computers within a reasonable computation time [52]. This class of BQP problems occurs in various fields such as cryptography [53], machine learning [54], medicine [55], computational chemistry [56] or in the financial sector [57], and thus attracts a large interest from different stakeholders in industry and research alike.

QCs are based on the fundamental concept of quantum bits (so called *qubits*), the basic units in quantum information technology, which are analogous concepts to binary bits, which are well known from classical computers. Mathematically, a qubit is a linear combination (superposition) $\alpha |0\rangle + \beta |1\rangle$ of two orthogonal basis vectors $|0\rangle$ and $|1\rangle$. This superposition is typically represented by a Bloch sphere, as can be seen in Fig. 1.2 (left). By using qubits, it is possible to build quantum logic gates, the basis of quantum circuits [52]. A single-qubit gate operation can be represented by a 2×2 unitary matrix having the effect of a rotation of the Bloch vector on the spherical surface. While the mathematical description of single qubits is rather simple, the engineering task of building physical qubits is still extremely challenging. Since qubits use fundamental properties of quantum mechanics, such as superposition and entanglement, they also suffer from decoherence. Decoherence can be interpreted as an irreversible information loss of quantum states as soon as there is any type of interaction with the environment. To minimize this interaction, qubits are operated at cryogenic temperatures. Despite the many difficulties for the operation of qubits, there have been first major successes. The QCs of Google, IBM, and Intel/QuTech operate between 50 and 100 qubits and successfully run some simple quantum algorithms [21–23]. In 2019 Google even announced that they reached quantum supremacy. They claimed that it would take their quantum processor Sycamore with 54 qubits (53 functional ones) about 200 seconds to sample one instance of a quantum circuit one million times, a task for which a state-of-the-art supercomputer would need around 10,000 years [21]. However, this claim turned out be be rather controversial, because IBM published on their research blog that the task could also be done in 2.5 days on a classical computer with a far greater fidelity [58]. Independent of whether quantum supremacy has already been reached or not, the whole discussion addresses a problem which is highly artificial and currently not relevant for real-world applications. To run more useful algorithms like Shor's algorithm for prime factorization [53] or Grover's algorithm for searching an unordered list [59], it is necessary to operate thousands or millions of qubits. To enable this, there is a multitude of challenges which still need to be overcome.

One of the most important ones is the choice of the physical qubits. In theory, any two-state quantum system could be considered as a basis, however, no leading qubit technology has been established withing the QC community yet. Therefore, many different technologies are currently explored, for example trapped ion qubits [61], qubits based on crystal impurities [62, 63] or single photons [64]. The qubit-technologies used by the big tech-companies are superconducting qubits [21, 28] and silicon qubits [23, 37]. These technologies have been demonstrated already on small

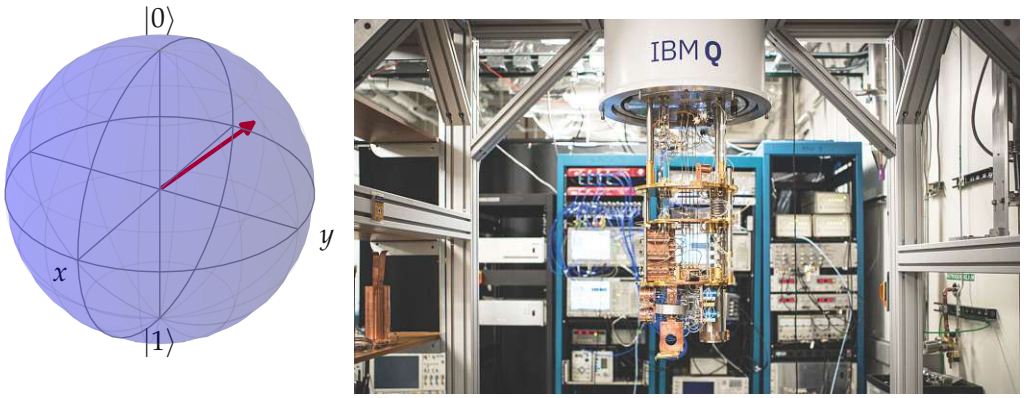

Figure 1.2. Left: Mathematical representation of a qubit as Bloch sphere, which represents all possible states as superposition of $\alpha\,|0\rangle + \beta\,|1\rangle$. Operations on a qubit are represented by unitary matrices, resulting in a rotation of the vector on the sphere. Right: Quantum computer from the *IBM Q* project. The golden chandelier in the photograph shows different stages, from the mK-stage at the bottom to room temperature at the top. Figure taken from [60].

scales and they can be produced in industrial cleanrooms for 300mm wafers [65–67]. For these reasons, they are currently also in the focus of our research partner *imec* [67].

Superconducting Qubits

Among all qubit technologies, the superconducting qubit approach is currently the most advanced one in terms of the number of qubits, as can be seen in Fig. 1.1. Here, the qubits are based on Josephson Junctions (JJs), see Fig. 1.3 (a). Typically, JJs consist of two superconducting stripes, which are connected via a tunnel oxide in between. At deep-cryogenic temperatures (in the mK-regime) JJs act as non-linear inductors [68]. While a linear inductor in an LC-circuit would lead to equidistant energy levels in the corresponding quantum harmonic oscillator, the anharmonicity of a non-linear JJ leads to non-equidistant energy levels. This allows to take the ground state $|0\rangle$ and the first excited state $|1\rangle$ as two-state quantum system without allowing a higher state to become accessed. The manipulation of the qubit (which is the rotation on the Bloch sphere) can be induced by microwave pulses sent to an antenna which is coupled to the qubit. The frequency of this wave determines the rotation between the different energy levels. It has to be noted that there are various realizations of superconducting qubits based on non-linear inductors operating at cryogenic temperatures. A more detailed overview which realizations exist and how superconducting qubits work can be found in [68–71].

Electron Spin Qubits

Another class of qubits are silicon spin qubits, which is a collective name for various technologies, based upon using the spin of a charge carrier as two-state quantum system. There are various types of spin qubits, e.g. based on Si/SiGe [74], Si SOI [75], or based on Si MOS [76]. The latter ones are also in the focus of our research partner

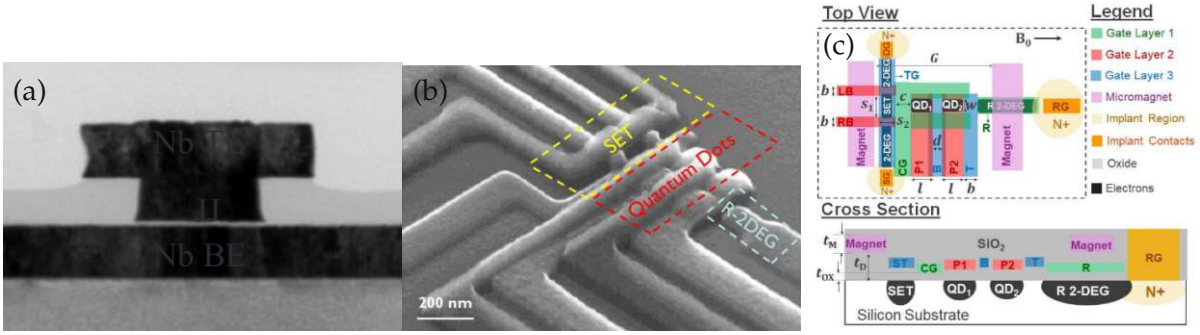

Figure 1.3. (a) Cross-section transmission electron microscope (TEM) of a fully patterned Josephson junction. **(b)** The scanning electron microscope (SEM) image of a Si electron spin qubit shows a single electron transistor (SET) for readout, two quantum dots (QDs) at the Si/SiO_2 interface, and the 2-dimensional electron gas reservoir for loading the QDs. **(c)** The schematic of a silicon spin qubit shows multiple gates for the qubit confinement and two magnets for operating the qubits. Figures taken from [67, 72, 73].

imec [65, 67, 73, 77] and are therefore discussed in more detail in the following. A SEM image of the qubit can be seen in Fig. 1.3 (b), where the single electron transistor (SET), the two quantum dots and the 2-dimensional electron gas (2-DEG) are highlighted. Several gates are needed to electrostatically confine the two quantum dots QD_1 and QD_2 located at the Si/SiO_2-interface as can be seen in Fig. 1.3 (c). The 2-DEG serves as a charge reservoir enabling the loading of the QDs. The qubits can then be manipulated with a magnetic field created by micromagnets via electric dipole spin resonance [78] (alternatively, electron spin resonance can be used). Using the SET it is possible to readout the spin states using spin-charge conversion [73, 79]. A series of qubit operations using the magnetic field allows then the execution of quantum algorithms. Unlike superconducting qubits, the fabrication of Si spin qubits could benefit from many advantages of modern CMOS scaling which makes the concept highly interesting for commercial applications [65]. In theory, silicon spin qubits are ideal candidates for a co-integration of CMOS control hardware and qubits on the same chip [76]. However, state of the art projects are dominated by superconducting qubits (see Fig. 1.1) while spin qubits lag behind. The main reason for this is the limiting gate fidelity which describes how close two realized quantum states are. The fidelity of spin qubits is below the threshold required for complex quantum algorithms which leads to error rates of gate operations around 1% and thus makes the up-scaling impossible [65]. One of the main limitations for fidelity is the non-uniformity of qubits, which arises from the fabrication process. Among other processes, interface defects between the Si and the SiO_2 layer can be introduced during manufacturing, which has the consequence that every QD behaves slightly different and it is thus necessary to find for every qubit the right operation voltage regime individually [80]. Here, knowledge from the semiconductor industry, specially from the field of device reliability, can be used to improve the interface quality and to further specify the limiting factors.

Classical Control Interface

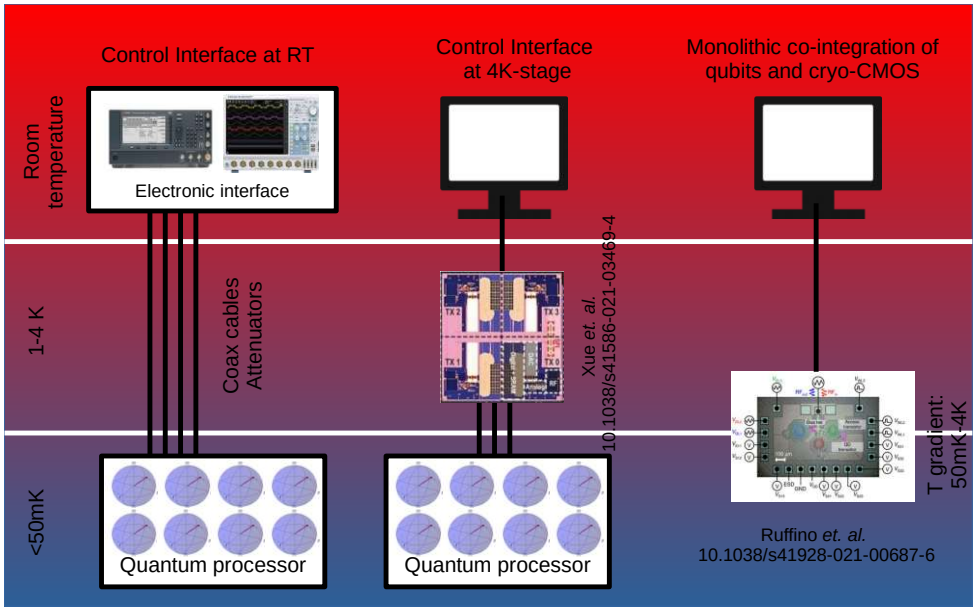

Figure 1.4. The first successful QCs use a dilution refrigerator for bringing the quantum processor to the sub-K regime. The qubits are controlled via coaxial cables with electronics at RT. Since this leads to a wiring bottleneck, lots of effort has been spent to design cryo-CMOS control interfaces working at the 4 K-stage [81]. Recently, first successes have been published bringing a monolithic co-integration of qubits and the cryo-CMOS control hardware to the cryogenic stage, where a large temperature gradient between the qubit layer at 50 mK and the control layer at 1 K to 3 K occurs [82, 83].

While the importance of the choice of the qubit technology is crucial, another less obvious bottle neck appears at the classical control interface for qubits. Until recently, existing QCs with a maximum of ∼100 qubits have used racks full of electronic equipment (RF sources, DC sources, etc.) at room-temperature, where each individual qubit is connected via several coaxial cables to electronic instruments [81]. For such small systems, fidelities of above 99% can already be reached [84, 85]. However, for QCs which should be able to solve real-world applications the operation of thousands or millions of physical qubits is necessary to perform calculations together with quantum-error correction schemes, such as surface codes [86, 87]. For such large-scale QCs the research approach of room-temperature electronics fails for multiple reasons. The signal integrity across the RF lines for controlling the qubits gets too low, the number of possible physical connections at the test chip is limited, and the damping by the high resistivity of the vacuum throughputs becomes too high. Therefore, the classical control interface must be integrated within the cryostat and located as close as possible to the quantum processor. For this, CMOS technology is an ideal candidate for the control interface, because it can be produced in a high-quality commercially, allows the integration of billions of transistors, has a low power consumption and is functional even in the sub-Kelvin regime [88–92]. However, to avoid heat transfer to the qubits,

the power budget of the control interface is extremely limited to approximately 1 W at the 4-K stage and to less than 10 µW at the mK stage [89, 93]. Additionally, noise leads to decoherence which reduces the qubit fidelity [78]. With these given restrictions near-quantum-limit amplifiers [94–96], ultralow-loss resonators [97, 98], circulators [99], multiplexers [100], cryogenic filters [101], wirebonds [102], DAC/ADCs [103–105], and many more components have been developed to realize cryogenic circuits. These building blocks have recently enabled the design of fully-fledged cryo-CMOS controllers for qubits, located at the 4-K stage [81, 92, 106]. This is an important intermediate step which not only allows to drastically reduce wiring but is also an important step towards fully integrated control electronics and qubits enabling up-scaling, as can be seen in Fig. 1.4. Recently, a first proof-of-concept of a monolithically integrated quantum–classical hybrid circuit was realized [82, 83]. This approach of unifying spin qubits in Si quantum dots and CMOS circuit technology could have the potential of using very large-scale integration (VLSI) and the experience collected over decades in the semiconductor industry for the production of monolithical qubits.

1.2.2 High Performance Computing

Another potential application for cryo-CMOS application-specific integrated circuits (ASICs) is the field of high performance computing. The changes in the $I_D(V_G)$ characteristics of the transistors which are the core of the ASICs towards cryogenic temperatures, can be deliberately used to increase logic switching speed or to decrease the supply voltage V_{DD}, which allows a reduction of the total energy consumption [15–17]. Fig. 1.5 shows qualitatively how the $I_D(V_G)$ curve of a MOSFET changes between room temperature and 4.2 K. At cryogenic temperatures, phonon scattering becomes reduced which leads to a higher charge carrier mobility, and as a consequence to an increasing on-state current I_{ON}. Also the subthreshold slope gets steeper leading to an increasing transconductance g_m. Additionally, the threshold voltage V_{th} increases due to the shifting semiconductor Fermi level and the temperature dependent band gap widening. A detailed discussion of these properties including certain saturation effects is given in Section 6.

While at room temperature V_{DD} can not be scaled down due to the constraining V_{th} and a large SS which would result in a significant off-state current, the steep transition at cryogenic temperatures opens new possibilities. For instance, a scaling of V_{th} allows an operation at considerably lower V_{DD}, as indicated in Fig. 1.5 [16]. There are multiple approaches for achieving the V_{th} scaling, e.g. it is possible to optimize the gate metal by choosing a metal with a band edge work function which meets V_{th} at the reduced target [15]. Additionally, lower band gap channel materials such as SiGe or Ge can be used to reduce V_{th} [15]. Another approach aims for using metal-oxide cap induced interfacial dipole layers [15] to optimize V_{th}. Overall, a device with an optimized V_{th} operating at cryogenic temperatures allowing a low-V_{DD} operation has

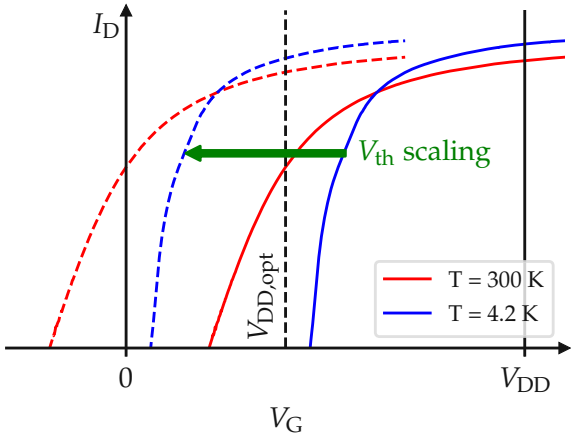

Figure 1.5. The temperature dependence of the charge carrier mobility, the Fermi level, the band gap widening and other MOSFET parameters affects the $I_D(V_G)$ transition curve. Towards cryogenic temperatures the $I_D(V_G)$ curve gets steeper and the on-state current increases. This can be used in combination with threshold voltage V_{th} scaling to decrease the supply voltage V_{DD} which allows to increase the logic switching rate and to reduce total power consumption [15].

the potential of a significant higher logic switching rate. Chiang *et al.* proposed for an advanced FinFET technology a speed increase of around 50% when operating at 77 K with a constant operating power [15]. For a planar 28 nm high-κ metal gate technology, Saligram *et al.* claimed a 90% performance/Watt improvement while operating at 6 K, however, in their calculation the cost of cooling has not been included [107]. But even with including the power of the needed refrigeration, a net power reduction of approximately 30% can be achieved [15].

Next to V_{th} scaling, the operation of MOSFETs at a very low-V_{DD} leads to additional reliability challenges which have to be considered for circuit design for cryo-CMOS applications. With the lowering of V_{DD} the margins for V_{th} and SS also become more narrow meaning that the variability in the transfer curve must be minimized to guarantee a reliable logical switching. Furthermore, reliability issues such as bias temperature instability (BTI) and random telegraph noise (RTN), which are introduced in the upcoming section, must be reduced to a minimum, because low-V_{DD} applications do not leave room for large threshold voltage shifts [108].

1.3 Reliability and Variability Issues in Cryo-CMOS

By operating devices at cryogenic temperatures, new challenges for CMOS engineering arise. Performance and device characteristics are strongly affected by the operation temperature and models developed for room temperature and above considerably loose in accuracy at low temperatures. Therefore, a large effort was put into the development of cryogenic models for transfer characteristics [109] and derived

quantities as subthreshold swing [110], threshold voltage [111], on-state current [112], etc. While the proposed models steadily improve, reliability issues are rarely addressed. However, the stable operation can be affected by a wide range of stress mechanisms, for example caused by radiation, heat, mechanical or electrical stress. These diverse stress mechanisms can have an impact on the transfer characteristics by reducing the performance, increasing leakage currents, or leading in the worst case to a complete oxide breakdown.

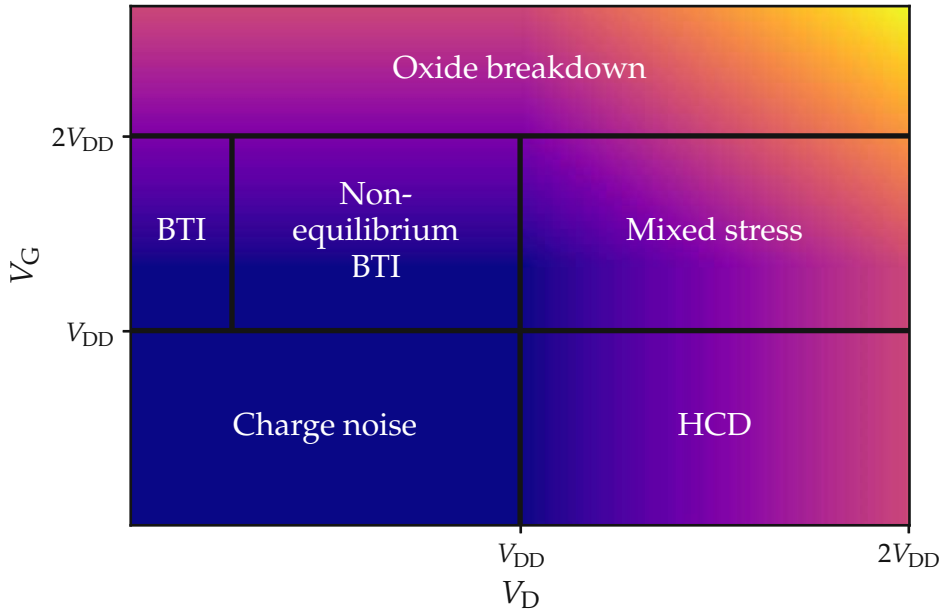

Figure 1.6. A schematic representation of the $\{V_G, V_D\}$-space shows the dominating degradation mechanisms for the different bias regimes.

Device aging is inevitable during regular operation, thus, a highly stable dielectric layer is an important building stone for meeting the lifetime-criterion of electronic circuits. Therefore, the focus of this work will lie on aging mechanisms caused by electrical stress, and the impact of this stress on pre-existing or generated defects in the MOSFET oxide. These defects can capture and emit a charge and thus affect the transfer characteristic of devices. Oxide defects and defects at the interface between oxide and substrate can not only occur in transistors, they can also disturb the stable operation of quantum dots. As discussed in the previous section, QDs are complex structures with multiple oxide layers. Since reaching a high fidelity is one of the main challenges in the engineering of qubits, guaranteeing high quality interfaces between substrate and these oxide layers is essential to avoid charge trapping.

For understanding the mechanisms of charge trapping, it is necessary to differentiate between various degradation mechanisms, which will be dominant at different bias conditions as can be seen in Fig. 1.6

- **Bias temperature instability (BTI)**: BTI occurs, whenever a bias is applied on the gate side of a transistor. The applied bias reduces the barrier between the

charge reservoir, which can be either the channel or the gate, allowing the traps to capture charge. The trapped charges then affect the electrostatics and consequently change the device transfer characteristics. On large-area devices, this can be seen as a continuous shift of the threshold voltage, caused by thousands of defects capturing and emitting charges simultaneously when the threshold voltage is monitored over time. By scaling the geometry of a transistor, discrete steps can be seen in the measured current, which reveal single charge capture and emission events. To examine this effects, typically elevated temperatures and voltages above the operational condition are used to accelerate the degradation. Since elevated temperatures do not make sense for the study of BTI in cryogenic environments, studies in this work will often be based on elevated voltage conditions. Also, it has to be noted that generally BTI is classified into negative BTI (NBTI) and positive BTI (pBTI), determined by the sign of the gate voltage applied during the stress, which can be measured on both nMOS and pMOS. However, typically the focus lies on NBTI on pMOS devices and PBTI on nMOS devices corresponding to the typically used bias conditions. A more detailed introduction to BTI can be found in Section 8 and in [113–115].

- **1/f noise and random telegraph noise (RTN)**: Oxide defects can exchange charges with the device substrate or the gate, even under non-accelerated conditions. The ongoing mechanisms are the same as in BTI, i.e. defects can capture and emit charges, altering the device electrostatics and leading to fluctuations in the device parameters. However, in contrast to BTI, less defects can contribute to the noise. But similar to BTI, while on large-area devices the fluctuations are continuous and known as 1/f noise, there are discrete steps on scaled devices, which are known as random telegraph noise (RTN) [116–118].

- **Hot carrier degradation (HCD)**: Compared to BTI which focuses on elevated gate insulator fields while no drain-source bias is applied to result in a uniform carrier distribution along the channel during stress, HCD describes the degradation under non-equilibrium conditions, i.e. at accelerated drain bias conditions. A high drain-source field accelerates the carriers to either drain or source (depending on the carrier type) leading to carriers with a high kinetic energy. These so-called *hot carriers* cause damage in the channel or at the Si/SiO_2 interface where they can create interface defects which can capture and emit charges. However, unlike BTI, the damage caused by the high energetic carriers is mostly irreversible. A more detailed overview to HCD can be found in [119, 120].

The above mentioned degradation mechanisms have been well studied over the past decades and a profound knowledge has been gained on how the behavior of MOSFETs is affected in various ways. In addition, the community of defect physics has made lots of progress in understanding the role of defects in MOSFET degradation, and fundamental theories like nonradiative multiphonon (NMP) theory have been established to describe

the interaction of defects with charge reservoirs, as will be discussed in detail in Section 3. While there are innumerable reliability studies on reliability issues in electronics in above room-temperature environments, there is a lack of degradation studies in cryogenic environments. However, with the ever increasing interest in QC the research for robust electronic systems operated in cryogenic environments becomes increasingly important. As one part of it, the gained knowledge from defect studies at cryogenic temperatures helps developing devices operating stably in such environments. This is of special importance, because many cryogenic applications are extremely sensitive to reliability issues. As discussed, the classical control interface of quantum computing should be as close as possible to the qubits and thus at the cryogenic stage. At the same time, qubits are extremely noise sensitive, therefore charge noise must be reduced as much as possible in the CMOS control interface. This gets even more critical in monolithical qubits, where the CMOS control unit is as close as possible to the qubits. Here, charge noise is very likely to accelerate dephasing of the qubits and can therefore have a negative impact on the qubit fidelity.

Reliability issues play a crucial role in high performance computing. A low-V_{DD} is important for electronics in HPC to increase the logical switching rate and to decrease the total power consumption. However, it comes with the price that circuits become very sensitive to variability. A large variability of the devices has the effect that a switching between on- and off-state can not be guaranteed anymore within low-V_{DD} applications. Additionally, small ΔV_{th} shifts can already disturb the stable switching and thus compromise a reliable operation. Therefore, it is essential to include variability and reliability considerations in design of optimized CMOS applications [121].

Knowledge on defect physics gained from reliability studies on cryo-CMOS applications can be further used in related fields. Since cryogenic environments are often used to reduce noise, charge noise caused by oxide defects or by interface defects between the numerous different processed layers (see Fig. 1.3 (c)) is a major concern for qubits. In the qubit design, one of the main engineering challenges is reaching a high fidelity and avoiding decoherence. Defects which get created during device manufacturing can affect the readout of the spin in superconducting qubits [122–125] or have an impact on the confinement of the QD forming the electron spin qubits causing dephasing [78, 80, 126–132]. Therefore, understanding the physics of the occurring traps from theoretical calculations [133–135] and from experiments as shown in this work can deliver a valuable contribution in reducing the loss of decoherence in solid state qubits.

1.4 Scope and Outline of this Work

In this work, two approaches for developing a better understanding of MOSFET reliability at croygenic temperatures are followed. A theoretical, modeling approach presented in Part I *Modeling and Simulation of Defects*, in which an efficient charge trapping model for cryogenic temperatures is developed. In Part II *Defect Characterization at*

Cryogenic Temperatures, various characterization, reliability and variability studies on multiple technologies are presented and analyzed with the models from Part I.

Part I – Modeling and Simulation of Defects

- ***Chapter 2 – Defect Candidates***
  This chapter reviews specific defect candidates which have been found to play a major role in device degradation during operation. The defect candidates for SiO_2, SiON and HfO_2 insulators have been studied in detail with theoretical tools such as density functional theory (DFT) or technology computer-aided design (TCAD), as well as experimentally. This allows the calculation and verification of various trap properties like densities, trap levels, relaxation energies or their typical spatial positions.

- ***Chapter 3 – Charge Transfer Models***
  Charge trapping in oxides is a stochastic process which describes the exchange of charge carriers between a trap and a reservoir. A mathematical description of this trapping and detrapping process is essential for reliability studies. First models developed in the 1950s have been steadily improved for allowing precise simulations of measurements. In this work, the 2-state nonradiative multiphonon (NMP) models is used. This model is valid even in the limit of cryogenic temperatures, however, it is computationally expensive and thus not suitable for reliability simulations. Therefore, an efficient WKB-based charge transition model is derived in this work. This model is numerically superior and gives very precise results compared to the full model. An even more efficient model is presented, which is based on the assumption of non-deforming potential energy surfaces during charge transitions.

- ***Chapter 4 – Reliability Simulations***
  The charge transition models presented in Chapter 3 are now used for reliability simulations. For this, the WKB-based approximation of the 2-state NMP model was implemented in the device reliability simulator *Comphy*. *Comphy* uses a surface potential based 1-dimensional electrostatics. By sampling thousands of defects in the oxide, the electrical response of these defects after applying a gate voltage allows to compute the caused threshold voltage shift. These simulations are used to compare the theoretical studies with measurements.

Part II – Defect Characterization at Cryogenic Temperatures

- ***Chapter 5 – Measurement Setup and Technologies***
  Two different measurement setups have been used within this work to characterize

various MOSFET technologies. On cryogenic probe stations, electrical measurements between 4.2 K and room temperature have been conducted using a Helium cooling system. Manual probe arms allowed the characterization of large-area planar high-κ-metal gate devices, planar SiON devices and planar MoS_2 devices. Since manual probing is a very time-consuming task and only a handful of devices can be characterized at once, a different approach was chosen to obtain statistical data representative for a certain technology. *SmartArrays* with thousands of transistors which can be addressed digitally have been characterized. This allows collecting large data sets and enables the study of different designs at cryogenic temperatures. In this work, time-zero variability and mismatch and their dependence on temperature and device geometry have been studied, which can be very critical especially for low-V_{DD} applications.

- ***Chapter 6 – Time-Zero Characterization***
  Time-zero studies form the foundation for every reliability study. The conduction of $I_D(V_G)$ curves allows the analysis of various MOSFET parameters such as threshold voltage, subthreshold swing or on-state current and their temperature dependence. The modeling of the temperature dependence of these parameters is extremely complex and must not only include the shifting Fermi level, the changing Fermi-Dirac statistics or bandgap widening but also interface and band-edge states. Time-zero characterization on *SmartArrays* allows studying the variability and the mismatch of the extracted parameters and their temperature dependence.

- ***Chapter 7 – Charge Noise Characterization***
  Cryogenic environments are often used to minimize noise in the measurements. Thus, it is extremely important to understand the impact of low frequency noise (1/f noise) caused by traps at cryogenic temperatures. For this, the temperature and gate voltage dependence of RTN in various technologies has been studied. The findings have then be modeled with the quantum mechanical 2-state nonradiative multiphonon theory discussed in Chapter 3, which can correctly describe the temperature activation of charge trapping.

- ***Chapter 8 – Bias Temperature Instability***
  In this chapter BTI studies on various technologies are presented. The different freeze-out behavior of large area SiON and high-κ metal-gate devices is discussed and modeled using the reliability simulator *Comphy*. For this, the quantum mechanical version of 2-state NMP theory has been used. Furthermore, BTI was studied on scaled MoS_2 devices using time-dependent defect spectroscopy (TDDS) measurements at cryogenic temperatures. Here, it is explicitly shown that single charge capture and emission events causing threshold voltage shifts are clearly visible at cryogenic temperatures.

Part I

Modeling and Simulation of Defects

Chapter 2

Defect Candidates

Reliability issues such as BTI, HCD, or charge noise, i.e. RTN, are caused by electrically active defects that can be located in the oxide or at its interface to the semiconductor bulk. These defects are introduced in the device during manufacturing, even in very optimized processes, or can be created during device operation at nominal bias conditions and temperatures and their formation can not be avoided. Since these defects typically have a negative impact on the device performance or, in the worst case, can even lead to the failure of the device, it is of major importance to understand how defects emerge and how they influence the device performance. Over the years, the manufacturing processes have considerably improved and as a consequence the number of active defects could be significantly reduced. Still, there are many applications which rely on an extremely low noise level and very narrow reliable operation conditions which is further motivation to investigate the physical processes involved in charge trapping. Since defects occur in a large variety, many different approaches are used for examinations, from theoretical tools such as DFT or TCAD, electrical examination tools as presented in Part II of this work, or physical characterization methods such as electron paramagnetic resonance (EPR) spectroscopy, X-ray photoelectron spectroscopy (XPS), or secondary ion mass spectroscopy (SIMS).

The combination of all these tools allows to search for specific defect candidates, which may play a major role in the device degradation during operation. The trapping properties of suitable defect candidates, such as relaxation energies, defect densities, trap level, and spatial position can be theoretically studied and linked to various experimental approaches using reliability simulations as presented in Chapter 4. The most important defect candidates for electrical reliability studies are introduced in the following sections.

2.1 Interface Defects

Silicon dioxide is the most common insulator in MOSFET technologies. During the fabrication of devices, SiO_2 grows natively on silicon wavers when exposed to air via

oxidation, which makes it an ideal candidate as an insulator. The resulting thin layer (with a thickness below 1 nm) is called native SiO_2. Alternatively, nitrited SiO_2 (SiON) is used as insulator to decrease the gate leakage current and to prevent dopants from diffusing from the channel into the insulator. The defect behavior in SiON insulators is similar to the one in pure SiO_2 , which typically allows using the simplified SiO_2 system in simulations [136].

The lattice mismatch between the amorphous structure of the insulator layer and the semiconductor makes the presence of a certain amount of defects close to the interface inevitable. The interface defect density strongly depends on the combination of the materials used for the semiconductor and insulator and the processing conditions. Note that for Si/SiO_2 systems densities on the order $10^{10}\,cm^{-2}eV^{-1}$ can be achieved [137]. For this material system, the best understood interface defect is the P_b center, which can appear on both, the (111) and the (100) oriented surface [138, 139]. In general, the P_b centers have been observed to occur in different chemical variations: The P_{b0} center of the form ($\cdot Si{\equiv}Si_3$) and the P_{b1} center, whose atomistic structure is not completely understood yet, are depicted in Fig. 2.1. Due to the unpaired electron at the Si side, the defects are also called *dangling bonds*.

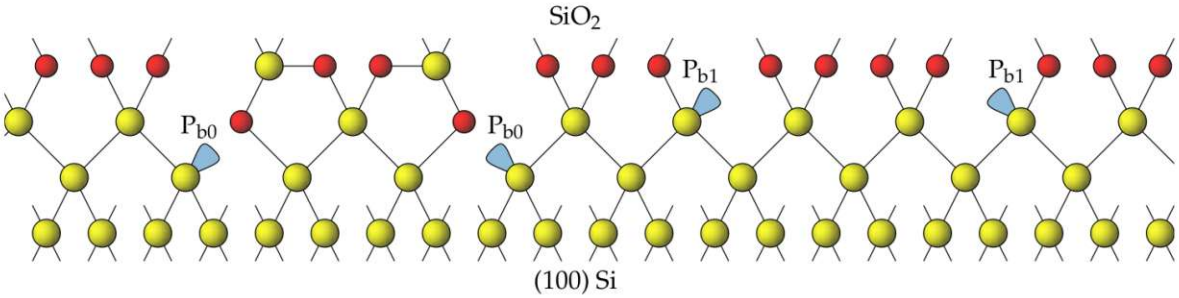

Figure 2.1. P_b centers can occur in two chemical variations at the (100) oriented Si/SiO_2 interface: As P_{b0} center of the form ($\cdot Si{\equiv}Si_3$) and as P_{b1} center, whose atomistic structure is not completely understood yet. The unpaired *dangling bond* is shown represented by the bluish triangles. Figure reprinted from [140].

There are at least two trapbands corresponding to the P_b centers, one above and one below Si midgap, which allow electron capture and emission from and to the channel, respectively [141]. The trap levels of such interface defects lie within these trap bands, and they typically have small relaxation energies which leads to small relaxation times compared to bulk oxide defects [142]. Thus, the different relaxation energies can be used as an indicator which defects might be responsible for measured signals, as done e.g. in Section 7.4.

The oxidation of amorphous SiO_2 on the Si substrate leads not only to interface defects but also to a continuous change in the density of states and the corresponding band edges [136], as shown in Fig. 2.2. This behavior can be explained by a gradual change in the ratio of silicon and oxygen atoms between bulk SiO_2 and the Si substrate. The gradual change in the density of states and linear band edge structure has to be be

considered in device simulations, e.g. by correcting the WKB tunneling factors [136]. Fig. 2.2 shows the transition layer from amorphous bulk SiO_2 to the Si substrate which shows a thickness of 0.3 nm to 0.4 nm. This has the consequence that in high-κ devices, which typically have a very thin SiO_2 layer of below 1 nm between the substrate and the high-κ material, no bulk SiO_2 occurs.

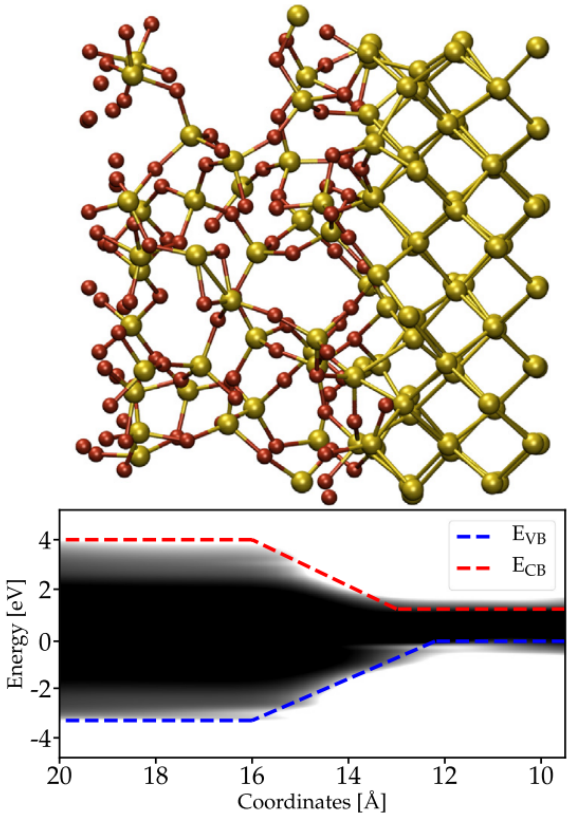

Figure 2.2. The density of states does not change abruptly between amorphous bulk SiO_2 and Si at the interface but is rather continuous. This can be explained by the continuously changing ratio of silicon (yellow spheres) and oxygen (red spheres) atoms from the Si substrate to bulk SiO_2 and the influence of the Si surface. Figure taken from [136].

2.2 Bulk Oxide Defects

While in the early days of MOSFETs interface defects have been in the center of research, later also defects in the bulk oxide were added to a more complete description of device degradation.

2.2.1 Oxide Defects in Amorphous SiO_2 (a-SiO_2)

Bulk SiO_2 is very well understood from both experimental characterization and theoretical studies. One intensively investigated defect candidate for a-SiO_2 is the *oxygen vacancy* (OV) [143–145]. It forms naturally during oxidation and is essentially a missing

O atom in a SiO_2 network (where an O atom ideally bonds with two neighbored Si atoms) which leads to a Si-Si bond. The defect is shown in the neutral (1) and positive (2) charge state in Fig. 2.3, whereby the prime $'$ denotes a metastable configuration which can be accessed via a thermal transition, as explained in Section 3.3. When the neutral OV traps a hole, the distance between the two Si atoms increases. As a consequence, the Si-Si bond can break, resulting in a Si dangling bond [146].

As shown in Fig. 2.3, the bond breaking and relaxation into a new stable configuration of the oxygen vacancy could be theoretically described by a 4-state model with the states $1, 1', 2'$, and 2. However, various first principle studies show that the energetic position of the defect inside the band structure of the insulator is too far below the valence band or too far above the conduction band of the Si substrate [147–149]. Therefore, oxygen vacancies are very unlikely to capture holes and stay neutral in pMOS transistors with relatively thin a-SiO_2 oxides and thus do most likely not contribute to BTI or RTN.

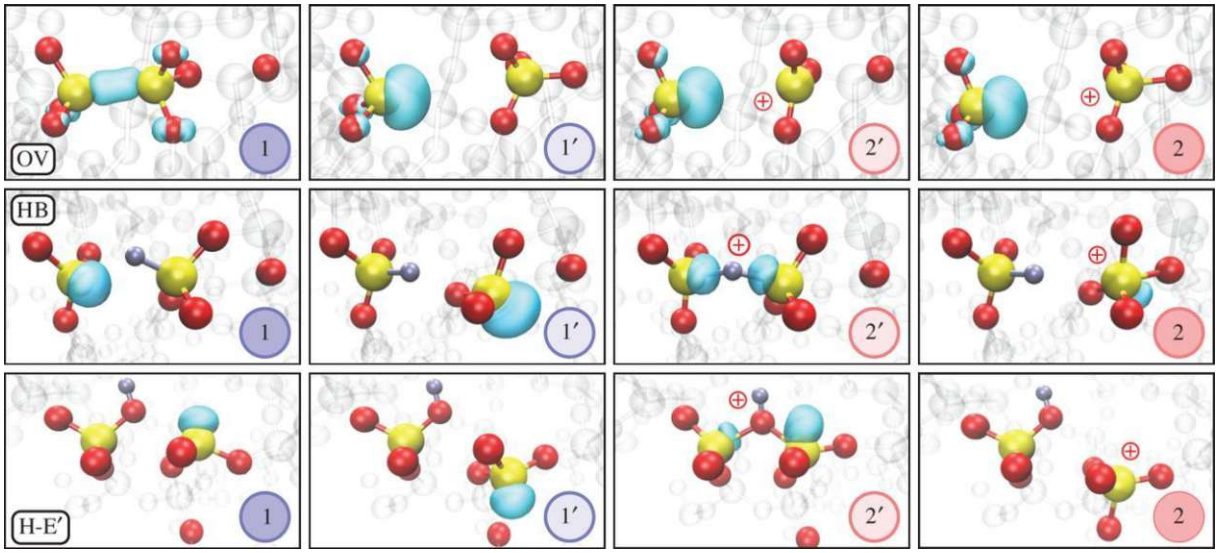

Figure 2.3. The atomistic configurations of the oxygen vacancy (OV), the hydrogen bridge (HB) and the hydroxyl-E′ center (H-E′) show that these defects can be described with a 4-state NMP model. The defects show two stable configurations 1 and 2 and can switch between theses stable states via the meta-stable states $1'$ and $2'$. The turquoise bubbles show the highest occupied molecular orbital (HOMO) electron distributions. Figure reprinted with permission of the Royal Society from [149].

Recent studies suggested that hydrogen related defects are responsible for reliability degrading phenomena like BTI and RTN [150, 151]. As hydrogen is commonly used in device processing to passivate dangling bonds at the interface between semiconductor and SiO_2, the formation of defects involving H is very likely [151]. In addition, H has the ability to easily diffuse in the SiO_2 layer where it can form electrically active defects such as the *hydrogen bridge* (HB) and the *hydroxyl-E′ center* (H-E′) [152, 153].

The H atoms diffusing into the SiO_2 layer can react with oxygen vacancies, which other ways would stay neutral. When the H atom binds to a Si atom, the resulting defect

configuration which involves two Si atoms is called hydrogen bridge (HB). In its neutral configuration, a bond is formed between the H atom and a single Si atom. As can be seen in Fig. 2.3, the Si bonded to the H atom can transit through the triangle spanned by oxygen atoms and the hydrogen atom can then bind to the second Si atom, leaving the first one uncoordinated. By capturing a hole, the HB can form the stable positive configuration 2, which is called puckered. The transition from the unpuckered to the puckered configuration takes place by passing one of the two metastable states 1′ or 2′.

While in the HB configuration the hydrogen atom binds to the silicon atom, it can also have a binding with the oxygen atom and form a hydroxyl-E′ center (H-E′ center). These defects can only occur in a-SiO_2, because they need strained Si-O bonds to get formed. The hydrogen atom can break such a strained bond and bind to the oxygen, forming a hydroxyl group (OH). The same puckering configuration which occurred in the OV and the HB can also occur for the H-E′ center and the trap dynamics can be approximated with a 4-state model.

While OV have trap levels far below the valence band or far above the conduction band of Si, hydrogen related defects show trap levels close close to the Si band edges, as can be seen in Fig. 2.4. Therefore, they are much more important in terms of understanding the mechanisms behind RTN and BTI.

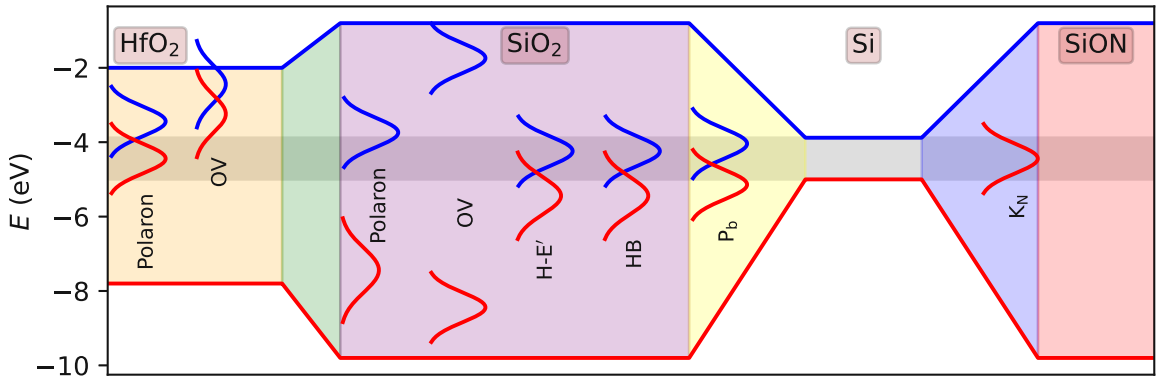

Figure 2.4. The trap levels of various studied defect candidates in HfO_2, SiO_2 and SiON are spread across the band gap of all host materials. The trap levels of oxygen vacancies (OV) are far away from the Si conduction band and therefore play a minor role in BTI and RTN. The trap levels of hydrogen related defects, as the hydroxyl-E′ center or the hydrogen bridge, and interface defects, as the P_b center or the K_N center are on the other hand close to the Si band edges, which makes them highly relevant in terms of BTI. In addition, intrinsic charge trapping caused by polarons is responsible for electron trapping in high-κ devices.

While most defect studies focus on charge trapping caused by defects, the explanation of some experimental studies require trap bands which can not directly be linked to any known defect candidates [154, 155]. This type of charge trapping can play a major role in wide bandgap devices as SiC. It can happen spontaneously in a-SiO_2 and is caused by structures with elongated Si-O bonds, which allow the localization of an additional electron at a certain Si atom [156]. Their trap band lies close to the Si conduction band edge as can be seen in Fig. 2.4.

2.2.2 Oxide Defects in SiON

Plasma-nitrided oxides (PNO) are routinely employed in Si technology to decrease the gate leakage current [157]. However, this comes at the price of a high concentration of nitrogen atoms at the oxide-semiconductor interface. For these oxides, K_N centers become the dominating defect configuration instead of P_b centers [158–160]. K_N centers occur in silicon nitride Si_3N_4 and form during nitration. The central Si atom is back-bonded to an N atom and forms a dangling bond, in which the dangling Si atom is connected to three neighboring N atoms [161, 162]. SDR and SDT magnetic resonance measurements show that K_N centers are located in the vicinity of the interface region of PNO devices and suppress P_b centers. Thus, while in SiO_2 devices P_b centers are central for the NBTI behavior, NBTI is dominated by K_N centers in PNO devices [157]. While the defect candidates of near interface defects changes, hydrogen related defects can still play a major role in nitrided oxides, because H atoms are used for passivating P_b centers [113, 163]. In contrast to P_b centers, which follow a broad energy distribution with at least two peaks, one directly above and one directly below midgap, K_N centers are very narrowly distributed and the distribution of their energetic trap level is very close to the Si-midgap, see Fig. 2.4 [157].

2.2.3 Oxide Defects in HfO_2

The continuous miniaturization of MOSFETs has lead to a constant scaling of the insulator thickness to increase the gate capacitance and thus the drive current. However, with decreasing insulator thickness, leakage currents increase, which leads to several problems, such as higher steady state power consumption which in turn results in heat dissipation and a reduced device reliability. Therefore, the development of alternative insulators with comparably large dielectric constants (SiO_2: $\varepsilon_r = 3.9$), such as the so-called high-κ materials HfO_2 ($\varepsilon_r = 25$), ZrO_2 ($\varepsilon_r = 25$) or Al_2O_3 ($\varepsilon_r = 9.1$) became necessary to increase the gate capacitance without introducing leakage effects.

Compared to SiO_2 , HfO_2 shows a higher degree of disorder as can be seen in Fig. 2.5. Thus, intrinsic charge trapping is suspected to be a central reliability issue in HfO_2 based oxides [165]. Contrary to SiO_2, HfO_2 is a non-glass forming oxide, which means that it has varying coorination numbers depending on the oxide phase. Furthermore, it is known that HfO_2 can (partially) crystallize during annealing which introduces additional structural uncertainties [166]. This opens possibilities for various defect candidates and makes modeling of HfO_2 insulators extremely challenging. It is proposed that HfO_2 can form electron polarons, which can be stable at room temperature and could be responsible for PBTI. The formation is driven by uncoordinated Hf-ions which can trap electrons. A trapped electron can localize at multiple Hf sites, enabling the formation of a ring where the Hf atoms are pulled towards the center while the O anions are pushed outwards [165, 167].

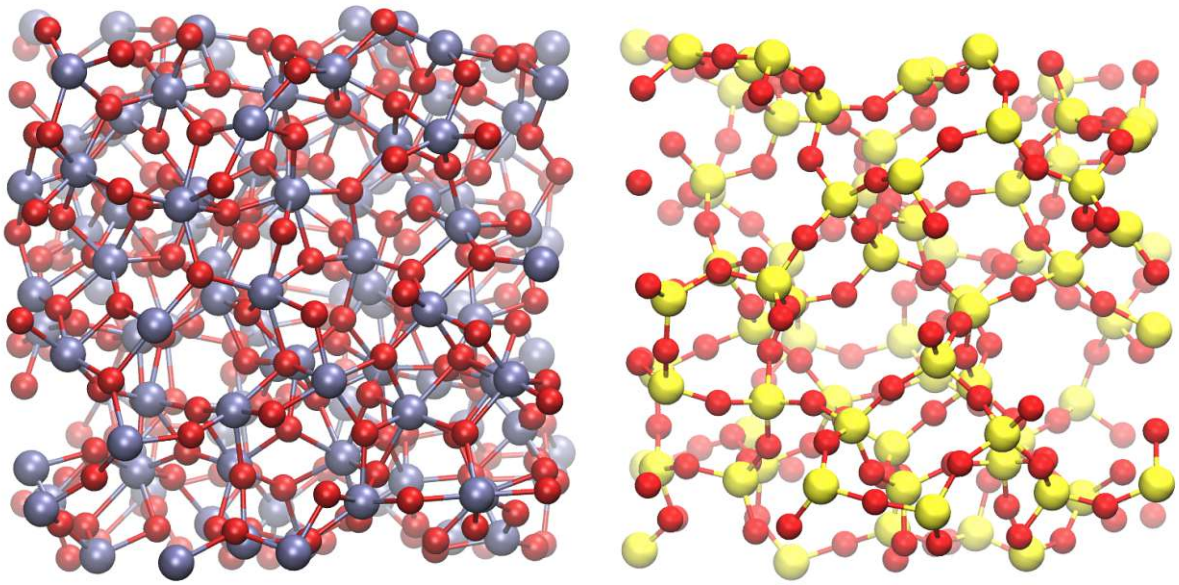

Figure 2.5. The atomistic structure of a-HfO_2 (Hf - silver, O - red) shows a higher degree of disorder than a-SiO_2 (Si - yellow, O - red). This is a source for intrinsic charge trapping and could explain the PBTI behavior of high-κ devices. Figures reproduced from [164].

Some of the defect types discovered in a-SiO_2 can also be found in a-HfO_2. For example, the OV occurs in HfO_2, where it is known to form neutral, positive and negative states [168]. It is assumed that the negative states could be responsible for the high density of electron traps observed in high-κ-devices [136, 169]. Various works show trap levels of 1.2 eV to 1.8 eV below the conduction band of HfO_2 which have been associated with oxygen vacancies [169, 170]. However, recent first principle studies suggest that the energy level of oxygen vacancies lies closely below the conduction band of HfO_2. According to these calculations, the defect bands measured in [136, 169, 170] can not be linked to oxygen vacancies.

As has been discussed, several trap candidates have been identified that affect the high-perfomant operation of MOSFETs. These defects can exists in various configurations which exhibit different properties. In order to describe the defects and make their impact accessible to device simulators, various models have been discussed in literature which will be presented in the following chapter.

Chapter 3

Charge Transfer Models

Charge trapping is a stochastic process and refers to the exchange of charge carriers between a defect and a charge reservoir. A mathematical description of the trapping and detrapping probabilities is central for making this process available to simulators and to enable reliability studies of MOSFETs. Note that charge trapping is a dynamic process which is determined by the applied biases, temperatures as well as intrinsic properties of the particular defect like its trap level and relaxation energy. As such, the development of a theoretical model which covers all features of charge trapping is a formidable challenge.

The very first model for describing charge transition rates was proposed by Hall [171] and Shockley and Read [172] in the 1950s, and is thus called Shockley-Read-Hall (SRH) model. This model is well suited for describing trapping at the interface with little relaxation, which was one of the biggest fabrication issues in the first generation of transistors as described in Section 1.1. This model will be outlined in Section 3.1. However, the SRH model can not accurately describe the bias and temperature dependence observed in many BTI and RTN experiments, where oxide defects are involved, which typically show a significant relaxation upon charge transfer. In order to account for such relaxation effects, the nonradiative multiphonon (NMP) model is introduced, which describes charge transition events accurately. While for many applications a 2-state NMP model is sufficient for describing the defect kinetics, there are phenomena such as anomalous RTN (aRTN), which prove the existence of additional metastable states and thus require an extension to a 3 or 4-state model.

In Section 3.2 the 2-state NMP model will be introduced. Since the calculation of the full quantum mechanical transition rate is computationally expensive, a classical approximation is typically used for above room temperature applications. However, this classical approximation is not suitable for cryogenic applications, because it does not take nuclear tunneling into account. Therefore, a WKB-based, numerically efficient approximation has been developed, which is presented and benchmarked in Section 3.2.1 and Section 3.2.2, respectively, based on [MJJ1]. An even more efficient approximation can be used when assuming that the potential energy surfaces which represent a defect

are undistorted before and after a charge transition, see Section 3.2.3. In Section 3.3 the 4-state NMP model will outlined, which can be necessary for describing single defect measurements.

3.1 Shockley-Read-Hall Model

With the Shockley-Read-Hall model, defects are described by a neutral and a charged configuration state. The defects can either be acceptor- or donor like. Acceptor-like defects accept an electron during charge trapping and thus have a negative charged state, while donor-like defects donate an electron and become positively charged. Additionally, the terminology of electron- and hole-like defects is frequently used, depending on the charged state. This leads to four different processes shown in Fig. 3.1:

- Electron capture: A defect traps an electron from the conduction band, which is represented by $0 \rightarrow -$. This is equivalent to a hole transitioning from the defect to the conduction band, represented by $+ \rightarrow 0$.
- Electron emission: A defect releases an electron to the conduction band, which is represented by $- \rightarrow 0$. This is equivalent to a hole captured from the conduction band, which is represented by $0 \rightarrow +$.
- Hole capture: A defect traps a hole from the valence band, which is represented by $0 \rightarrow +$. This is equivalent to an electron release from the defect to the valence band represented by $- \rightarrow 0$.
- Hole emission: A defect emits a hole to the valence band, which is represented by $+ \rightarrow 0$. This is equivalent to an electron capture from the valence band, represented by $0 \rightarrow -$.

Analogously to the derivation in [140], the hole capture process is used in the following to present the mechanics of the SRH model. However, the derivation can be done in a similar fashion for hole emission and electron trapping. In the case of hole capture, the defect can be in the initial neutral configuration 0 and in the charged configuration $+$. In the SRH model, capturing a hole from the valence band is possible without overcoming a barrier, leading to a hole capture probability c_p which is often written as being directly proportional to the thermal velocity $v_{\text{th}} = \sqrt{8k_{\text{B}}T/(\pi m_{p,\text{eff}})}$ of the charge carriers

$$c_p \approx v_{\text{th}}\sigma, \tag{3.1}$$

with the electron temperature T, the effective hole mass $m_{p,\text{eff}}$ and the capture cross section σ. By taking into account the hole occupancy $f_p(E)$ which follows a Fermi-Dirac

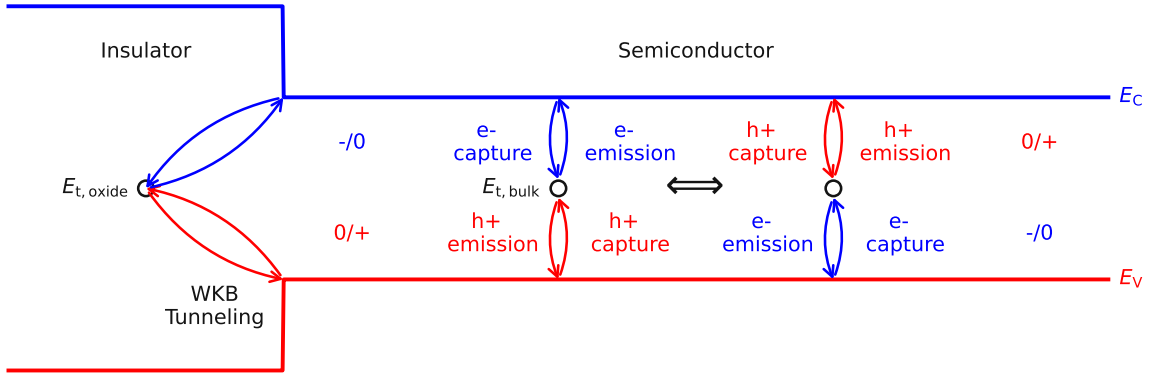

Figure 3.1. Charge transitions between a neutral and a charged configuration in the bulk can be divided into four different processes: Electron capture and emission from/to the conduction band, which is equivalent to hole emission/capture to/from the CB and hole capture and emission from/to the valence band, which is equivalent to electron emission/capture to/from the VB.

distribution, and the density of states D_p, the capture rate can be computed by

$$k_{0+} = \int_{-\infty}^{E_\mathrm{v}} c_p(E) f_p(E) D_p(E) \mathrm{d}E. \tag{3.2}$$

Similarly, the hole emission rate is given by

$$k_{+0} = \int_{-\infty}^{E_\mathrm{v}} e_p(E)(1 - f_\mathrm{p}(E)) D_p(E) \mathrm{d}E, \tag{3.3}$$

by replacing the capture probability with the emission probability e_p and considering the Fermi-Dirac distribution of the electrons f_n. Using

$$\frac{f(E)}{1 - f(E)} = \mathrm{e}^{-\beta(E - E_\mathrm{F})} \tag{3.4}$$

with E_F being the Fermi level and $\beta = 1/k_\mathrm{B}T$, the rate can be expressed as

$$k_{+0} = \int_{-\infty}^{E_\mathrm{v}} e_p(E) f_p(E) \mathrm{e}^{-\beta(E - E_\mathrm{F})} D_p(E) \mathrm{d}E. \tag{3.5}$$

The principle of detailed balance states that for an arbitrary energy E the probability of being in state 0, $f_0 = f(E_\mathrm{T})$ times the transition rate from state 0 to $+$ is equal to the reversed process, which is the probability of being in state $+$ times the transition rate from $+$ to 0

$$f_0 \mathrm{d}k_{0+}(E) = f_+ \mathrm{d}k_{+0}(E). \tag{3.6}$$

From detailed balance the following relation between emission and capture probability can be derived directly as

$$e_{\mathrm{p}}(E) = c_{\mathrm{p}}(E)\mathrm{e}^{\frac{E-E_{\mathrm{T}}}{k_{\mathrm{B}}T}}. \tag{3.7}$$

This allows expressing the hole emission rate

$$k_{+0} = \int_{-\infty}^{E_{\mathrm{v}}} c_p(E)\mathrm{e}^{\beta(E-E_{\mathrm{T}})} f_p(E)\mathrm{e}^{-\beta(E-E_{\mathrm{F}})} D_p(E)\mathrm{d}E. \tag{3.8}$$

For non-degenerated semiconductors, the Fermi-Dirac statistics can be approximated by the Boltzmann statistics. If additionally the band edge approximation is used, which allows to assume that all carriers are concentrated at the band edges, the integral in (3.8) can be carried out analytically leading to the simplified transition rates

$$\begin{aligned} k_{0+} &= p v_{\mathrm{th}} \sigma \\ k_{+0} &= N_{\mathrm{v}} v_{\mathrm{th}} \sigma \mathrm{e}^{\beta(E_{\mathrm{v}}-E_{\mathrm{T}})} \end{aligned} \tag{3.9}$$

with p being the hole concentration and N_{v} the valence band weight. The SRH model is typically used for bulk and interface defects. To account for oxide traps, the capture cross section σ can be expressed as $\sigma = \sigma_0 \lambda$ with λ being a Wentzel-Kramers-Brillouin (WKB) tunneling factor which describes the elastic tunneling from the charge reservoir to the defect, see Fig. 3.1. This WKB-tunneling factor allows in principle to include oxide defects [173]. However, bias and temperature dependence of transition rates of oxide defects can not be modeled accurately with the SRH model, therefore a more precise description is needed [174].

3.2 2-State Nonradiative Multiphonon Model

The reason for the inaccurate bias and temperature dependence within the SRH model is that it ignores the structural reconfiguration at the defect site arising from changing charge states and the electrostatic shift of the trap level due to the oxide field [174]. Since this deformation involves the emission/absorption of phonons, it is necessary to include the coupling to a surrounding phonon bath in the physical defect model to provide a physical description of the whole system including all involved electrons and nuclei. Within first order perturbation theory, non-adiabatic transition rates in a system with both electrons and nuclei can be represented by their electronic and vibrational wavefunctions $|\phi_i\rangle$, $|\phi_j\rangle$ and $|\eta_{i\alpha}\rangle$, $|\eta_{j\beta}\rangle$, respectively. The non-adiabatic charge transition rates can be obtained using Fermi's golden rule

$$k_{i\alpha,j\beta} = |M_{i\alpha,j\beta}|^2 \delta(E_{i\alpha} - E_{j\beta}). \tag{3.10}$$

where the matrix element M is

$$M_{i\alpha,j\beta} = \langle \eta_{i\alpha} | \langle \phi_i | \hat{H}' | \phi_j \rangle | \eta_{j\beta} \rangle . \tag{3.11}$$

Here, i and j represent a neutral and a charged state, α and β are the corresponding vibrational eigenstate of i and j, respectively, with eigenenergies $E_{i\alpha}$ and $E_{j\beta}$. The coupling operator $\hat{H}'$ describes the perturbation, while the Dirac delta distribution $\delta(E_{i\alpha} - E_{j\beta})$ ensures energy conservation. The form of the vibrational wavefunctions depends on the Hamiltonian. Commonly a harmonic oscillation is used as a model Hamiltonian to describe the system as shown in Fig. 3.2. According to the Born-Oppenheimer approximation the electronic and vibrational wavefunctions can be treated separately, because the nuclei are much heavier than the electrons and thus there is a large difference in their time scales of motion. Hence the electrons can always be assumed to be instantaneously in the ground state of the potential produced by the positions of the nuclei, while the nuclei themselves move in an effective potential produced by the ground state electron distribution. Therefore, the Hamiltonian can be split in an electronic and a vibrational component

$$\hat{H}' = \hat{H}'_{\text{el}} + \hat{H}'_{\text{vib}}, \tag{3.12}$$

which allows to factorize the transition matrix element as follows

$$\begin{aligned} M_{i\alpha,j\beta} &= \langle \eta_{i\alpha} | \langle \phi_i | \hat{H}'_{\text{el}} + \hat{H}'_{\text{vib}} | \phi_j \rangle | \eta_{j\beta} \rangle \\ &= \langle \eta_{i\alpha} | \hat{H}'_{\text{vib}} | \eta_{j\beta} \rangle \langle \phi_i | \phi_j \rangle + \langle \phi_i | \hat{H}'_{\text{el}} | \phi_j \rangle \langle \eta_{i\alpha} | \eta_{j\beta} \rangle \\ &= \langle \phi_i | \hat{H}'_{\text{el}} | \phi_j \rangle \langle \eta_{i\alpha} | \eta_{j\beta} \rangle . \end{aligned} \tag{3.13}$$

This final simplification is called Franck-Condon principle and $\langle \eta_{i\alpha} | \eta_{j\beta} \rangle$ are referred to as Franck-Condon factors [175–177]. Using this simplification the transition rate (3.10) can be represented by

$$k_{i\alpha,j\beta} = A_{ij} I_{i\alpha,j\beta}, \tag{3.14}$$

with the overlap integral

$$I_{i\alpha,j\beta} = | \langle \eta_{i\alpha} | \eta_{j\beta} \rangle |^2. \tag{3.15}$$

The electronic matrix element A_{ij} describes the coupling between the respective electronic wavefunctions in the initial and final state, i.e. the coupling between the localized defect state and the delocalized valence/conduction band state of the charge reservoir. The vibrational overlap $I_{i\alpha,j\beta}$ is a measure for the nuclear tunneling probability. For a transition from one charge state to another all vibrational modes of the charge states must be taken into account. These vibrational modes can be described as a canonical

ensemble resulting in the total transition rate

$$k_{ij} = A_{ij} \underset{\alpha}{\text{ave}} \sum_{\beta} I_{i\alpha,j\beta} \tag{3.16}$$

where

$$\underset{\alpha}{\text{ave}} = \frac{1}{Z} \sum_{\alpha} \mathrm{e}^{-E_{i\alpha}/k_\mathrm{B}T} \tag{3.17}$$

is the weight of the excited initial vibrational wavefunction $i\alpha$ and Z is the partition function

$$Z = \sum_{\gamma} \mathrm{e}^{-E_{i\gamma}/k_\mathrm{B}T}. \tag{3.18}$$

This expression allows the definition of the lineshape function f_{ij}^{LSF}

$$f_{ij}^{\mathrm{LSF}} = \underset{\alpha}{\text{ave}} \sum_{\beta} |\langle \eta_{i\alpha} | \eta_{j\beta} \rangle|^2 \delta(E_{i\alpha} - E_{j\beta}) \tag{3.19}$$

which contains all interactions of the vibrational states of the nuclei. The total transition rate (3.16) describes the transfer of a charge carrier between two electronic states. However, in a MOSFET device the states within the conduction and valence band of the semiconductor or in the metal gate form a continuum of states. To take all these states into account it is necessary ot integrate over all states of a band obtaining the rate equations

$$\begin{aligned}
k_{ij}^{\mathrm{CB}} &= \int_{E_\mathrm{C}}^{\infty} D_n(E)(1 - f_n(E)) A_{ij}(E, E_\mathrm{t}) f_{ij}^{\mathrm{LSF}}(E, E_\mathrm{t}) \mathrm{d}E \\
k_{ji}^{\mathrm{CB}} &= \int_{E_\mathrm{C}}^{\infty} D_n(E) f_n(E) A_{ji}(E, E_\mathrm{t}) f_{ji}^{\mathrm{LSF}}(E, E_\mathrm{t}) \mathrm{d}E \\
k_{ij}^{\mathrm{VB}} &= \int_{-\infty}^{E_\mathrm{V}} D_p(E) f_p(E) A_{ij}(E, E_\mathrm{t}) f_{ij}^{\mathrm{LSF}}(E, E_\mathrm{t}) \mathrm{d}E \\
k_{ji}^{\mathrm{VB}} &= \int_{-\infty}^{E_\mathrm{V}} D_p(E)(1 - f_p(E)) A_{ji}(E, E_\mathrm{t}) f_{ji}^{\mathrm{LSF}}(E, E_\mathrm{t}) \mathrm{d}E
\end{aligned} \tag{3.20}$$

with D_n, D_p being the density of states in the conduction and valence band, respectively, and f_n, f_p the Fermi-Dirac distribution for electrons and holes. Since in a semiconductor the charge carriers are located close to the conduction band edge E_C and the valence band edge E_V, it can be assumed that in a first order approximation the electronic coupling elements $A_{ij/ji}$ and the lineshape functions $f_{ij/ji}^{\mathrm{LSF}}$ only depend on the band edges. This is called *band edge approximation* and allows to factor out $A_{ij/ji}(E_{\mathrm{C/V}}, E_\mathrm{t})$ and $f_{ij/ji}^{\mathrm{LSF}}(E_{\mathrm{C/V}}, E_\mathrm{t})$ from the integrals in (3.20).

The electronic coupling element $A_{ij}(E, E_t)$ can be approximated by [174]

$$\begin{aligned} A_{ij}(E_C, E_t) &\approx v_{th,n}\sigma_{0,n}\vartheta_n \\ A_{ij}(E_V, E_t) &\approx v_{th,p}\sigma_{0,p}\vartheta_p \end{aligned} \tag{3.21}$$

for electrons and holes, respectively, with $v_{th} = \sqrt{8k_B T/(\pi m_{eff})}$ the thermal velocity, σ_0 the capture cross section and ϑ being a tunneling factor, accounting for the spatial separation of the defect and the charge reservoir. The tunneling factor can be determined by solving the stationary Schrödinger equation. However, in most cases a simplified WKB approximation is used to increase computational efficiency [178]. In the case of tunneling between the charge reservoir to the defect, the potentials can be approximated linearly. For computing the WKB factor it is necessary to distinguish between *direct tunneling* and *Fowler-Nordheim tunneling* [179] depending on the shape of the potential barrier, as explained in more detail in [180].

An important parameter withing the entire model is the lineshape function that depends on the Frank-Condon factors and thus on the shape of the potentials of the defects. This potentials can be approximated as 1-dimensional [181, 182] harmonic, justified by the fact that in a Taylor expansion of the potentials the first order terms vanish and thus the second order term are dominant close to the equilibrium configuration [183]. Consequently, the potentials can be represented by harmonic potential energy curves (PECs) in the configuration coordinate (CC) diagrams, as can be seen in Fig. 3.2.

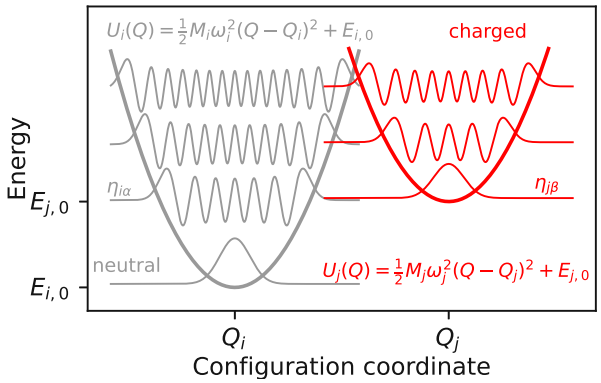

Figure 3.2. The harmonic potential energy curves can be used to represent the neutral and the charged defect configuration. The analytic eigenfunctions of the potentials are well known from the quantum harmonic oscillator.

The PECs are described by the parabolic potentials

$$\begin{aligned} U_i(Q) &= \frac{1}{2}M_i\omega_i^2(Q-Q_i)^2 + E_{i,0} \\ U_j(Q) &= \frac{1}{2}M_j\omega_j^2(Q-Q_j)^2 + E_{j,0} \end{aligned} \tag{3.22}$$

with ω_i, ω_j being the vibrational frequency of the harmonic potential, which effectively describes the curvature of the parabola, and M_i, M_j being the corresponding masses. The analytical eigenfunctions of the harmonic oscillator are well known from literature and are given by

$$\begin{aligned}\eta_{i\alpha}(Q) &= \frac{1}{2^\alpha \alpha!}\left(\frac{M_i\omega_i}{\pi\hbar}\right)^{1/4} \mathrm{e}^{-\frac{M_i\omega_i Q^2}{2\hbar}} H_\alpha\left(\sqrt{\frac{M_i\omega_i}{\hbar}}Q\right)\\ \eta_{j\beta}(Q) &= \frac{1}{2^\beta \beta!}\left(\frac{M_j\omega_j}{\pi\hbar}\right)^{1/4} \mathrm{e}^{-\frac{M_j\omega_j Q^2}{2\hbar}} H_\beta\left(\sqrt{\frac{M_j\omega_j}{\hbar}}Q\right)\end{aligned} \tag{3.23}$$

with H_α, H_β being the Hermite polynomials, defined by $H_\gamma(Q) = (-1)^\gamma \mathrm{e}^{\gamma^2}\frac{\mathrm{d}^\gamma}{\mathrm{d}Q^\gamma}\mathrm{e}^{-Q^2}$. Using the analytic solutions for the vibrational states allows the computation of f_{ij}^{LSF}, however, this computation is expensive, even when using optimized algorithms which determine the vibrational overlap based on a recurrence relation as proposed by Schmidt [184]. Thus, in Section 3.2.1 a WKB-based approximation for the lineshape function is derived, which is computationally superior. However, at room temperature and above the WKB-based approximation is not necessary, because the Frank-Condon factors, which are in essence the overlap of the vibrational wave functions of state i and j, are dominated by the overlaps close to the intersection point of the diabatic potentials U_i and U_j [185]. Note that the intersection point can be computed analytically. For this, typically the following parameterization of the potentials is used

$$\begin{aligned}E_{\mathrm{R}} &= \mathcal{S}_{ij}\hbar\omega_i = c_i\Delta Q^2,\\ R_{ij} &= \sqrt{\frac{c_i}{c_j}},\\ \Delta E_{ij} &= E_{i,0} - E_{j,0}.\end{aligned} \tag{3.24}$$

E_{R} is called relaxation energy and defined using the Huang-Rhys factor $\mathcal{S}_{ij}$ and is a measure for the electron-phonon coupling strength [186]. $c_i = \frac{1}{2}M_i\omega_i^2$ and $c_j = \frac{1}{2}M_j\omega_j^2$ are the curvatures of the harmonic potentials and define the curvature ratio R_{ij} while ΔE_{ij} describes the bias dependent energy difference of the minima of the two states, see Fig. 3.3.

With this parameterization the energy barrier is obtained as

$$\varepsilon_{ij} = \frac{E_{\mathrm{R}}}{(R_{ij}-1)^2}\left(1 - R_{ij}\sqrt{\frac{E_{\mathrm{R}} - \Delta E_{ij}(R_{ij}^2-1)}{E_{\mathrm{R}}}}\right) \tag{3.25}$$

which reduces to

$$\varepsilon_{ij} = \frac{(E_{\mathrm{R}} - \Delta E_{ij})^2}{4E_{\mathrm{R}}} \tag{3.26}$$

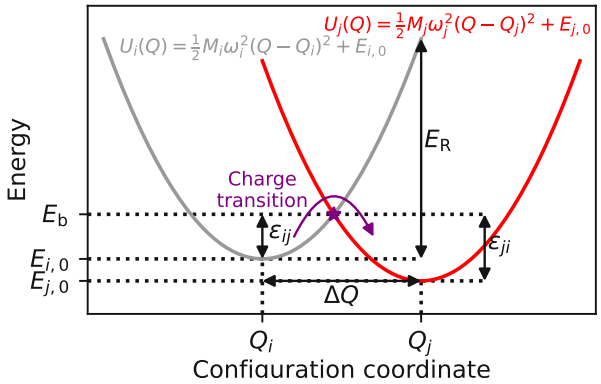

Figure 3.3. In the classical limit the transition rate can be approximated by using the classical barrier ε_{ij} instead of computing the overlaps of the eigenfunctions. The classical barrier can be computed analytically.

for $R_{ij} = 1$. Note that in this classical description, the parameter ΔQ becomes immaterial, since the lineshape function is only determined by the classical barrier height. Using this closed form expression for the intersection point of the potentials, the lineshape functions reduce to

$$\begin{aligned} f_{ij}^{\mathrm{LSF}}(T) &= \mathrm{e}^{-\beta\varepsilon_{ij}} \\ f_{ji}^{\mathrm{LSF}}(T) &= \mathrm{e}^{-\beta\varepsilon_{ji}} \end{aligned} \tag{3.27}$$

with $\varepsilon_{ij} = \varepsilon_{ji} - \Delta E_{ij}$. In the limit of high temperatures, the transition rates in (3.20) can then be simplified using the band edge approximation [180] to

$$\begin{aligned} k_{ij}^{\mathrm{CB}} &= n v_{\mathrm{th,n}} \sigma_{0,n} \theta_{\mathrm{n}} \mathrm{e}^{-\beta(\varepsilon_{ij}^{\mathrm{CB}} - E_{\mathrm{F}} + E_{\mathrm{CB}})} \\ k_{ji}^{\mathrm{CB}} &= n v_{\mathrm{th,n}} \sigma_{0,n} \theta_{\mathrm{n}} \mathrm{e}^{-\beta\varepsilon_{ji}^{\mathrm{CB}}} \\ k_{ij}^{\mathrm{VB}} &= p v_{\mathrm{th,p}} \sigma_{0,p} \theta_{\mathrm{p}} \mathrm{e}^{-\beta\varepsilon_{ij}^{\mathrm{VB}}} \\ k_{ji}^{\mathrm{VB}} &= p v_{\mathrm{th,p}} \sigma_{0,p} \theta_{\mathrm{p}} \mathrm{e}^{-\beta(\varepsilon_{ji}^{\mathrm{VB}} + E_{\mathrm{F}} - E_{\mathrm{VB}})} \end{aligned} \tag{3.28}$$

This closed form allows a numerically fast computation of transition rates.

In this simplified classical approximation of the 2-state NMP model it is necessary to distinguish between strong electron-phonon coupling (SEPC) and weak electron-phonon coupling (WEPC), see Fig. 3.4. SEPC is the dominating mechanism for BTI degradation and describes charge transitions where the classical intersection point of the initial and the final potential energy surface lies between the two minima of the PECs. All other cases are subsumed as WEPC. For WEPC, which typically plays a minor role for BTI degradation [180], classical barriers are given as shown in Table 3.1.

This derived classical approximation for the 2-state NMP model works well for room temperature and above, however, in the limit of low temperatures the simplified

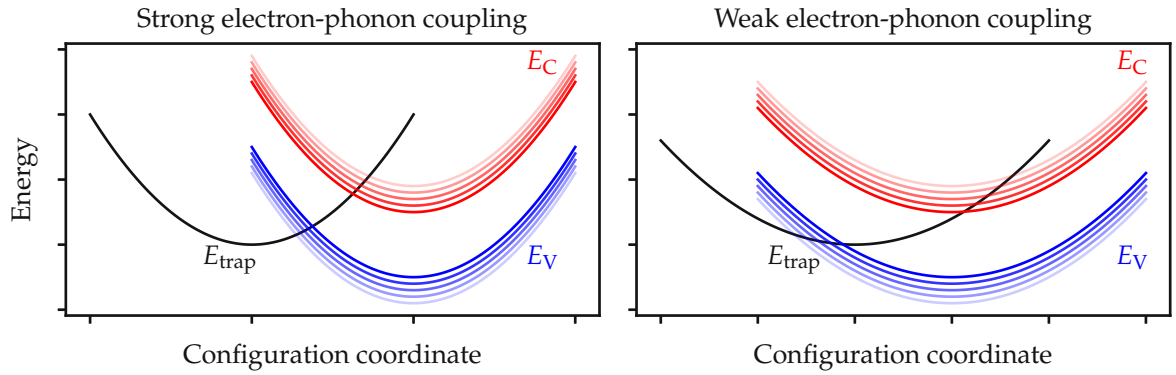

Figure 3.4. Depending on trap parameters and applied gate voltage, the charge transition takes place in the strong (left) or in the weak (right) electron-phonon coupling regime. In the SEPC regime, the classical intersection lies in between the minima of the two PECs, while in the WEPC regime it lies outside of the minima.

ΔE_{ij}	$\varepsilon_{ij}^{\mathrm{CB}}$	$\varepsilon_{ij}^{\mathrm{VB}}$	$\varepsilon_{ij}^{\mathrm{metal}}$
$\Delta E_{ij} > 0\,\mathrm{eV}$	ΔE_{ij}	Eq. (3.25)	ΔE_{ij}
$\Delta E_{ij} < 0\,\mathrm{eV}$	Eq. (3.25)	$0\,\mathrm{eV}$	$0\,\mathrm{eV}$

Table 3.1. In the case of weak electron-phonon coupling predefined barriers are used for taking the FD-distribution into account correctly [180].

lineshape functions (3.27) would freeze out completely

$$\lim_{T\to 0\,\mathrm{K}} f_{ij}^{\mathrm{LSF}}(T) = 0 \tag{3.29}$$

and there would be no charge trapping in cryogenic environments. However, measurements show that there are still electrically active traps at cryogenic temperatures, as will be discussed in detail in Part II of this work. The reason for this is that at low temperatures the assumption that the overlaps of the vibrational wavefunctions close to the intersection points of the diabatic potentials dominate does not hold anymore, because those states cannot be excited anymore at low temperatures. Thus, vibrational states at lower energies dominate the overlap and hence the lineshape function. This makes it necessary to include all vibrational states in the computation of the transition rate.

3.2.1 WKB Approximation

The computation of the full quantum mechanical charge transition rate (3.19) is inefficient and thus not suitable for calculating the defect model within device or circuit simulation tools. However, the classical approximation breaks down at cryogenic temperatures and can thus also not be used. Therefore, an approximation based on the Wentzel–Kramers–Brillouin (WKB) method combined with the saddlepoint method was

proposed by Holstein [187] and further developed by Markvart [188–190]. These approximations where developed for the weak electron-phonon coupling regime, however, for charge trapping most transitions happen within the strong electron-phonon coupling regime. To make this method suitable for reliability studies of MOS devices a more generalized form of the approach is developed. In the following the approximation of the full quantum mechanical lineshape function is based on the idea that the vibrational wavefunctions can be approximated by a WKB-based terms. This approximation allows switching from a summation over the exact discrete eigenenergies to a continuous integral form. The integrals can be evaluated using the saddlepoint method. This allows finding an effective barrier which is lower than the classical barrier and thus prevents the transition rates from freezing out. The full derivation is given in the following.

For deriving a general formulation for the WKB approximation a transition $\alpha \rightarrow \beta$ is singled out by introducing

$$\xi_{i\alpha,j\beta} = \mathrm{e}^{-\beta\Omega_{i\alpha}} |I_{i\alpha,j\beta}|^2 \delta(E_{i\alpha} - E_{j\beta}), \tag{3.30}$$

with the vibrational overlap integral $I_{i\alpha,j\beta}$ given by

$$I_{i\alpha,j\beta} = \Big| \int_{-\infty}^{\infty} \eta_{i\alpha}(Q)\eta_{j\beta}(Q)\mathrm{d}Q \Big|^2. \tag{3.31}$$

By approximating the PECs as harmonic, the vibrational wave functions can be expressed analytically by (3.23). The parabolic PECs for a charged and a neutral configuration are shown in Fig. 3.5, together with their eigenfunctions, and the vibrational wavefunctions which are shown with dashed lines. While this analytical expression is suitable for the lower eigenfunctions, the Hermite polynomials becomes very large for higher energy levels leading to numerical instabilities.

For solving the integration in (3.31) efficiently, a WKB-based approximation of the wavefunction $\eta_{i\alpha}$ introduced in (3.23) is used. The idea of the WKB approximation is a semi-classical series expansion of a corresponding action with respect to $\hbar$. In principle, this expansion could be carried out at any point, however, typically it is done at the classical turning point Q_i. This case is also shown in Fig. 3.5. The expansion can be classified in three regions, a classically allowed region between the turning points Q_i and Q_j, which is described by an oscillating function and the classically inaccessible regions which can be described by exponentially decaying functions (solid lines). The oscillating function cancels out approximately when integrated and thus can be neglected [190]. In a similar manner, the exponentially decaying function which is far away from the classical intersection point can be neglected. Therefore, the WKB wavefunction reduces

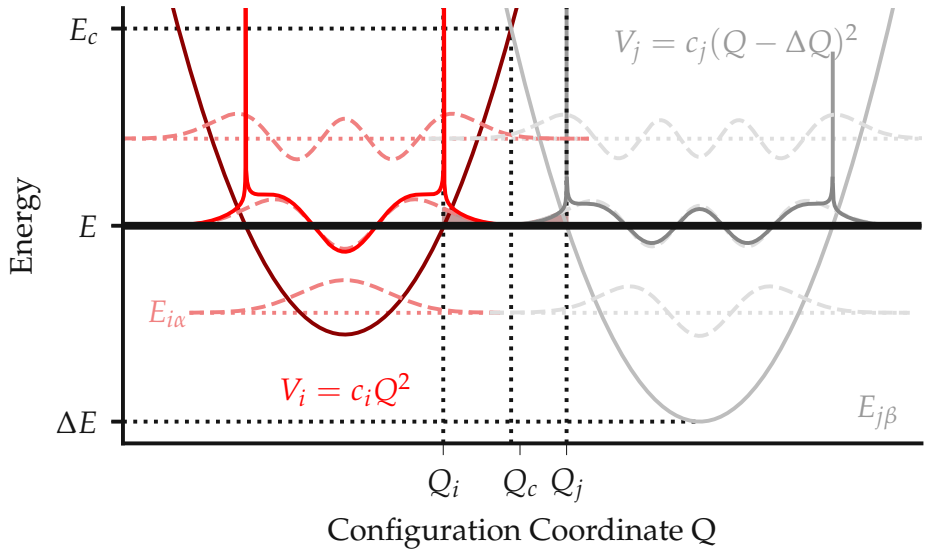

Figure 3.5. The vibrational wavefunctions of harmonic potential energy surfaces can be computed analytical. Since the analytical expressed wavefunctions (dashed lines) are rather complicated and unwieldy for simplifications, a WKB approximation (solid lines) is used which describes the wavefunctions in the classically forbidden regions well by capturing the exponential decay.

to a single exponentially decaying function in the classically prohibited region

$$\eta_{i\alpha}(Q) \approx \frac{(-1)^{\alpha} C_i}{\sqrt{|k_{i\alpha}(Q, E_{i\alpha})|}} \mathrm{e}^{\frac{1}{\hbar}\int_{Q_i}^{Q} k_{i\alpha}(Q', E_{i\alpha})\mathrm{d}Q'}, \tag{3.32}$$

with $k_{i\alpha}(Q, E_{i\alpha}) = \sqrt{2m(U_i(Q) - E_{i\alpha})}$ being the classical momentum. The full derivation of the WKB method is given in Appendix A.1. By normalizing $\eta_{i\alpha}(Q)$ the constant $C_i = (m_i\omega_i/8\pi^2)^{(1/4)}$ can be computed, with ω_i being the oscillator frequency [190]. The idea can also be seen in Fig. 3.5: The dashed lines show the analytical vibrational functions (3.23), while the solid line is divided in the three regions, two exponential decaying and a oscillating region. The integration is only done in the shaded regions below the classical intersection point, since all other terms cancel out approximately. Close to the classical turning points a singularity can be seen. Such singularities occur due to the Taylor expansion used for the WKB approximation, as explained in Appendix A.1. However, the overlap is dominated far away from the singularity, thus the singularity gets neglected in the integration when using the saddlepoint method, as explained in detail later in this section. For the calculation of the exact LSF given in (3.19), one does not have to explicitly distinguish between weak and strong electron-phonon coupling, hence the calculation is always the same regardless of the relative positions of the PEC. However, by reduction to the exponential decaying functions in the classically prohibited region, the sign of the decaying part needs to be considered. In the following,

the deviation is done for the strong-electron-phonon coupling case, all occurring cases are shown in Fig. 3.6 and are considered in the sign of the final result.

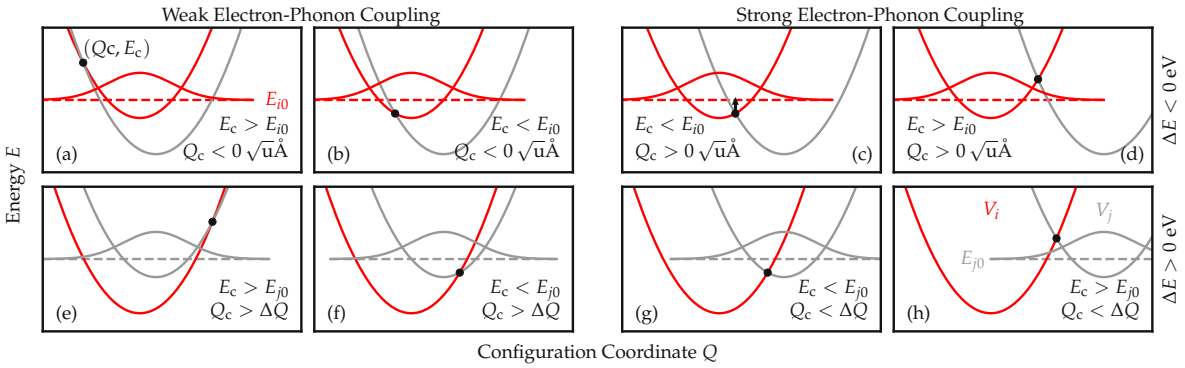

Figure 3.6. There are 8 different relative positions of the potential energy surfaces which need to be considered in the WKB based approximation of the transition rate: 4 positions in the weak electron-phonon coupling regime and 4 in the strong coupling regime. Furthermore, it is necessary to consider if the initial state is energetically above or below the final state and if the classical barrier E_c is above or below the ground state E_{j0}.

As previously mentioned, the overlap integral (3.31) can be simplified using the saddlepoint method [191]. For this, the overlap integral is first simplified using the WKB wavefunction (3.32) to get

$$I_{i\alpha,j\beta} \cong \left| \frac{C_i C_j}{\sqrt{|k_{i\alpha}(Q,E_{i\alpha})|}\sqrt{|k_{j\beta}(Q,E_{j\beta})|}} \mathrm{e}^{\frac{\varphi_{i\alpha j\beta}(Q,E_{i\alpha},E_{j\beta})}{2}} \right|^2, \tag{3.33}$$

where the phase $\varphi_{i\alpha j\beta}(Q,E_{i\alpha},E_{j\beta})$ is given by

$$\varphi_{i\alpha j\beta}(Q,E_{i\alpha},E_{j\beta}) = \frac{2}{\hbar}\int_{Q_i}^{Q} k_{i\alpha}(Q',E_{i\alpha})\mathrm{d}Q' - \frac{2}{\hbar}\int_{Q_j}^{Q} k_{j\beta}(Q',E_{j\beta})\mathrm{d}Q'. \tag{3.34}$$

Exploiting the fact that the integral is dominated by the region near the point of stationary phase, which is given by Q_c, which can be verified by solving for $\mathrm{d}\varphi_{i\alpha j\beta}(Q)/\mathrm{d}Q = 0$, the integral can be expressed as a function of Q_c only:

$$I_{i\alpha,j\beta}(E_{i\alpha},E_{j\beta}) \cong \left| \frac{C_i C_j}{\sqrt{|k_{i\alpha}(Q_c,E_{i\alpha})|}\sqrt{|k_{j\beta}(Q_c,E_{j\beta})|}} \times \mathrm{e}^{\frac{\varphi_{i\alpha j\beta}(Q_c,E_{i\alpha},E_{j\beta})}{2}} \sqrt{\frac{\pi}{\varphi''_{i\alpha j\beta}(Q_c,E_{i\alpha},E_{j\beta})}} \right|^2. \tag{3.35}$$

For this, the saddlepoint method is used, which is derived in Appendix A.2 and allows the simplification of a general integral I of the form

$$I = \int_{-\infty}^{\infty} \mathrm{e}^{-f(x)}\mathrm{d}x \cong \mathrm{e}^{-f(x_0)}\sqrt{\frac{2\pi}{f''(x_0)}} \tag{3.36}$$

when $f'(x_0) = 0$ and $f''(x_0) > 0$ hold. Note that the overlap integral is not exactly of this form, due to the Q dependence of the non-exponential prefactor. However, this can be neglected, since the integral is dominated by the exponential term. A further simplification requires to convert the quantized lineshape function containing α and β to a continuous form

$$\xi_{i,j}(E, E') = \mathrm{e}^{-\beta E_i}|I_{i,j}(E, E')|^2\delta(E - E') \tag{3.37}$$

by using $E_{i\alpha} \to E$ and $E_{j\beta} \to E'$. The same can be done for the summation of the eigenenergies

$$\sum_{\alpha=0}^{\infty} f(E_\alpha) \to \int_{E_0}^{\infty} f(E)\frac{\mathrm{d}n\alpha}{\mathrm{d}E}\mathrm{d}E. \tag{3.38}$$

Here, $\mathrm{d}\alpha/\mathrm{d}E$ is the density of states of the quantum mechanical harmonic oscillator

$$\frac{\mathrm{d}\alpha}{\mathrm{d}E} = \frac{\mathrm{d}}{\mathrm{d}E}\left(\frac{E}{\hbar\omega} - \frac{1}{2}\right) = \frac{1}{\hbar\omega}. \tag{3.39}$$

The continuous lineshape function (3.37) can be used to express equation (3.19) as

$$k_{ij}(T) = \frac{2\pi}{\hbar}\frac{|\theta_{ij}|^2}{Z}\iint \mathrm{e}^{-E/k_\mathrm{B}T}|I_{i,j}(E, E')|^2\delta(E - E')\frac{1}{\hbar\omega_i}\frac{1}{\hbar\omega_j}\mathrm{d}E\mathrm{d}E'. \tag{3.40}$$

Evaluating the Dirac delta distribution and using the simplified version of the overlap integral (3.35), the rate equation (3.40) simplifies to

$$k_{ij}(T) = \int C_2(E)\mathrm{e}^{-E/k_\mathrm{B}T+\varphi(E,Q_\mathrm{c})}\mathrm{d}E, \tag{3.41}$$

with $\varphi(E, Q_\mathrm{c})$ being the continuous form of $\varphi_{\alpha\beta}(Q_\mathrm{c})$. The term $C_2(E)$ contains all non-exponential terms depending on E

$$C_2(E) = \frac{2\pi}{\hbar}\frac{|\theta_{ij}|^2}{Z}\frac{1}{\hbar\omega_i}\frac{1}{\hbar\omega_j}\left|\frac{C_iC_j}{\sqrt{2m(E_\mathrm{c} - E)}}\right|^2\frac{2\pi}{\varphi''(Q_\mathrm{c})}. \tag{3.42}$$

By applying the saddlepoint method a second time, the integral (3.41) can be evaluated at the extremum E^*

$$\frac{\mathrm{d}\varphi(E)}{\mathrm{d}E}\Big|_{E=E^*} = \frac{1}{k_B T} \tag{3.43}$$

leading to

$$k_{ij}(T) = C_2(E^*)\mathrm{e}^{-E^*/k_\mathrm{B}T+\varphi(E^*)}\sqrt{\frac{2\pi}{\varphi''(E^*)}}. \tag{3.44}$$

After insertion of all terms the final result reads

$$\begin{aligned} k_{ij}(T) = \frac{2\pi}{\hbar}\frac{|\theta_{ij}|^2}{Z}\frac{1}{\hbar\omega_i}\frac{1}{\hbar\omega_j}\left|\frac{C_i C_j}{\sqrt{2m(E_\mathrm{c}-E^*)}}\right|^2 \times \\ \mathrm{e}^{-E^*/k_\mathrm{B}T+\varphi(E^*,Q_\mathrm{c})}\sqrt{\frac{2\pi}{\varphi''(E^*)}\frac{2\pi}{\mathrm{d}^2\varphi/\mathrm{d}Q^2(Q_\mathrm{c})}}. \end{aligned} \tag{3.45}$$

For obtaining the occurring phase $\varphi(E)$ in the final result it is necessary to compute the integral (3.34). This can be done analytically, resulting in the expression

$$\begin{aligned} \varphi_{i\alpha j\beta}(Q_\mathrm{c}, E_\alpha, E_\beta) = \frac{-\sqrt{2}}{\hbar}\left(r_1\left(\frac{\sqrt{E_\mathrm{c}}\sqrt{E_\mathrm{c}-E_\alpha}}{\sqrt{c_i}} - \frac{E_\alpha}{\sqrt{c_i}}\ln\frac{\sqrt{E_\mathrm{c}-E_\alpha}+r_2\sqrt{E_\mathrm{c}}}{r_2\sqrt{E_\alpha}}\right)\right. \\ \left. + s_1\left(\frac{\sqrt{E_\mathrm{c}-\Delta E}\sqrt{E_\mathrm{c}-E_\beta}}{\sqrt{c_j}} - \frac{E_\beta-\Delta E}{\sqrt{c_j}}\ln\frac{\sqrt{E_\mathrm{c}-E_\beta}+s_2\sqrt{E_\mathrm{c}-\Delta E}}{s_2\sqrt{E_\beta-\Delta E}}\right)\right). \end{aligned} \tag{3.46}$$

From this expression it is possible to get an analytic expression for the second derivative $\mathrm{d}^2\varphi/\mathrm{d}Q^2$ and $\varphi''(E)$. The factors r_1, r_2, s_1 and s_2 are necessary to differentiate between WEPC and SEPC. Furthermore it is necessary to differentiate between an initial state having a higher or a lower energy level than the final state and if the classical intersection point is above or below the ground state of the initial or final state. The different occurring cases are shown in Fig. 3.6 and the signs for the different cases are listed in Tab. 3.2.

	r_1	r_2	s_1	s_2
$Q_c < 0$	1	-1	-1	-1
$0 < Q_c < \Delta Q$	1	1	1	-1
$\Delta Q < Q_c$	-1	1	1	1

Table 3.2. Factors in analytic expression (3.46) of the overlap integral

3.2.2 Benchmark of the WKB-Based 2-State NPM Approximation

The derived WKB-based approximation of the 2-state model is developed for the limit of low temperatures. However, since characterization is often done in a broad temperature window (in this work specially from room temperature to 4 K) it is important to define the limits of the approximation and to check carefully if the models remain valid over the whole characterization range.

Lifetime Broadening

For benchmarking the WKB based approximation against the full quantum mechanical model it is necessary to introduce a lifetime broadening of the Dirac delta distribution $\delta(E_{i_\alpha} - E_{j\beta})$ occurring in the lineshape function (3.19) [192]. Without this broadening, the overlaps of the wave functions would only be nonzero for perfect energy level alignment $E_{i\alpha} = E_{j\beta}$. Therefore, the Dirac delta distribution is replaced by a Gaussian distribution

$$\delta(E_{i\alpha} - E_{j\beta}) = \frac{1}{\sqrt{2\pi}\sigma} \mathrm{e}^{-\frac{(E_{i\alpha}-E_{j\beta})^2}{2\sigma^2}} \tag{3.47}$$

with σ being the width of the broadening. The choice of σ is crucial, however, not trivial. Since the broadening is a thermal effect, σ should show a temperature dependence, thus the simplest form of a linear dependence $\sigma = k_\mathrm{B}T$ was chosen. At cryogenic temperatures this broadening would disappear, however, excited states still show a finite lifetime due to phonon emissions, and thus must be somehow affected by lifetime broadening. Therefor, a sigma of the form

$$\sigma = \max(\sigma_c \hbar\omega_i, k_\mathrm{B}T) \tag{3.48}$$

was chosen, where σ_c is a constant normalized to the level energy spacing $\hbar\omega_i$ which has the role of representing temperature independent lifetime broadening due to phonon emissions. A variation of this constant is shown in Fig. 3.7 for σ_c between $1/5$ and 3. Different values for σ_c lead to different saturation values of the lineshape function towards cryogenic temperatures. The impact of σ_c gets smaller in the classical limit where the lineshape functions approach the classical model for all chosen σ_c. As can be seen, σ_c shows a saturation towards smaller values. The difference between $\sigma_c = 1/5, 1/3$ and 1 is quite small compared to the larger σ_c values. In this work $\sigma_c = 1/3$ was chosen, to guarantee that there is always at least one overlap of vibrational functions in the range of 3σ.

Temperature Dependence

The main motivation for using the full quantum mechanical model or the WKB based approximation in favor of the much simpler classical approximation is to ensure

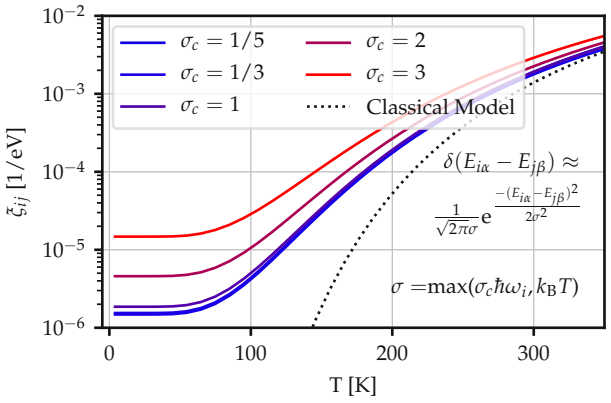

Figure 3.7. The benchmarking of the WKB based approximation against the full quantum mechanical model makes it necessary to smear out the Dirac delta distribution $\delta(E_{i_\alpha} - E_{j\beta})$ in the lineshape function (3.19) to a Gaussian function. The impact of different smearing widths can be seen for different minimal widths σ_c between 1/5 and 3.

the validity of the computed transition rates at cryogenic temperatures. Therefore, the benchmarking of the WKB based approximation against the full quantum mechanical model is specially relevant for temperatures below room temperature.

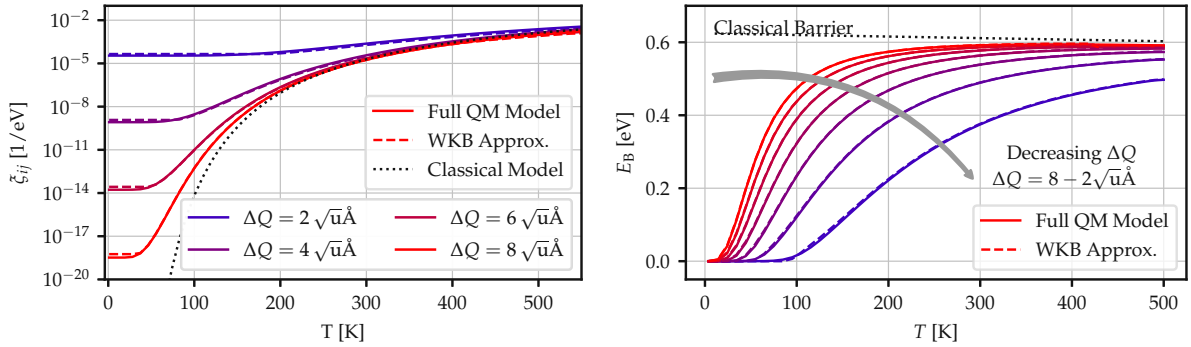

Figure 3.8. The full quantum mechanical lineshape function (solid) and the WKB based approximation (dashed) for $R_{ij} = 1$, $E_R = 2.5\,\text{eV}$ and $\Delta E_{ij} = 0\,\text{eV}$ do not freeze out for different configuration coordinate displacements ΔQ, while the ΔQ independent classical approximation (dotted) freezes out completely towards 4 K (left). This can be mapped to an effective barrier lowering, which shows that the effective barrier is 0 eV at cryogenic temperatures, while the classical effective barrier is almost temperature independent (right).

The lineshape function can be seen for a variation of temperatures between 4 K and 500 K in Fig. 3.8 for different configuration coordinate displacements ΔQ between $2\sqrt{\text{u}}\text{Å}$ and $8\sqrt{\text{u}}\text{Å}$. The full quantum mechanical model (solid) and the WKB based approximation are almost perfectly on top of each other. Depending on the displacement ΔQ the lineshape function starts to get constant below a certain temperature, while the classical approximation of the lineshape function, shown as black dotted line, is ΔQ independent and freezes out completely towards cryogenic temperatures. Using the

computed lineshape function, it is possible to plot the effective barrier

$$E_{\text{eff}} = -k_{\text{B}} \frac{\partial \ln(\xi(T))}{\partial(1/T)}. \tag{3.49}$$

which can be derived directly from the Arrhenius law [185]. The effective barrier decreases to 0 eV at low temperatures as can be seen in Fig. 3.8 (right). Towards high temperatures the effective barrier approaches the classical barrier. However, at room temperature there is still a significant difference between the classical barrier and the effective barrier for small ΔQ, which means that the classical approximation underestimates charge transition rates of defects with small configuration coordinate displacements by neglecting nuclear tunneling. The classical barrier shows only a small temperature dependence arising from the temperature dependent prefactor in the classical lineshape function.

The lineshape function can be seen for a variation of the energetic displacements in Fig. 3.9. For low temperatures, the classical lineshape function gives reasonable values only in a very small displacement window, in which the classical barrier is close to 0 eV. The full quantum mechanical lineshape function and the WKB based approximation on the other hand give comparable large values over a wide range of energetic offsets, which explains the observed freeze-out of transition rates in the classical approximation.

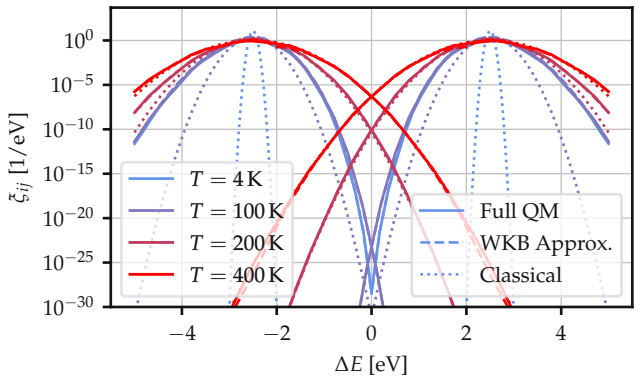

Figure 3.9. The forward and reverse lineshape functions ξ_{ij} and ξ_{ji} at $T = 4\,\text{K}$ (blue), $T = 100\,\text{K}$ (violet), $T = 200\,\text{K}$ (magenta) and $T = 400\,\text{K}$ (red) for different energetic displacements ΔE_{ij} show that the classical approximation (dotted) gives large values only in a very narrow interval where the classical barrier is close to 0 eV, while the full quantum mechanical version (solid) and the WKB based approximation (dashed) have large values for a comparable wide interval.

Parameter Scan

For a systematic comparison of the WKB based approximation and the full quantum mechanical lineshape function the parameters E_{R}, R_{ij}, ΔQ, ΔE_{ij}, and T are being varied. The variation of these parameters is done using the grid given in Tab. 3.3

Parameter space			
Parameter	min.	max.	gridpoints
E_R [eV]	0.5	6	10
ΔE [eV]	-3	3	10
ΔQ [$\sqrt{\text{u}}$Å]	1	6	10
R_{ij} [1]	0.7	1.3	3
T [K]	4	500	10

Table 3.3. For benchmarking the WKB-based approximation against the full quantum mechanical transition rate of the 2-state NMP model the parameters defining the potential energy curves, and thus the transition rates, are varied in the parameter space given above.

The lineshape function is computed for every combination of grid point parameters. Since the lineshape function varies over many orders of magnitude, the logarithmic lineshape function is used for defining the relative error made by the WKB-based approximation compared to the full quantum mechanical computation

$$\delta_{\text{rel}} = \frac{\ln(\xi_{\text{WKB}}) - \ln(\xi_{\text{FQM}})}{\ln(\xi_{\text{FQM}})}. \tag{3.50}$$

For the parameter variation shown in Tab. 3.3 more than 97% of the parameter combinations result in a δ_{rel} in the interval $[-1, 1]$, which corresponds to a difference of less than one order of magnitude (for a quantity which varies over tens of orders of magnitudes). This is illustrated in Fig. 3.10. Note that larger differences between the WKB-based approximations and the full quantum mechanical 2-state model arise for temperatures above 300 K. Here the overlap integral is not dominated by the exponential tail and therefore the WKB-approximation leads to larger deviations.

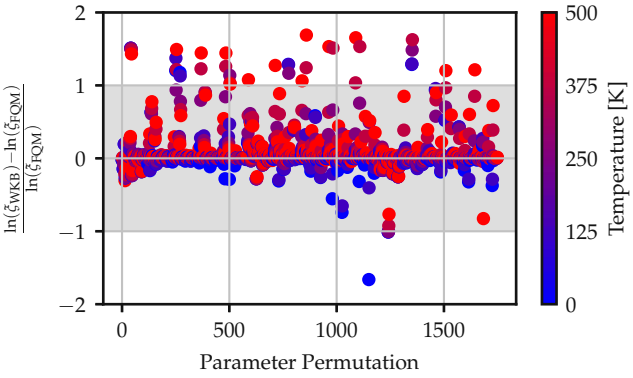

Figure 3.10. More than 97% of the parameters listed in 3.3 result in a relative error of the logarithmic lineshape function in the interval $[-1, 1]$. This corresponds to a difference of one order of magnitude which is considered acceptable in the context of transition rates which typically vary over many decades.

Computational Costs

As shown in the previous section, the WKB based approximation gives very accurate results for the lineshape function over many orders of magnitude and a wide range of parameters. Here the focus is set on analyzing the computational performance of the WKB approximation, which is important in the context of compact modeling. The computation time of the full quantum mechanical model depends on the number of vibrational states which need to be considered, see Fig. 3.11. Large relaxation energies E_{R} and large energetic offsets ΔE_{ij} lead to a higher number of vibrational states contributing to the lineshape function. Thus, with an increasing ΔE_{ij} the number of not negligible elements in the overlap matrix grows, scaling with $O(n^2)$ where n is the index of the highest included vibrational state. Since the computation of the overlap matrix is the bottleneck of the full quantum mechanical model, the overall runtime of the exact model also scales with $O(n^2)$, as is clearly shown in Fig. 3.11. The bottleneck of the WKB based approximation is to minimize (3.43) for finding $E^\star$. Finding $E^\star$ can be done by using e.g. Newton's method or other root-finding algorithm. Here, the energetic offset has no impact on the root-finding algorithm as can be seen in the Fig. 3.11.

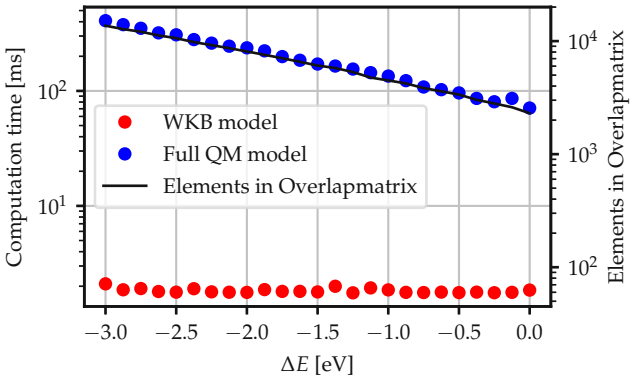

Figure 3.11. The number of vibrational states contributing to the lineshape function increases with increasing ΔE_{ij} and with it the number of elements in the overlap matrix (black line). Since the computation of the overlap matrix is the bottleneck of the full QM lineshape function (blue), the computation time increases with the same slope as the elements in the overlap matrix. The numerical bottleneck of the WKB bases model (red) is finding the energy E^*, which is independent of the number of vibrational functions.

With this numerical advantage of the WKB based model over the full quantum mechanical 2-state NMP model defect sampling of thousands of defects is possible. However, the WKB based model is not a closed form expression in the form of an equation set, because a root-finding algorithm needs to be applied for finding the effective energy E^*. By assuming that the PECs of the two states have identical curvatures, the WKB-based approximation can be further simplified to a real closed form expression, as discussed in the next section.

3.2.3 Approximation for Undistorted Potential Energy Curves

The WKB-based approximation for the 2-state NMP model is very accurate and numerically superior to the full quantum mechanical summation. However, since finding the effective barrier E^* in the WKB-based approximation requires a Newton solver, it is not a closed form solution and is thus numerically expensive compared to the classical approximation. To further improve the computational performance, the number of free parameters in the defect sampling can be reduced by assuming that the deformation of PECs during charge trapping can be neglected, which is done by fixing the curvature ratio of the PECs to $R = 1$. This is a reasonable approximation and known from the classical model since R and E_R are correlated in the 2-state NMP model, allowing to fix $R = 1$ by adapting E_R. By developing a Taylor expansion of the classical barrier ε_{12} with respect to the energetic offset ΔE_{12} [MJJ2]

$$\varepsilon_{12} = \frac{E_\mathrm{R}}{(1+R)^2} + \frac{R}{1+R}\Delta E_{12} + \mathcal{O}(\Delta E_{12}^2) \tag{3.51}$$

it can be seen that for a wide range of R values a corresponding E_R can be found resulting in the same ε_{12}, see Fig. 3.12. This allows finding (R, E_R) pairs which are experimentally almost indistinguishable. While in previous studies large curvature ratios like $R = 2.59$ [136] have been used, this correlation allows using smaller ratios in the value range of $R = 0.8 - 1.2$, which is the typical range for oxide defects as extracted from DFT calculations [185, MJJ2].

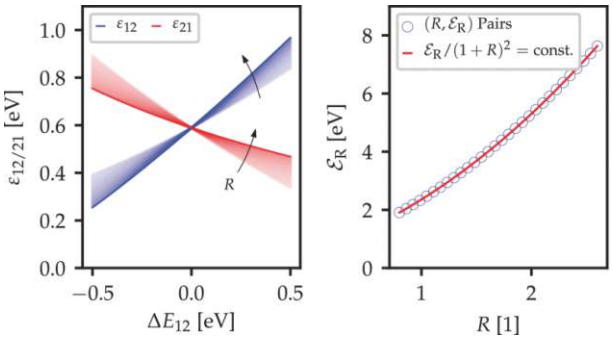

Figure 3.12. The charge transition barriers $\varepsilon_{12/21}$ are experimentally very similar for a large set of (R, E_R) pairs along the curve $E_\mathrm{R}/(1+R)^2 = \text{const}$. This can be derived using a Taylor expansion of the classical transition barrier (3.51) and allows to fix either R or E_R.

By fixing $R = 1$, the WKB-based approximation can then be considerably simplified [193]. The lineshape function ξ can be computed by

$$\xi = \omega\frac{\sqrt{2\pi}}{2D\hbar}\left((\coth\gamma + 1)\mathrm{e}^{-E_\mathrm{l}/2D^2\hbar^2} + (\coth\gamma - 1)\mathrm{e}^{-E_\mathrm{h}/2D^2\hbar^2}\right). \tag{3.52}$$

with

$$D = \sqrt{\frac{\omega^2}{2}\Delta^2(2n+1)} \tag{3.53}$$

and the dominant mode $n = 1/(\exp(2\gamma) - 1)$ where

$$\gamma = \frac{\hbar\omega}{2k_\mathrm{B}T} \tag{3.54}$$

and

$$\Delta = \sqrt{\frac{2E_\mathrm{R}}{\hbar\omega}}. \tag{3.55}$$

The oscillation frequency ω can be calculated directly from the relaxation energy E_R for a given configuration coordinate offset ΔQ

$$\omega = \frac{\sqrt{2E_\mathrm{R}}}{\Delta Q}. \tag{3.56}$$

Since $R = 1$, $\omega_i = \omega_j$ holds and is referred to as ω. The energy levels E_l and E_h in (3.52) are given by

$$\begin{aligned} E_\mathrm{l} &= \Delta E - \hbar\omega - E_\mathrm{R} \\ E_\mathrm{h} &= \Delta E + \hbar\omega - E_\mathrm{R}. \end{aligned} \tag{3.57}$$

This set of equations allows a direct computation of the transition rate in the strong electron-phonon coupling for the case that the final state is lower than the initial state. The backward rate (which can not be computed with this set of equations, because here the final state would be energetically higher than the initial state) can be computed directly from detailed balance as it is done in classical 2-state NMP theory [180]. For the case of weak electron-phonon coupling, the effective barrier can be replaced by the fixed values derived in [136, 180].

The lineshape function can also be used for computing the effective barrier E_eff by

$$E_\mathrm{eff} = -\log(\xi)/\beta(T). \tag{3.58}$$

The closed form expression for the lineshape function for undistorted potential energy curves matches well with the full quantum mechanical 2-state model as can be seen in Fig. 3.13 for $E_\mathrm{R} = 4\,\mathrm{eV}$ and $\Delta E = 0\,\mathrm{eV}$ for various configuration coordinate offsets.

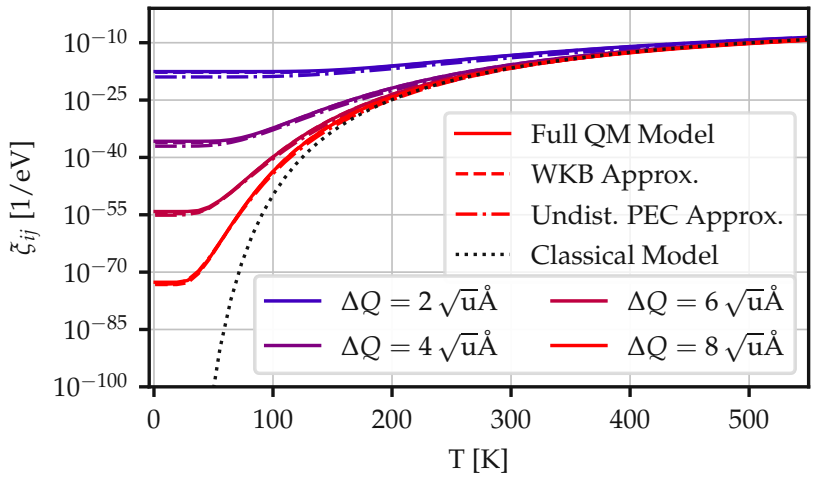

Figure 3.13. The lineshape function for undistorted potential energy curves (dashdoted) matches well with the full quantum mechanical model (solid) and the WKB-based approximation (dashed) across various temperatures and configuration coordinate offsets for $E_R = 4\,\text{eV}$ and $\Delta E = 0\,\text{eV}$.

3.3 4-State Nonradiative Multiphonon Model

The 2-state NMP theory is very efficient for describing BTI on large-area devices, where large ensembles of defects are simultaneously active. However, when looking at the single defect level (RTN, TDDS), there are certain phenomena which can not be explained with 2 states. A prime example for failing of the 2-state model is the observation of anomalous RTN (aRTN), in which the RTN signal is interrupted for longer periods of time [118]. This behavior can only be explained by introducing an additional third state [194]. As can be seen in Fig. 3.14, a defect captures and emits a charge with a high frequency transitioning between state 1 and 2. After a certain time, a transition between state 2 and 2′ happens and the defect stays for a comparably long period in state 2′.

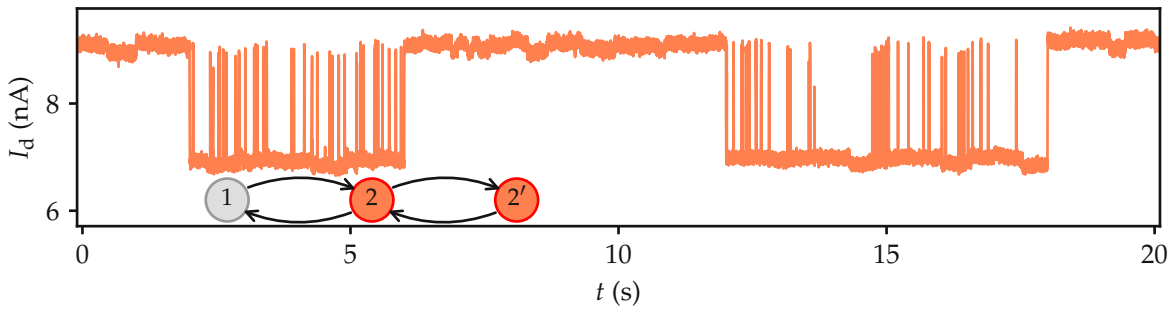

Figure 3.14. Anomalous RTN can be explained by introducing an additional third state.

This observation combined with the knowledge from DFT simulations (see Section 2) motivates the introduction of the 4-state NMP model. Additionally to the two stable states 1 and 2 which occur in 2-state NMP theory, two additional meta-stable states 1′

and $2'$ are introduced. A direct transition from one stable state to the other stable state is only possible via one of the two meta stable states $1 \rightleftharpoons 1' \rightleftharpoons 2$ or $1 \rightleftharpoons 2' \rightleftharpoons 2$, see Fig. 3.15. Analogously to the stable states, one-meta stable state is neutral while the second meta-stable state is charged. The transition between a neutral and a charged state, which would be the transitions $1' \leftrightharpoons 2$ and $1 \leftrightharpoons 2'$ in Fig. 3.15 are modeled using NMP transition rates, while transitions between a state and the meta-stable state with the same charge $1 \leftrightharpoons 1'$ and $2 \leftrightharpoons 2'$ are thermal transitions which are described within classical transition state theory leading to transition rates of the form $k_{11'} = \nu_0 e^{-\beta\varepsilon_{11'}}$ where $\varepsilon_{11'}$ is a constant, i.e. gate voltage independent barrier.

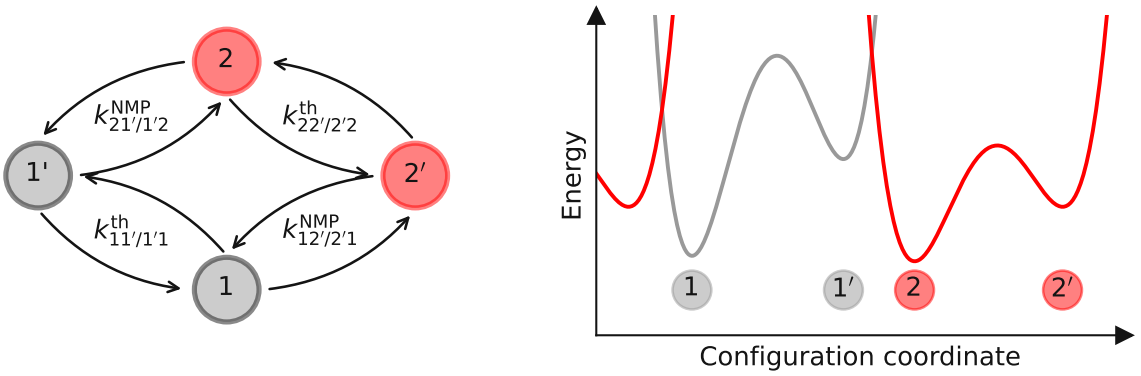

Figure 3.15. The state diagram for the 4-state NMP model shows two stable states 1 and 2 and two meta-stable states $1'$ and $2'$. A charge transition between the stable states is only possible via the meta-stable states. A transition between a two states with the same charge ($1 \leftrightharpoons 1'$ and $2 \leftrightharpoons 2'$) corresponds to a structural relaxation while a transition with a charge exchange ($1 \leftrightharpoons 2'$ and $2 \leftrightharpoons 1'$) can be modeled as NMP transition. The corresponding configuration coordinate diagram (right) shows the potential energy surfaces for the 4-state model.

Using the more complex potential energy surfaces shown in Fig. 3.15 allows to express the four NMP and the four thermal transition rates as shown in [174]. This rates allow obtaining the overall capture and emission time from one stable state to the other. Using Markov theory, the probability p_i of a defect to be in state i can be derived enabling reliability simulations.

Chapter 4

Reliability Modeling

First models for MOS transistors to calculate the potential across the devices were developed more than 50 years ago [195, 196]. In 1966 the Pao-Sah model using a double integral formulation for solving the Poisson equation [197] was published. However, this model was not feasible for integration in circuit simulations, because it was numerically inefficient [198]. Therefore, later in 1978 the Pao-Sah model was simplified by Brews employing the charge-sheet approximation [199] forming the basis for threshold-voltage based models like BSIMv3 or MM9 [200]. For several years threshold-voltage based models were mainly used in device simulations, however, due to discontinuities and inaccuracies in inversion, surface-potential based models like PSP [201] or MM11 and charge-based models like EKV [202], ACM [203] or BSIM-Bulk [204] became dominant.

The simulation framework used in this work is called *Comphy* (short for *compact physics*) [136] and is based on surface potential computation, such as PSP and MM11. *Comphy* has been developed for calculating the threshold voltage V_{th} changes for transient input parameters, that are the gate voltages V_{G} and temperatures T. This is done by defining a gate stack by its geometry and fundamental material parameters and by using the electrostatics derived for a one-dimensional device. By sampling defects in the oxide, degradation curves, e.g. the drift of the threshold voltage ΔV_{th} extracted from BTI measurements, can be simulated using a 2-state NMP model as discussed in Chapter 3.

4.1 Electrostatics of MOS Structures

By combining the one-dimensional Poisson equation with the charge neutrality equation, a surface potential φ_{S} based expression for the space charge per unit area can be derived [205]

$$Q_{\mathrm{S}} = \pm \frac{\sqrt{2} k_{\mathrm{B}} T}{q L_{\mathrm{D}}} \left(\left(\mathrm{e}^{-\beta \varphi_{\mathrm{S}}} + \beta \varphi_{\mathrm{S}} - 1 \right) + \frac{n_0}{p_0} \left(\mathrm{e}^{\beta \varphi_{\mathrm{S}}} - \beta \varphi_{\mathrm{S}} - 1 \right) \right)^{1/2}. \tag{4.1}$$

Here, the positive sign refers to $\varphi_{\mathrm{S}} > 0$ while the negative sign is used for $\varphi_{\mathrm{S}} < 0$. The abbreviation $L_{\mathrm{D}} = \sqrt{k_{\mathrm{B}} T \varepsilon_0 \varepsilon_{\mathrm{r,chan}} / p_0 q^2}$ is the Debye length for holes with $\varepsilon_{\mathrm{r,chan}}$

being the relative permitivity of the channel. n_0 and p_0 are the electron and the hole concentration in thermal equilibrium which can be efficiently computed using the Joyce-Dixon approximation [206]. q is the elementary charge, and $\varepsilon_{\mathrm{r,chan}}$ is the permittivity of the channel. For computing the temperature dependence of the Si band gap the model proposed by Bludau [207] is used. For effective masses for the valence band and the conduction band the models proposed by Lang *et. al.* [208] and Green [209] are used, respectively. With this, effective densities of states, Fermi level, intrinsic Fermi level, carrier concentrations, etc. can be computed as shown in [205].

By using simple electrostatic considerations [180] the following equation can be derived

$$\frac{Q_{\mathrm{S}}(\varphi_{\mathrm{S}})}{C'_{\mathrm{ox}}} + \varphi_{\mathrm{S}} - V_{\mathrm{G}} + \Delta E_{\mathrm{W},0} + E_{\mathrm{F},0} = 0, \tag{4.2}$$

where $\Delta E_{\mathrm{W},0}$ is the intrinsic work function difference and $E_{\mathrm{F},0}$ is the Fermi level in the channel at thermal equilibrium. The oxide capacitance per area is given by

$$C'_{\mathrm{ox}} = \frac{\varepsilon_0 \varepsilon_{\mathrm{r,ox}} W L}{d_{\mathrm{ox}}}, \tag{4.3}$$

where $\varepsilon_{\mathrm{r,ox}}$ is the relative permittivity of the oxide, W and L are channel width and length, and d_{ox} is the thickness of the oxide. Equation (4.2) can be solved by using an iterative scheme, e.g. a Newton solver to obtain φ_{S}. By considering the trapped charges in the oxide, (4.2) can be extended to (4.4) which then allows to be minimized for φ_{S} with

$$\frac{Q_{\mathrm{S}}(\varphi_{\mathrm{S}})}{C'_{\mathrm{ox}}} + \varphi_{\mathrm{S}} - V_{\mathrm{G}} + \Delta E_{\mathrm{W},0} + E_{\mathrm{F},0} + V_{\mathrm{traps}} = 0, \tag{4.4}$$

where V_{traps} stands for the voltage shift caused by the oxide charges. The calculation of φ_{S} is the central operation in surface potential based compact models and allows to compute the potential curve across the insulator which is approximated as being defect free. This enables computing the shift of trap levels and then the computation of charge capture and emission rates using 2-state NMP theory, which is discussed in detail in Section 3.2. Depending on the defect occupation and the defect density, a mean captured charge can be computed at every time step for an ensemble of defects. The sum of all charge states then leads to a shift of the threshold voltage. This finally results in a transient ΔV_{th} characteristics which can be compared to experimental data.

4.2 Defect Distribution in the Oxide

Using the electrostatics described in the previous section, the defect population can be sampled in the oxide in a subsequent step. Every defect has a certain probability of capturing and emitting a charge, which is determined by the transition rates $k_{ij}^{\mathrm{CB/VB}}$ and

$k_{ji}^{\mathrm{CB/VB}}$ in (3.20) from 2-state NMP theory. Using these transition rates it is possible to compute the occupancy of each trap according to the Master equations

$$\begin{aligned}\frac{\mathrm{d}p_i(t)}{\mathrm{d}t} &= -p_i \sum_{r=1}^{N_r} k_{ij,r} + p_j \sum_{r=1}^{N_r} k_{ji,r} \\ \frac{\mathrm{d}p_j(t)}{\mathrm{d}t} &= p_i \sum_{r=1}^{N_r} k_{ij,r} - p_j \sum_{r=1}^{N_r} k_{ji,r},\end{aligned} \tag{4.5}$$

where $p_{i,j}$ are the probabilities that a defect is in the initial/final state. These equations include the interaction of each defect with all charge reservoirs r, which are in the case of a MOSFET the valence and conduction band of the channel and the gate. The solution of the differential equation system is

$$p_j(t) = \frac{k_{ij}}{k_{ij}+k_{ji}} + \left(p_j(0) - \frac{k_{ij}}{k_{ij}+k_{ji}}\right)\mathrm{e}^{-t(k_{ij}+k_{ji})} \tag{4.6}$$

$$p_i(t) = 1 - p_j(t) \tag{4.7}$$

and can be used to weigh the impact on the potential caused by the defects by

$$V_{\mathrm{traps}} = -\varepsilon_0\varepsilon_{\mathrm{r}} WL \sum_{n=1}^{N_{\mathrm{T}}} q_{\mathrm{T},n} d_{\mathrm{ox}}\left(1 - \frac{x_{\mathrm{T}}}{d_{\mathrm{ox}}}\right) \tag{4.8}$$

with

$$q_{\mathrm{T}}(t) = q p_j(t) \tag{4.9}$$

being the charge stored at a defect. Subsequently, the surface potential of the device can be computed using (4.2). The shift of the potential compared to the initial time step then gives the threshold voltage shift. The total threshold voltage shift thus depends on the probability for a charge transition which itself depends on the defect parameters, the applied gate voltage and the temperature as shown in Fig. 4.1. Note that in (4.8) the contribution of a defect to the total V_{traps} depends on its depth and is considered according to the charge sheet approximation (CSA). However, recent studies have shown that the CSA considerably underestimates the impact of a single defect on the device performance, which might give rise to a slight overestimation of the trap densities in the current version of *Comphy* [210, MJJ3, MJC1, MJC2].

Specified Defect Distributions

Each defect can be described by a set of model parameters: the relaxation energy E_{R}, the curvature ratio R, the trap level E_{T}, the spatial position inside the oxide x_{T}, and the configuration coordinate offset ΔQ. The latter is only relevant when using a quantum mechanical 2-state NMP transition rate. Typically, in large area devices an

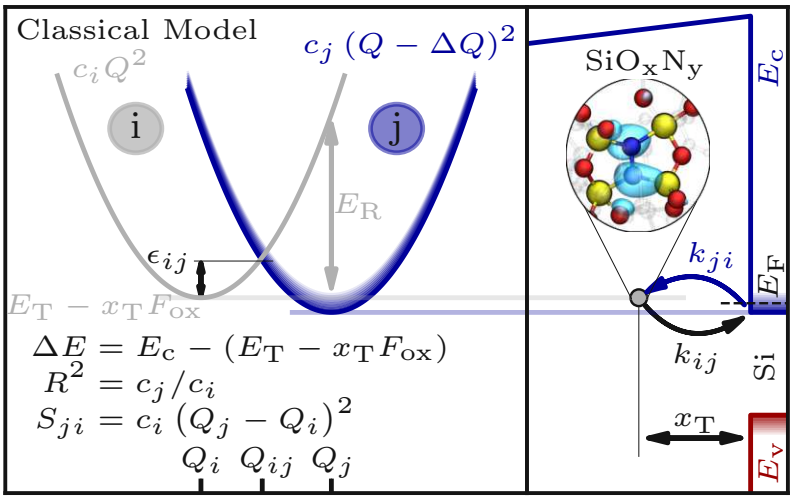

Figure 4.1. The mean charge state of a pre-existing oxide defect depends on the defect parameters trap level E_T, relaxation energy E_R, curvature ratio R and position in the oxide x_T, on the applied gate voltage and the temperature. With these parameters it is possible to compute transition rates k_{ij} and k_{ji} using NMP theory which gives then the impact on the electrostatics using the charge sheet approximation.

ensemble of defects is assumed to be distributed across a so called *defect band*. For this, it has often been assumed that the trap level of defects is normally distributed, that relaxation energies are normally distributed as well and that the spacial position is uniformly distributed [136]. Since the curvature of the PECs is correlated with the relaxation energy, it is not necessary to treat R as a stochastically distributed unit. Finally, a uniform spatial defect concentration N_T across parts of the dielectric is assumed. The parameters can be seen in Fig. 4.2. The sampling of the defects can either be done by Monte-Carlo sampling or more efficiently, the mean degradation of the defect band can be computed by sampling the parameters on a grid and introducing a weighting scheme.

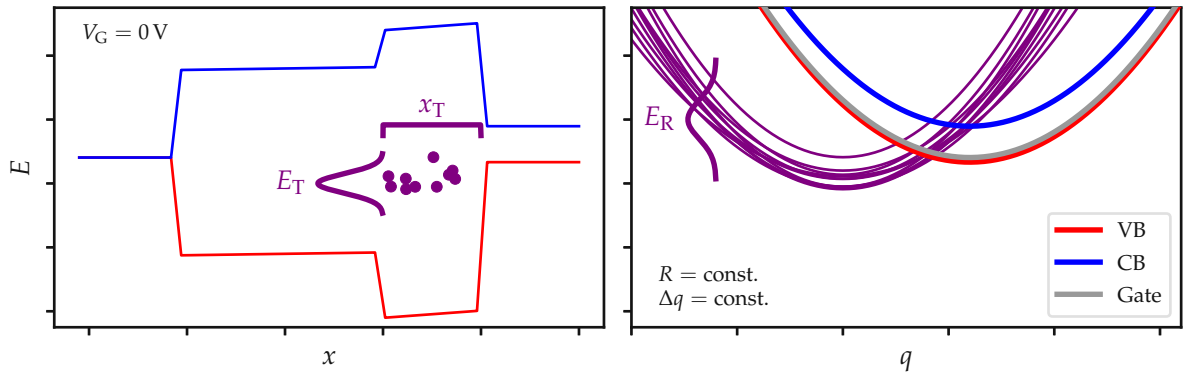

Figure 4.2. A defect is described by the relaxation energy E_R, the trap level E_T, the spatial position in the oxide x_T, the curvature ratio R, and the configuration coordinate offset ΔQ. While the latter two are correlated with the energies, x_T is assumed to be uniformly distributed and the energies are normally distributed in defect bands. With the specified distributions, defects can then be sampled in the band using Monte-Carlo sampling.

Effective Single Defect Decomposition

While the approach of using Gaussian trap bands has proven to be very successful [136], it has the disadvantage that large relaxation energies, compared to values obtained from DFT calculations of suitable oxide defect candidates [185], are required eventually [MJJ2]. This can give rise to unphysically large charge transition times, as exemplary shown for the shallow SiO_2 trapband of the 28 nm technology examined in [136] in Fig. 4.3, where the majority of all sampled defects shows transition times above 10^7 s, which is experimentally inaccessible. The smaller time constants correspond to the tail of the distribution of the relaxation energies. Thus, an alternative approach using an effective single defect decomposition (ESiD) has been developed to avoid artificially high relaxation energies. In this approach no assumptions on the defect parameter distributions are made when the model is calibrated to experimental data [MJJ2].

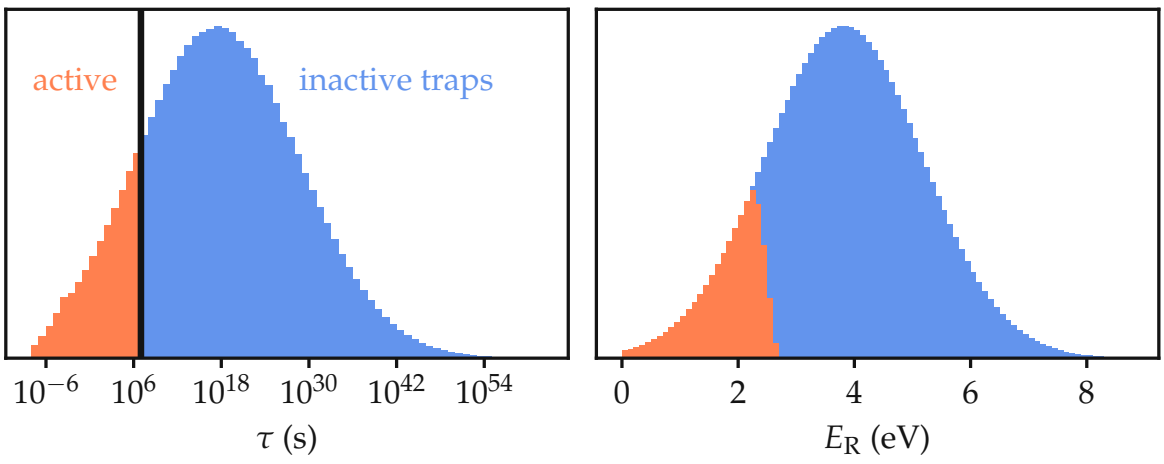

Figure 4.3. Gaussian trap bands as used in the 28 nm technology published in [136] show that only a small part of the distribution results in active defects. The majority of the extracted defect parameters of the trap band leads to charge transistion times above 10^7 s (left), and is thus experimentally inaccessible. These large charge transition times correspond to relaxation energies E_R larger than compared to those derived from DFT calculations (right) and are thus an artifact of the Gaussian distributions. Figure recreated from [MJJ2].

When using the ESiD approach, the total threshold voltage shift ΔV_{th} is expressed as superposition of the threshold voltage shifts caused by each defect δV_{th}

$$\Delta V_{th} = \sum_{E_T, E_R, x_T, \Delta Q} N(t; E_T, E_R, x_T, \Delta Q) \delta V_{th}(E_T, E_R, x_T, \Delta Q) \tag{4.10}$$

where $N(E_T, E_R, x_T, \Delta Q)$ is the weight of each defect. Instead of assuming Gaussian distributions for E_T and E_R, the ESiD algorithm employs a uniform parameter grid sampled for E_T, E_R, x_T and ΔQ. These parameters span a grid that is considered in the response matrix $(\boldsymbol{\delta V})$, where at every point i the response δV_{th} is computed. By

using the response matrix and defining the observation vector $(\boldsymbol{\Delta V})_{\boldsymbol{j}} = \Delta V_{\mathrm{th}}(t_j)$ a non-negative linear least square (NNLS) algorithm can be used for computing the weights $(\boldsymbol{N})_{\boldsymbol{i}} = N(E_{\mathrm{T},i}, E_{\mathrm{R},i}, x_{\mathrm{T},i}, \Delta Q_i)$ to

$$\boldsymbol{N} = \arg\min_{\hat{\boldsymbol{N}} \geq 0} \left\| (\boldsymbol{\delta V}) \cdot \hat{\boldsymbol{N}} - \boldsymbol{\Delta V} \right\|_2^2 . \tag{4.11}$$

Since this is mathematically an ill-posed problem that can lead to solutions with physically unrealistic high defect densities, a Tikhonov regularization [211] is added to ensure that the least-square solution results in smoother defect densities

$$\boldsymbol{N} = \arg\min_{\hat{\boldsymbol{N}} \geq 0} \left\| (\boldsymbol{\delta V}) \cdot \hat{\boldsymbol{N}} - \boldsymbol{\Delta V} \right\|_2^2 + \gamma^2 \left\| \hat{\boldsymbol{N}} \right\|_2^2 \tag{4.12}$$

where γ is the regularization parameter. The ESiD algorithm allows an efficient extraction of trap parameters from measurement data which can be compared for their agreement with DFT simulations [MJJ2, 212]

Part II

Defect Characterization at Cryogenic Temperatures

Chapter 5

Measurement Setup and Technologies

In the following section two measurement setups will be introduced, one setup where all devices have to be contacted manually and one setup with the additional option of measuring array structures, which allows the characterization of thousands of transistors which are digitally addressable. The two experimental setups are introduced in more detail in Section 5.1. Using the toolset, various different technologies have been characterized. An overview of all characterized technologies is given in Section 5.2. In Section 5.3 the working principle of dedicated *SmartArrays* is explained.

5.1 Measurement Setup

Two different measurement setups have been used for measurements presented in this work. A LakeShore CRX-4K cryogenic probe station is available, which uses a closed cycle refrigerator (CCR) for cooling to cryogenic temperatures. For this, a CCR compressor cools two separate stages via Helium lines. This allows to maintain elevated temperatures at the sample during the cooling phase of the system to avoid condensation at the sample. The temperature of the chuck and the stages can be controlled using a LakeShore Model 336 temperature controlling system. A turbo vacuum pump manufactured by Pfeiffer is used to guarantee a vacuum of around 10^{-6} hPa in the vacuum chamber. Up to six thermally shielded probe arms are mounted to the cryogenic stage which are equipped with Tungsten needles with a 10 µm tip radius to enable contacting small structures. This allows to contact the samples using micro-manipulated stages which enable translations in x, y and z direction. This setup has mainly been used for the characterization of single transistors. Every device thereby needs to be contacted manually, however, the probes need to be lifted during larger temperature changes to avoid scratching pads due to thermal expansion or contraction. The probe arms are connected via triaxial cables to a custom built measurement box, which is described in more detail in [140, 213, 214].

The other setup uses a LakeShore PS-VLT-CPX probe station, which allows cooling to 1.6 K using a conventional He dewar system. This can be done by lowering the

pressure inside the sample stage heat exchanger with an additional rotary vane pump. The probe arm assembly is the same as at the CRX-4K probe station (6 probe arms which can be manipulated in x, y and z direction), however, additionally there is a connection from the vacuum chamber to a custom designed 48-line break-out box. This allows for the characterization of programmable arrays of thousands of transistors using an adapted version of the custom built measurement box with ten synchronized source-measure-units (SMUs) [140, 213, 214].

Figure 5.1. A LakeShare CPX-4K cryogenic probe sation (left) was used for time-zero and reliability characterization of various technologies. A custom built measurement box with ten SMUs (right) was used for the characterization of transistor arrays. Left two images courtesy Lake Shore Cryotronics, Inc.

5.2 Investigated MOS Transistor Technologies

With the presented measurement setups the following technologies were used for the characterization of transient curves, random telegraph noise and bias temperature instability between 4 K and room temperature:

- *Technology A:* A commercial 28 nm bulk CMOS high-κ technology with a metal gate contact. In a rapid thermal oxidation process a thin interface layer is processed on the substrate. On top a HfO_2 high-κ layer is stacked exhibiting an equivalent oxide thickness (EOT) of 1.41 nm. The technology has been used for various characterization methods: Time-zero characteristics have been recorded on large area devices with the dimensions $W \times L = 10\,\mu m \times 1\,\mu m$ as presented in Section 6.1. The same

devices have been used for BTI measurements which are presented in Section 8.2. Scaled devices with dimensions of $W \times L = 100\,\text{nm} \times 28\,\text{nm}$ have been used for variability measurements presented in Section 6.4.1. Devices with $W = 100\,\text{nm}$ and different lengths of $L = 70\,\text{nm}, 100\,\text{nm}, 135\,\text{nm}, 170\,\text{nm}$ and $200\,\text{nm}$ have been used for variability measurement shown in Section 6.4.2.

- *Technology B:* A commercially available 180 nm SiON technology has been used as a simplified system to exclude trapping caused by the high-κ layer. Various device dimensions are available for this technology. Note that for the time-zero characterization the largest available geometry with $W \times L = 10\,\mu\text{m} \times 10\,\mu\text{m}$ has been used, see Section 6.1. At this geometry also BTI measurements have been conducted, which are shown in Section 8.2. For the characterization of single defects the smallest available geometry with designed dimensions of $W \times L = 220\,\text{nm} \times 180\,\text{nm}$ has been used, and the corresponding measurement results are presented in Section 7.4.

- *Experimental Batch C:* Novel nano-scaled MoS_2 devices have been fabricated by mechanical exfoliation of few-layer MoS_2 on 20 nm SiO_2/Si. Using electron beam lithogrphy and plasma etching, the channel was patterned on selected flakes. Afterwards, an evaporation and lift-off process has been used to create source and drain contacts. The device batch was fabricated in a university cleanroom. For RTN measurements presented in Section 7.5 devices with dimensions $W \times L = 70\,\text{nm} \times 70\,\text{nm}$ have been used, and for TDDS characterization presented in Section 8.3 devices with dimensions $W \times L = 100\,\text{nm} \times 70\,\text{nm}$ are selected.

5.3 Characterization of *SmartArray* Structures

The first generation of imec-built *SmartArrays* was developed in 2013 [215]. Since then, almost every year an updated version of the chip has been designed and taped out using different technology nodes, geometries, doping profiles, etc. [216–219, MJC3]. An overview of the different generations can be seen in Fig. 5.2. For this work, the designs from 2017 [MJC3] and from 2018 are used.

SmartArrays have several advantages compared to measurements at single devices where the devices have to be contacted manually in order to collect a statistical relevant amount of data for a certain technology. The array structures are mounted and bonded onto custom-built PCBs, which avoids the need of probes and allows to sweep over the whole temperature range within one measurement without having the problem of mechanical tension of the probes on the metal pads on the wafer chip. Furthermore, it reduces condensation and contacting issues which regularly occur during manual operation. However, the main advantage is the availability of thousands of transistors which can be addressed individually using shift registers. The circuit topologies of the 2017 design of the *SmartArray* can be seen in Fig. 5.3.

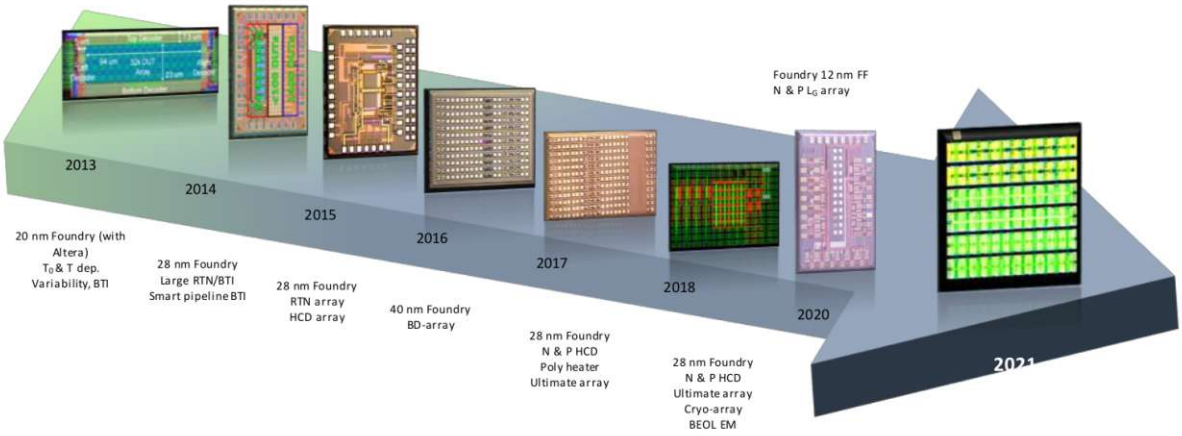

Figure 5.2. After the first generation of *SmartArrays* taped out in 2013 [215] containing 32000 commericallly available devices fabricated within a 20 nm technology node, almost every year an updated version with different devices, doping, geometries, etc. has been designed [216–219, MJC3].

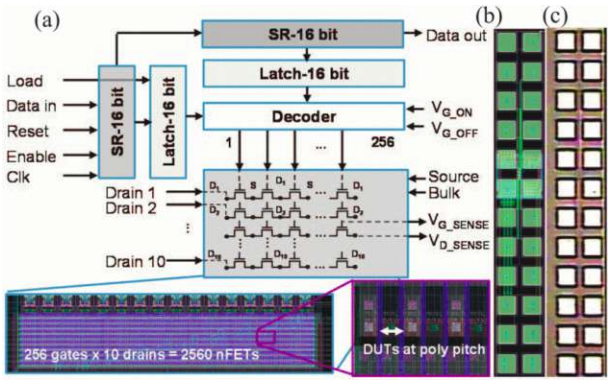

Figure 5.3. The circuit topology (a) of the *SmartArray* tape-out from 2017 allows to digitally address 2560 devices with dimensions $W \times L = 100 \times 28\,\mathrm{nm}^2$ via ten drain lines and 256 gate lines using a shift register. The circuit can be accessed via pads (b-c). Figure taken from [217].

The 2017 tape-out consists of twelve blocks with either 2560 nMOS or pMOS transistors of *Tech. A* each, with dimensions $W \times L = 100 \times 28\,\mathrm{nm}^2$ which can be addressed individually. Each block has 256 gate lines, which can be selected digitally via shift registers and latches and ten drain lines, which can be connected to the ten SMUs available in the custom-designed measurement tool. In the following sections, this tape out will be referred to as *SmartArray A*.

The 2018 tape-out is produced in two different flavors, nMOS and pMOS of *Tech. A* and will be referred to as *SmartArray B* in the following sections. One *SmartArray* contains 2500 transistors of one MOS type, see Fig. 5.4 (left). Five different geometries are available, with 500 transistors per geometry. The available devices exhibit $W \times L = 100\,\mathrm{nm} \times \{70, 100, 135, 175, 200\}\mathrm{nm}$. The first 50 drain lines allow the selection of 500 devices of the smallest geometry using ten gate lines. The next 50 drain lines correspond to the second smallest geometry and so on. To characterize the array

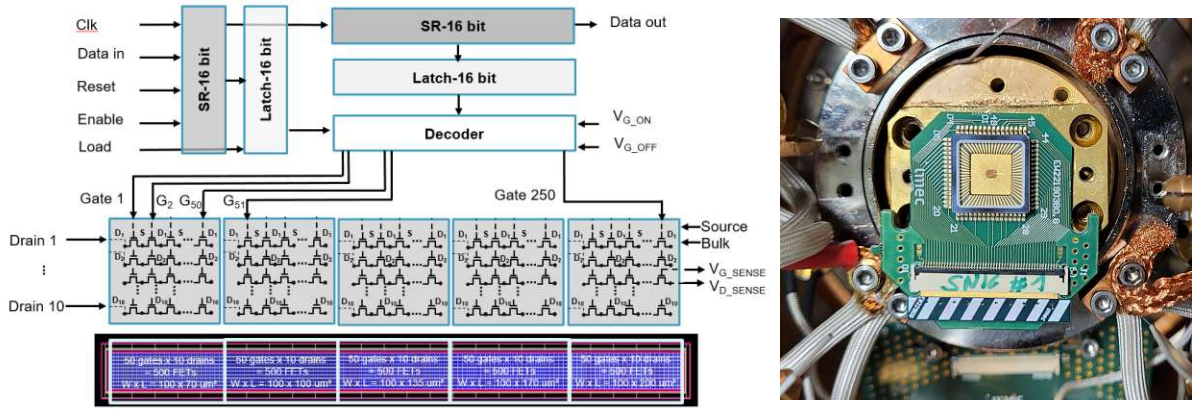

Figure 5.4. The circuit topology of the *SmartArray* tape-out from 2018 (left) shows how 2500 devices can be addressed via ten drain lines and 250 gate lines. This can be done using a shift register. The devices are organized in five blocks with 500 devices each, having different lengths values $W \times L = 100\,\text{nm} \times \{70, 100, 135, 175, 200\}\text{nm}$. The *SmartArray* is mounted and bonded on a PCB (right) which allows to connect the array to a breakout box providing BNC connectors.

it is encapsulated as shown in Fig. 5.4 (right) and bonded to a PCB. This allows to connect the pads to a breakout board in order to interface the bonded *SmartArray* in the cooled vacuum chamber to the measurement equipment at room temperature. The measurement equipment presented in this chapter has been used to measure time-zero characteristics and variability measurements which are shown in Chapter 6, charge noise characteristics presented in Chapter 7 and BTI characteristics discussed in Chapter 8.

Chapter 6

Time-Zero Characterization

Understanding the time-zero properties of a device under test is the basis for every variabiltiy and reliability study of a certain technology. The reason is that $I_D(V_G)$ and CV curves show not only a strong dependence on drain bias or device dimensions, but also on the device temperature. The temperature dependence is caused by reduced phonon-scattering, which increases the mobility of the charge carriers. Furthermore, the temperature-dependent shift of the Fermi-levels, the temperature dependence of the carrier distributions and the band gap also seriously affect the device characteristics. All these effects can change the shape of $I_D(V_G)$ and CV curves dramatically, which has in turn a strong impact on any reliability study. The impact of the temperature on the transition curves is discussed detailed in Section 6.1, the impact on CV curves is discussed in Section 6.2.

Additionally, quantum mechanical effects become dominant towards cryogenic temperatures, which do not occur in the classical limit. An example for this is resonant tunneling caused by quantum dots. This effect has been measured across various MOSFET technologies and is presented in Section 6.3.

After studying the temperature dependent time-zero characteristics on large-area devices, the characteristics are studied on *SmartArrays*, employing thousands of devices which can be addressed digitally. This enables the study of the temperature dependence of the variability of parameters which are an important measure for the MOSFET performance of the devices. This has been done for two different *SmartArray* generations with various geometries. The results are presented in Section 6.4.

6.1 Transfer Characteristics

The most widely used method for device characterization is the analysis of the $I_D(V_G)$ curve of a single MOSFET. This can be done for a pristine or a stressed device and allows the extraction of important quantities which determine the quality and performance of a MOSFET. To record an $I_D(V_G)$ curve, V_D is kept constant, while V_G is swept from deep depletion to inversion. From the recorded $I_D(V_G)$ curve it is possible to

extract multiple quantities like sub-threshold swing SS, threshold voltage V_{th}, on-state current I_{ON}, transconductance g_m, or leakage current I_{OFF}, as shown in Fig. 6.1. All these parameters depend not only on the device dimensions, but also show a strong temperature dependence as can be seen by a comparison of the $I_D(V_G)$-curve for 4 K and 298 K in Fig. 6.1. Depending on the examined quantity, $I_D(V_G)$ curves are plotted on a logarithmic scale (particularly for subthreshold quantities) or on linear scales as can be seen in Fig. 6.1 (right). The temperature dependence of the transfer characteristics can be seen for *Tech. A* in Fig. 6.2 (top) and *Tech. B* in Fig. 6.2 (bottom) for both, nMOS and pMOS. Towards 4 K the $I_D(V_G)$ curve becomes steeper and the threshold voltage and on-state current increase.

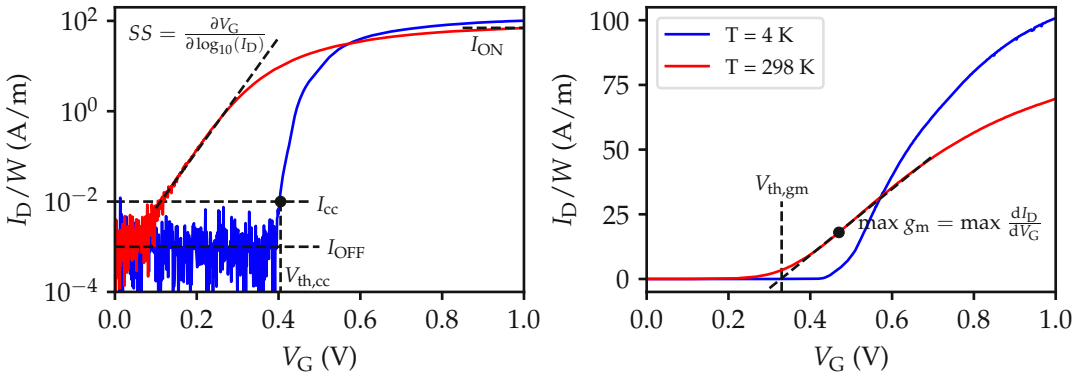

Figure 6.1. The schematic $I_D(V_G)$ curves for 298 K and 4 K show a strong temperature dependence which influences various important device parameters, such as the on-current I_{ON}, the subthreshold swing SS, the leakage current I_{OFF}, the (maximal) transconductance g_m, or the threshold voltage V_{th} which can be defined at a constant current I_{cc} or with the maximum transconductance max(g_m).

6.1.1 Subthreshold Swing

The steepness of the $I_D(V_G)$ is well-know as sub-threshold slope and is central for the switching between ON-state and OFF-state of a transistor and thus an essential performance criterion. For device characterization often the reciprocal value is used, which is known as subthreshold swing SS. In the classical limit, the SS is defined in the weak inversion regime and given by [220]

$$SS = \frac{\partial V_G}{\partial \log_{10}(I_D)} = \ln(10)\frac{k_B T}{q}\frac{C_{ox} + C_D + C_{it}}{C_{ox}} = m\ln(10)\frac{k_B T}{q}, \tag{6.1}$$

where C_{ox} is the capacitance of the oxide, C_D the depletion capacitance and C_{it} the capacitance of the interface states. The factor m which is defined by the capacitances is often called subthreshold-swing factor. The minimal value for SS at a certain temperature is reached in the limit $C_D \to 0$ F, $C_{it} \to 0$ F and $C_{ox} \to \infty$ F. Experimentally, SS can be extracted using a linear fit for the subthreshold region of the logarithmic

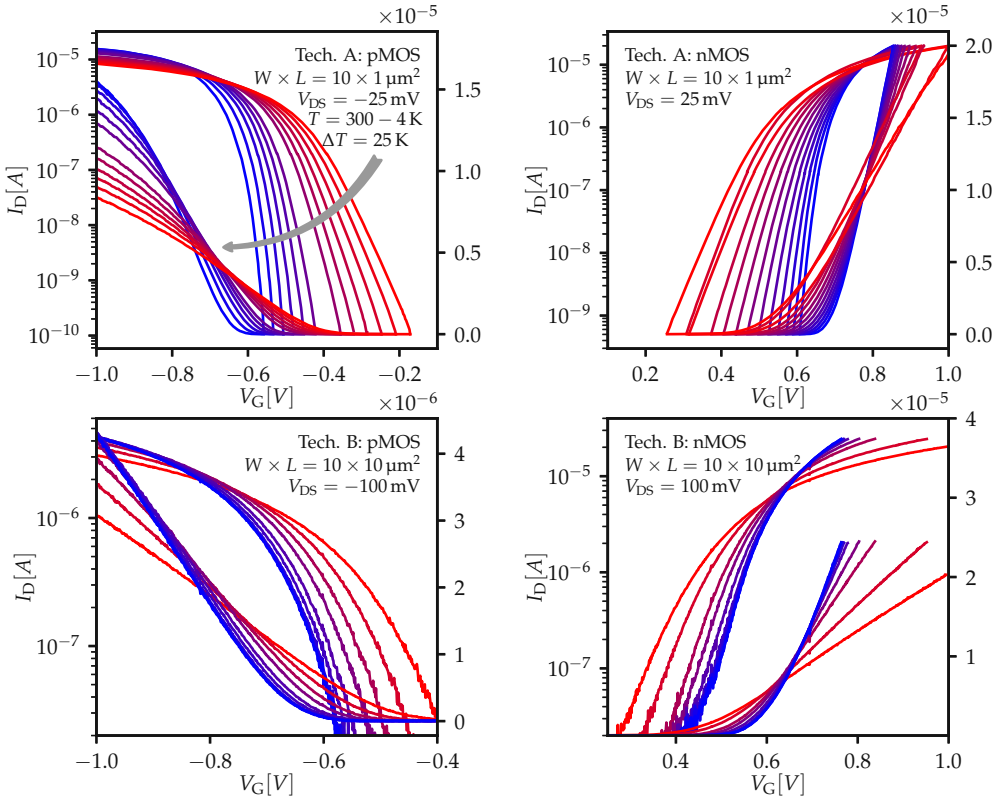

Figure 6.2. The transfer characteristics of *Tech. A* (top) and *Tech. B* (bottom) show qualitatively a similar behavior for both pMOS (left) and nMOS (right). Towards 4 K the $I_D(V_G)$ gets steeper and I_{ON} increases. All $I_D(V_G)$ curves intersect close to a zero temperature coefficient point.

$I_D(V_G)$ curve. By doing so the temperature dependence of SS can be extracted from $I_D(V_G)$ measurements, and is shown in Fig. 6.3.

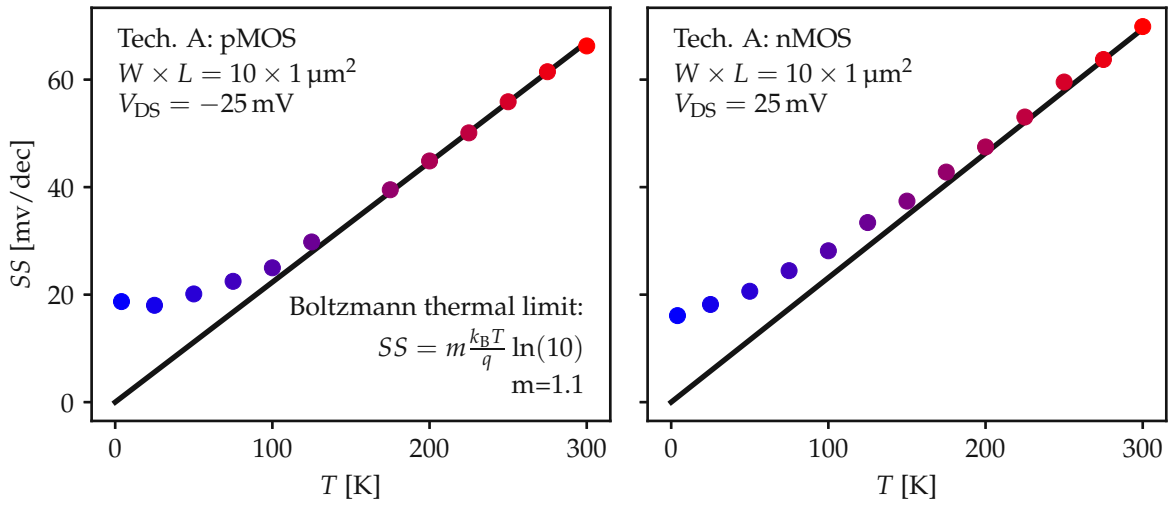

Figure 6.3. At temperatures above around 150 K, the subthreshold swing follows the Boltzmann relation (black line). Towards cryogenic temperatures it starts to saturate which can be seen for pMOS (left) and nMOS (right). This behavior has been observed for various technologies and can be explained by band tail states [110].

According to the classical Boltzmann thermal limit SS is defined as $m\ln(10)k_BT/q$. Here, m is the subthreshold-slope factor which can be approximated by $m = \partial V_G/\partial\varphi_S$ with φ_S being the surface potential. The classical Boltzmann thermal limit is experimentally not accessible below around 100 K, but starts to saturate towards a constant value. This can be seen in Fig. 6.3 for *Tech. A* for pMOS (left) and nMOS (right). This saturation effect of the subthreshold swing has been measured across multiple different FET technologies as shown in [110]. The deviation from the Boltzmann thermal limit can be explained by band tail states near the interface between the substrate and oxide. These band tail states occur for example due to static or dynamic disorder from interface defects, impurities or electron-phonon scattering and lead to a blurred band edge. This can be modeled precisely using Gauss hypergeometric functions for the band tails [221]. However, these functions require a large computational effort compared to a simplified model using exponential band tails $g(E)$, as can be seen in Fig. 6.4, which is derived in detail in [MJJ4].

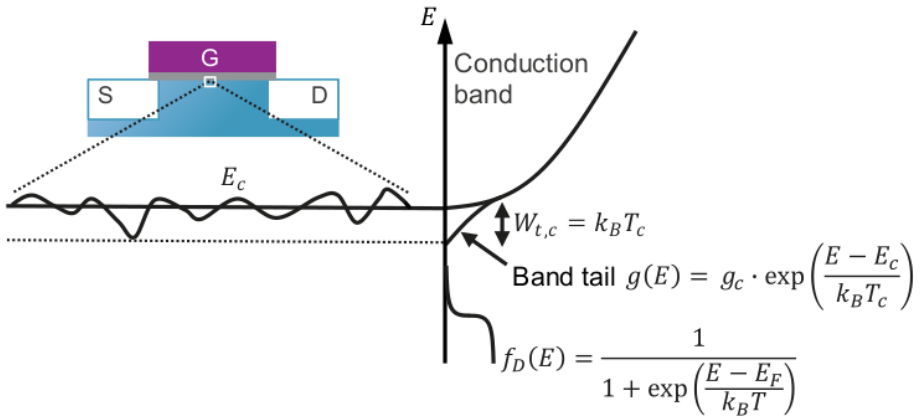

Figure 6.4. Crystalline disorder leads to blurred band edges. These can be modeled assuming exponentially decaying band tail states depending on a critical temperature T_c. This widened band edge contribute to the energy scan by the Fermi distribution and can explain the saturation of the subthreshold slope. Figure taken from [MJJ4].

As shown by Beckers *et al.* [MJJ4] the introduction of a critical temperature T_c, which is typically around 30 K to 50 K, allows the adaption of the classical SS to

$$SS = m\frac{k_BT}{q}\ln(10)b_c \tag{6.2}$$

with

$$b_c = \frac{1+f(\theta)\cdot y}{1+\theta\cdot f(\theta)\cdot y} \tag{6.3}$$

$$f(\theta) = \frac{\pi\cdot(\theta-1)}{\sin(\pi\theta)} \tag{6.4}$$

and

$$y = \exp\left[(1-\theta)\left(\frac{x}{k_B T}\right)\right]. \tag{6.5}$$

In this closed form expression $x = E_C - E_F$ is the energetic offset between the conduction band and the Fermi level and $\theta = T/T_c$. This closed-form simplification of the model presented in [221] describes the temperature dependence of SS and the occurring saturation efficiently and is thus well suited for usage in commercial device simulators.

6.1.2 Transconductance

The transconductance g_m (short for transfer conductance) relates the drain current with the gate voltage via

$$g_m = \frac{dI_D}{dV_G} \tag{6.6}$$

at a constant V_D. Towards lower temperatures, electron-phonon scattering gets reduced, leading to a higher carrier mobility. This has the effect of an increasing transconductance, as can be seen in Fig. 6.5 for *Tech. A* for both pMOS (left) and nMOS (right). The overall shape of g_m gets steeper because the Fermi-Dirac distribution becomes narrower at low temperatures. The point of maximum g_m which also gets continuously larger for lower temperatures, as can be seen in Fig. 6.5 (bottom), is widely used for the definition of the threshold voltage, as discussed in the upcoming section in detail.

It has to be noted that the g_m characteristics shows a maximum for each transistor. The point of $\max(g_m)$ can become degraded due to higher contact resistance or an increase in the number of interface states. The latter affect g_m, because the mobility of the carriers reduces due to Coulomb scattering at interface defects. This effect is deliberately used to link a decreasing $\max(g_m)$ to an increasing concentration of interface defects [180].

6.1.3 Threshold Voltage

The threshold voltage V_{th} of a MOSFET is the minimal gate voltage V_G which needs to be applied to turn the device on, which means creating a conduction path between source and drain [220]. While SS or transconductance are well defined, there are multiple definitions for the threshold voltage. In this work, two different methods for extracting the threshold voltage from measurements will be used: The constant current method and the $\max(g_m)$ method, referred to $V_{th,cc}$ and $V_{th,gm}$, respectively. For the constant current method, a constant current I_{cc} is defined at which the corresponding threshold voltage $V_{th,cc}$ is extracted, as can be seen in Fig. 6.1(a). For the $\max(g_m)$

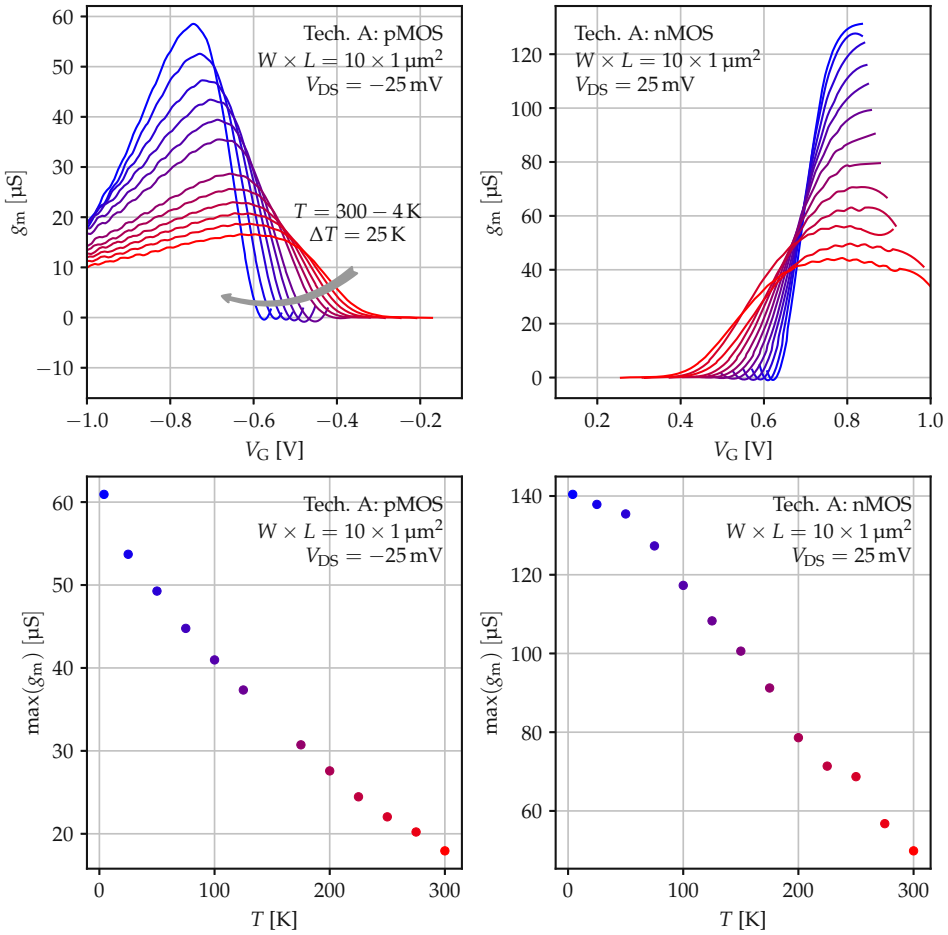

Figure 6.5. The shape of the transconductance shows a strong temperature dependence for both pMOS (left) and nMOS (right) of *Tech. A*. The steeper and sharper g_m at 4.2 K can be explained by the sharper Fermi-Dirac distribution at cryogenic temperatures; $\max(g_m)$ increases at lower temperatures because the mobility increases due to reduced phonon-scattering.

method the point of $\max(g_m)$ is extracted. Then the intersection point of the tangent line with $I_D = 0\,A$ is found. The corresponding voltage is defined as $V_{th,gm}$.

An extraction of $V_{th,cc}$ and $V_{th,gm}$ is shown for *Tech. A* in Fig. 6.6 for pMOS (left) and nMOS (right). As can be seen, with the two methods not the same threshold voltage values are extracted, in fact, they do not even exhibit the same temperature dependence. In the case of pMOS, $V_{th,cc}$ decreases by approximately 350 mV while $V_{th,gm}$ decreases only by approximately 200 mV when the temperature is changed from 300 K to 4 K. Even worse is the difference for nMOS. Here not only a difference in the absolute shift, but also a change of the shape for the T-dependence can be observed. While for pMOS the T-dependence is almost perfectly linear, the nMOS device shows a saturation of the threshold voltage, which is stronger for $V_{th,cc}$ than for $V_{th,gm}$. This rather complex T-dependence of the different V_{th} definitions is discussed in detail in [111]. By discussing the impact of the temperature dependence of the bulk Fermi potential, dopant freezeout, field-assisted dopant ionization, bandgap widening and interface traps, Beckers *et. al.*

model V_{th} and show that the saturation of V_{th} for nMOS towards 4.2 K can be explained by the presence of interface defects in combination with the shift of the bulk Fermi potential [111].

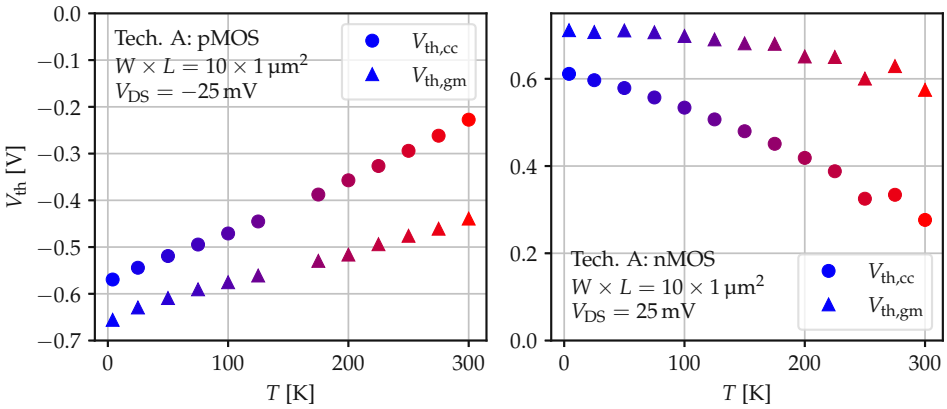

Figure 6.6. The threshold voltage increases towards cryogenic temperatures for both nMOS and pMOS of *Tech. A*. Different definitions and extraction methods of V_{th} lead to a different T-dependence which is discussed in detail in [111].

The threshold voltage plays an important role for the evaluation of the reliability of a certain technology. It is often used as reference before applying stress to a device. During stress the shape of the $I_D(V_G)$ curve can change, which is often simplified by assuming a parallel shift of the whole $I_D(V_G)$ curve. The change in V_{th} extracted from the initial and stress device is often compared. For this, V_{th} needs to be extracted from the $I_D(V_G)$ which makes it necessary to decide on a certain definition. These methods are discussed in detail in the upcoming sections about reliability characterization.

6.1.4 ON-State Current

The ON-state current I_{ON} is extracted in the saturation region, typically at $V_G = V_{DD}$. I_{ON} for *Tech. A* pMOS (left) and nMOS (right) can be seen in Fig. 6.7, extracted at $V_G = -1\,V$ and $V_G = 0.85\,V$, respectively. As can be seen, I_{ON} decreases for pMOS and increases for nMOS with decreasing temperatures, which can be explained by the increasing mobility of the charge carriers. This means that at cryogenic temperatures the efficiency of MOS devices increases, which is beneficial for the performance of circuits at these temperatures.

6.2 Capacitance-Voltage Measurements

Not only measuring of I_D, but also measuring the gate capacitance allows the extraction of defect related parameters. First a reference CV curve is measured and then the device is stressed under well defined conditions. Afterwards, a second CV curve is recorded and from the variation between the pre- and post-stress CV curves one can

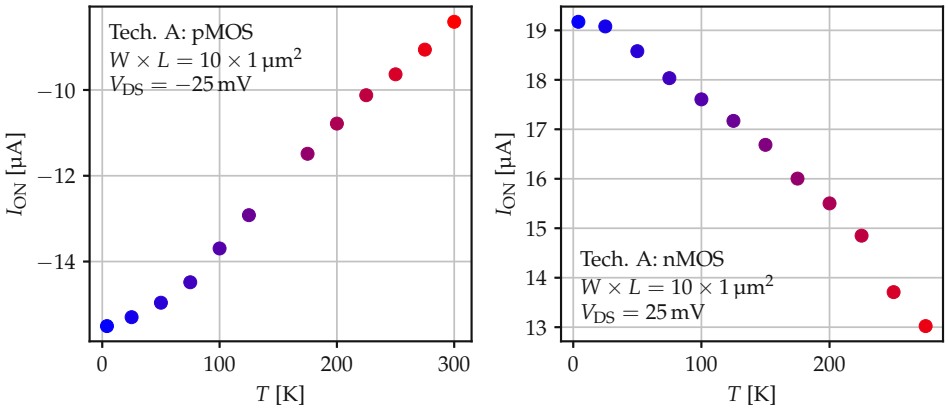

Figure 6.7. The ON-current of *Tech. A* decreases for pMOS (left) and increases for nMOS (right) towards 4.2 K. This can be explained by the increasing mobility due to reduced phonon-scattering.

extract various defect properties. However, in this work, CV curves have not been used for defect characterization, but are important for the calibration of the electrostatics of our reliability simulator *Comphy*, which will be introduced detailed in Section 4.

For recording a CV curve a DC-voltage V_G is applied at the gate of the device under test (DUT) which is superposed with small AC signal, which typically has a frequency between 10 kHz and 1 MHz. The mean voltage V_G is then swept from accumulation to inversion. By measuring the gate current at every V_G-step, the capacitance can be computed by

$$C = \frac{I_G}{2\pi f V_{AC}} \tag{6.7}$$

with f being the used frequency, V_{AC} being the amplitude of the AC-signal and I_G being the measured gate current. Depending on the defect properties and the charge trapping kinetics, the shape of the CV curves can change which makes CV measurements a powerful tool for reliability characterization which provides valuable information about the energetic position of the traps [140, 222].

The CV curves in Fig. 6.8 were recorded on *Tech. B* employing a Keithley 590 CV analyzer using a frequency of $f = 100$ kHz and are plotted normalized by C_{max}. The temperature has been swept from 4 K to 300 K. As can be seen for both nMOS and pMOS, the transition between depletion and inversion becomes sharper towards cryogenic temperatures. This is because the Fermi distribution becomes sharper and thus scans a smaller energy region. Figure taken from [MJC4].

6.3 Resonant Tunneling at Cryogenic Temperatures

Transistors used in digital applications should be able to switch between ON-state and OFF-state as quickly as possible. At cryogenic temperatures, the switching between

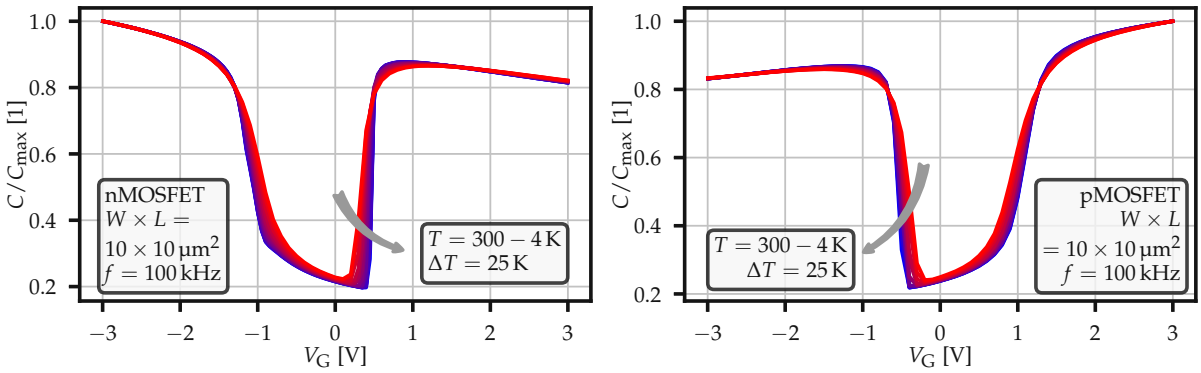

Figure 6.8. CV curves measured on *Tech. B* for temperatures between 4 K and 300 K using an AC-frequency of $f = 100\,\text{kHz}$. Towards cryogenic temperatures the transition between depletion and inversion gets sharper due to the sharper Fermi distribution.

ON- and OFF-state is superior compared to room-temperature, because of the increased ON-state current and the increased transconductance. However, at scaled devices resonance phenomena have been observed across various technologies including 40 nm bulk [223], 28 nm bulk CMOS [200, 224], 40 nm SOI [225], 22 nm FDSOI [226] and 16 nm FinFET [227], which can lead to large humps in the transition from the OFF- to ON-state. Such humps have been measured on *Tech. A* on nMOS *SmartArray A* (Fig. 6.9 (left)) with dimensions $W \times L = 100\,\text{nm} \times 28\,\text{nm}$ at 4.2 K with $V_\text{D} = 50\,\text{mV}$ [MJC5] and on a device of the nMOS-flavored *SmartArray B* in Fig. 6.9 (right) with dimensions $W \times L = 100\,\text{nm} \times 70\,\text{nm}$ at 4.2 K with $V_\text{D} = 5, 10, ..., 50\,\text{mV}$.

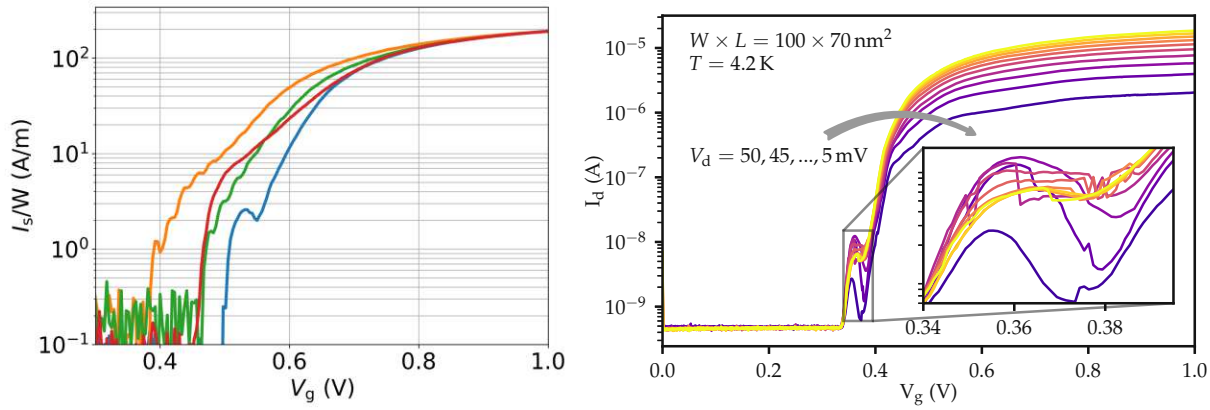

Figure 6.9. Resonant tunneling measured at cryogenic temperatures on nMOS of *Tech. A* using *SmartArray A* (left) and *SmartArray B* (right). The oscillations caused by resonant tunneling fade out with increasing V_D. Figure taken from [MJC5].

These humps in the $I_\text{D}(V_\text{G})$ curves can be extremely prominent and can even lead to a negative transconductance g_m, as can be seen in Fig. 6.10.

The resonances are caused by quantum confinement in the channel due to ionized dopants, impurities or defects. Resonant tunneling leads then to the occurring resonance in I_D. In literature, this is often referred to as *Coulomb oscillation* [223, 224, 227]. The resonance shows a large V_D and temperature dependence. A larger V_D leads to a fade

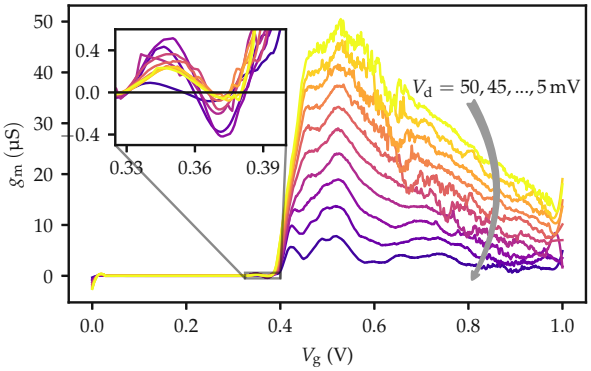

Figure 6.10. Distinctive oscillations, as can be seen in Fig. 6.9 (right), can lead to a negative transconductance. This effect fades out with increasing V_D and towards higher temperatures.

out of the resonance, as can be seen in Fig. 6.9 (left). This is caused by the drain-induced barrier lowering (DIBL) effect which allows more carriers to overcome the barrier and diminishes the reduced conductance caused by the resonance [227]. Towards higher temperature the resonance fades out, because the higher thermal energy allows more carriers to overcome the occurring tunneling barriers [227].

6.4 Variability Characterization

The characterization of time-zero properties is central for determining the performance of transistors. This characterization is typically done on single transistors using probe stations with manual needle arms. Using this approach, the collection of a large set of measurement data is extremely time consuming and makes variability studies unfeasible. However, a small variability is, specially for cryogenic applications which typically use a very low V_{DD}, inevitable. Therefore, a different approach is used based on *SmartArrays* with thousands of transistors which can be addressed individually as explained in detail in Section 5.3. This allows to conduct measurements on thousands of pristine devices efficiently and to characterize the variability and mismatch of DUTs between 4.2 K and room temperature.

SmartArray A and *SmartArray B* were characterized by Alexander Grill and the author of this thesis during research visits at *imec* in 2019 and 2021, respectively. Measurement results have been published in [MJC3, MJC5].

6.4.1 Variability Study on *SmartArray A*

SmartArray A consists of twelve Blocks with 2 560 transistors per block, in total 30 720 transistors of *Tech. A* with $W \times L = 100\,nm \times 28\,nm$, and is described in detail in Section 5.3. For the variability study, for each temperature of the set $T \in \{4.2\,K, 77\,K, 150\,K, 225\,K, 300\,K\}$ two blocks of pristine nMOS devices (5 120 transistors) have been characterized. For this, transfer characteristics between $V_G = 0\,V$ and

$V_G = 1\,V$ have been recorded in the linear and saturation region using $V_D = 50\,mV$ and $V_D = 1\,V$, respectively.

The recorded transfer characteristics in the linear region are shown for RT and 4 K in Fig. 6.11 (a,c). The red arrows in the subthreshold region indicate that the variability increases towards cryogenic temperatures. This holds specially true in the subthreshold regime. The same can be observed for the transconductance g_m in Fig. 6.11 (b,d). Also for g_m the variability increases towards 4 K. The variability has a maximum at the same V_G, at which g_m has a maximum.

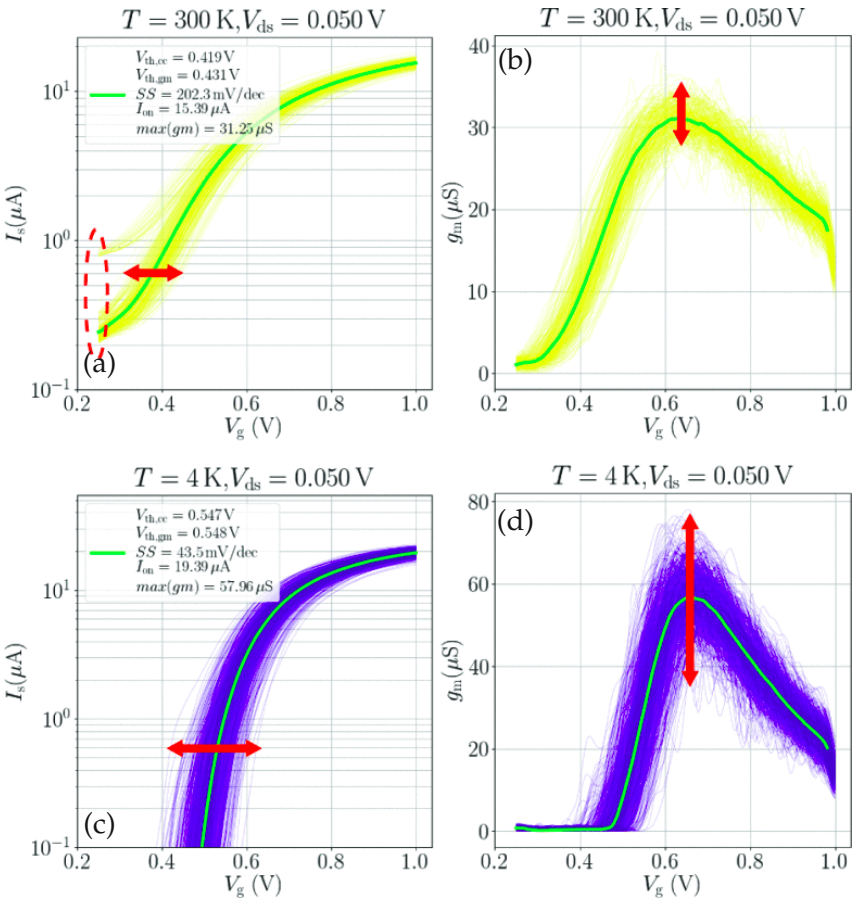

Figure 6.11. Transfer characteristics measured on 2 560 devices of *SmartArray A* at various temperatures. The variability in the subthreshold region at RT (a) is smaller compared to 4 K (c), as indicated by the red arrows. The same holds true for the derived transconductance in (b,d). Figures taken from [MJC3].

An increased variability in the subthreshold region has been reported before [228, 229]. Typically, it is attributed to resonant tunneling caused by surface roughness, defects or dopants. The resonance leads to humps in the transfer curves or to a double threshold behavior and therefore massively increases the variability [MJC3, MJC5]. The role of resonant tunneling is discussed in more detail in Section 6.3. The four selected transfer lines from this measurement set in Fig. 6.9 (left) show clearly how resonant tunneling affects the variability.

The variability of $V_{\mathrm{th,g_m}}$, SS, I_{ON} and $\max(g_{\mathrm{m}})$ extracted from the recorded transfer curves in the linear region can be seen in the quantile plots in Fig. 6.12 (a,b,c,d), respectively. A significant increase in the variability towards cryogenic temperatures can be seen only for $\max(g_{\mathrm{m}})$. $V_{\mathrm{th,g_m}}$, SS and I_{ON} show the well known temperature dependence discussed in detail in Section 6.1, however, the variability stays approximately constant. The measurement at 300 K shows an offset which can be most likely attributed to leakage stemming from broken devices.

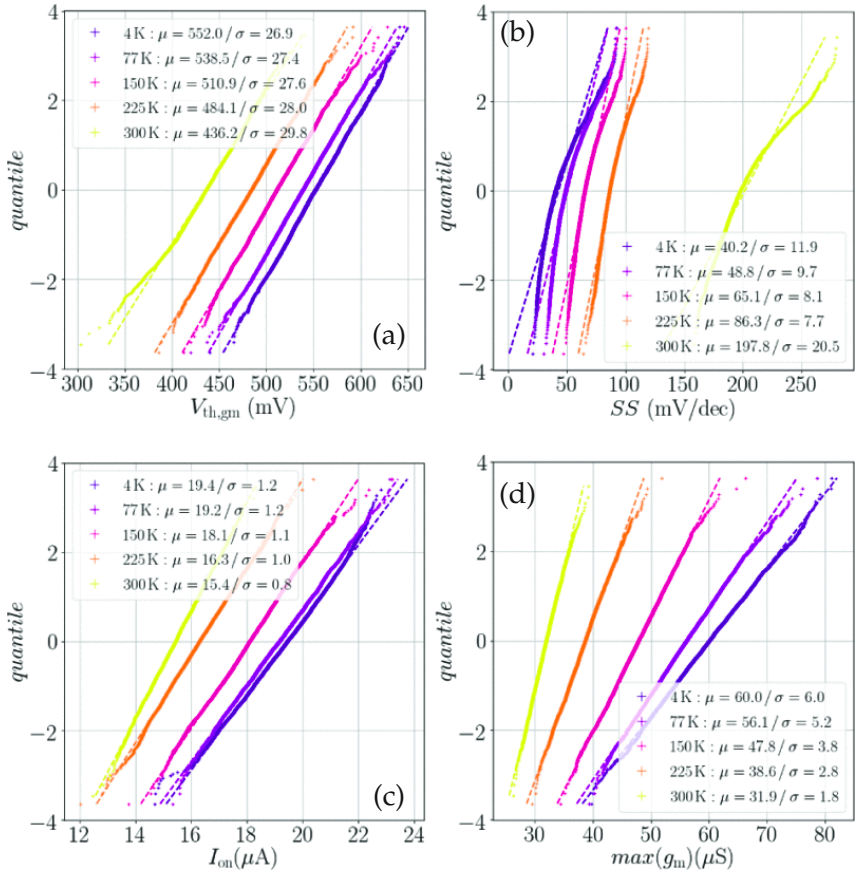

Figure 6.12. In the quantile plots of $V_{\mathrm{th,g_m}}$, SS, I_{ON} and $\max(g_{\mathrm{m}})$ extracted in the linear region all parameters show the well known temperature dependence discussed in Section 6.1. While the variability of $V_{\mathrm{th,g_m}}$, SS and I_{ON} shows no temperature dependence, it significantly increases for $\max(g_{\mathrm{m}})$ towards cryogenic temperatures. The offset of the measurement at 300 K can be most likely attributed to leakage stemming from broken devices. Figures taken from [MJC3].

The quantile plots in Fig. 6.12 correspond to the main diagonal in the correlation plot for the linear regime in Fig. 6.13 (top). The correlation plot in Fig. 6.13 (bottom) corresponds to an equivalent measurement set in the saturation region. While $\max(g_{\mathrm{m}})$ shows a strong variability increase in the linear region, the variability is rather constant across all temperatures in the saturation region. The same holds true for SS, $V_{\mathrm{th,g_m}}$ and I_{ON}.

The histograms above the main diagonal show a strong negative correlation between $V_{\mathrm{th,g_m}}$ and I_{ON} in both the linear and the saturation region. This can be explained by

the fact that $V_{\mathrm{th},g_\mathrm{m}}$ is defined using $\max(g_\mathrm{m})$ but I_ON is defined at a constant voltage $V_\mathrm{G} = 0.9\,\mathrm{V}$ instead of a constant overdrive. The strong positive correlation between I_ON and $\max(g_\mathrm{m})$ is caused by the effect that both are strongly affected by the increasing mobility. Both correlations slightly decrease towards cryogenic temperatures. This could arise from mobility saturation, but also from other effects like the more dominant contact resistance at 4.2 K [MJC5]. The leaking measurements can be clearly seen in the scatter plots below the main diagonal, where the cloud at 300 K shows a clear offset.

In essence, the same relations as in the correlation plots can be seen in the mismatch plots for the linear region and the saturation region in Fig. 6.14 (top) and (bottom), respectively. The parameter mismatch is defined by the difference of a parameter of two neighboring devices, e.g.

$$\Delta I_\mathrm{ON}(N+1) = I_\mathrm{ON}(N+1) - I_\mathrm{ON}(N). \tag{6.8}$$

This definition takes both pairs $\{N+1, N\}$ and $\{N, N-1\}$ into account.

As shown before in the correlation plots, the histograms above the main diagonal show strong correlations between $\Delta V_{\mathrm{th},g_\mathrm{m}}$ and ΔI_ON and between $\Delta \max(g_\mathrm{m})$ and ΔI_ON. The variability of the mismatch is slightly higher at cryogenic temperatures compared to RT, as can be seen in the main diagonals. This effect is most dominant for the mismatch of $\max(g_\mathrm{m})$. The offset in the measurement at 300 K cancels out in the mismatch plot.

For the set of measured transfer curves, mean SS, $\max(g_\mathrm{m})$, $V_{\mathrm{th},g_\mathrm{m}}$ and I_ON can be computed. As can be seen in Fig. 6.15, the mean values for both the linear and the saturation region show the same behavior discussed for large area devices in Section 6.1. The mean SS shows the typical saturation towards cryogenic temperatures, which is explained by band tail states, see Section 6.1.1. The mean $\max(g_\mathrm{m})$ and I_ON increase at lower temperatures due to the increasing charge carrier mobility caused by less phonon scattering, as discussed in Section 6.1.2 and 6.1.4. The bars showing the confidence interval of one-sigma clearly show the increasing variability of $\max(g_\mathrm{m})$. The increasing $V_{\mathrm{th},g_\mathrm{m}}$ and I_ON show a slight saturation behavior, which again can be explained by band tail states.

6.4.2 Variability Study on *SmartArray B*

SmartArray B exists in two variations, either with 2 500 nMOS or pMOS transistors of *Tech. A*. All transistors have a width of $W = 100\,\mathrm{nm}$, while there are sets of 500 devices with lengths $L \in \{70\,\mathrm{nm}, 100\,\mathrm{nm}, 135\,\mathrm{nm}, 170\,\mathrm{nm}, 200\,\mathrm{nm}\}$. A detailed description of *SmartArray B* can be found in Section 5.3. For every set of dimensions, transition curves have been recorded for the temperatures $T = 4.2\,\mathrm{K}, 77\,\mathrm{K}, 150\,\mathrm{K}, 225\,\mathrm{K}$, and 300 K in the linear region using $V_\mathrm{D} = 50\,\mathrm{mV}$ and in the saturation region using $V_\mathrm{D} = 1\,\mathrm{V}$.

The recorded transfer curves are shown in Fig. 6.16. For both the linear and the saturation region and across all gate lengths the $I_\mathrm{D}(V_\mathrm{G})$ curves shift and get steeper towards 4.2 K, as is well known from large area devices and discussed in detail Section 6.1. The

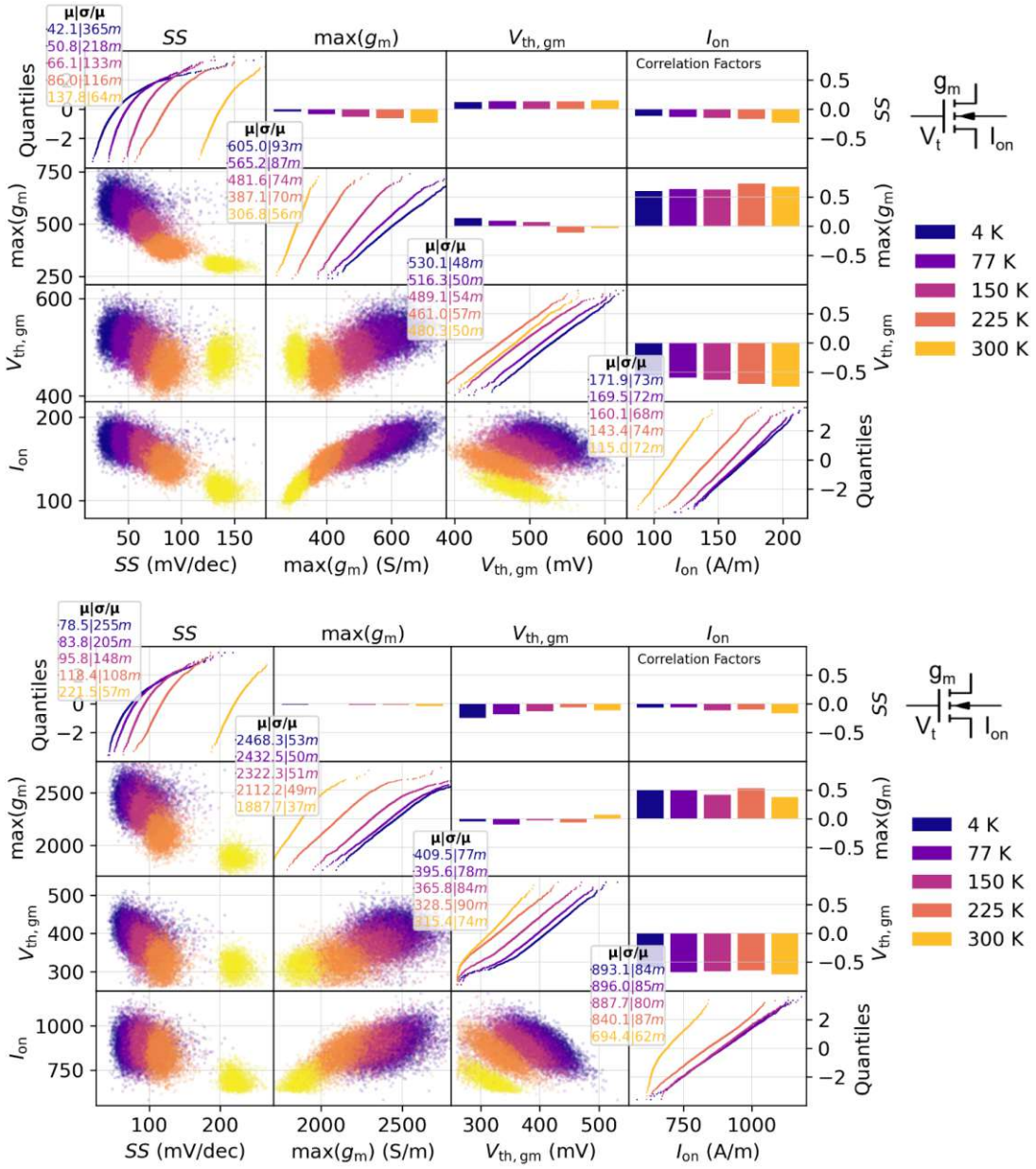

Figure 6.13. The correlation plots for the linear (top) and saturation (bottm) region show in the main diagonal that the variability of $\max(g_m)$ increase towards 4 K, while it stays approximately constant for SS, V_{th,g_m} and I_{ON}. The strong negative correlation between I_{ON} and V_{th,g_m} can be explained by the definition of I_{ON} at $V_G = 0.9$ V, while the strong positive correlation between $\max(g_m)$ and I_{ON} is caused by the fact that both depend on the increasing mobility towards 4 K. The scatter plots show the damaged devices used at 300 K. Figures taken from [MJC5].

transfer curves shift to higher gate voltages with increasing lengths, as it is well known from literature. The $I_D(V_G)$-curves show a large asymmetry between nMOS and pMOS, which is not predicted by the PDK. The reason for that is still debated.

Using the recorded transfer curves in Fig. 6.16, the transconductance g_m can be derived for all drawn dimensions and measurement regimes, as can be seen in Fig. 6.17. The g_m curves show clearly, that the variability in g_m increases towards cryogenic

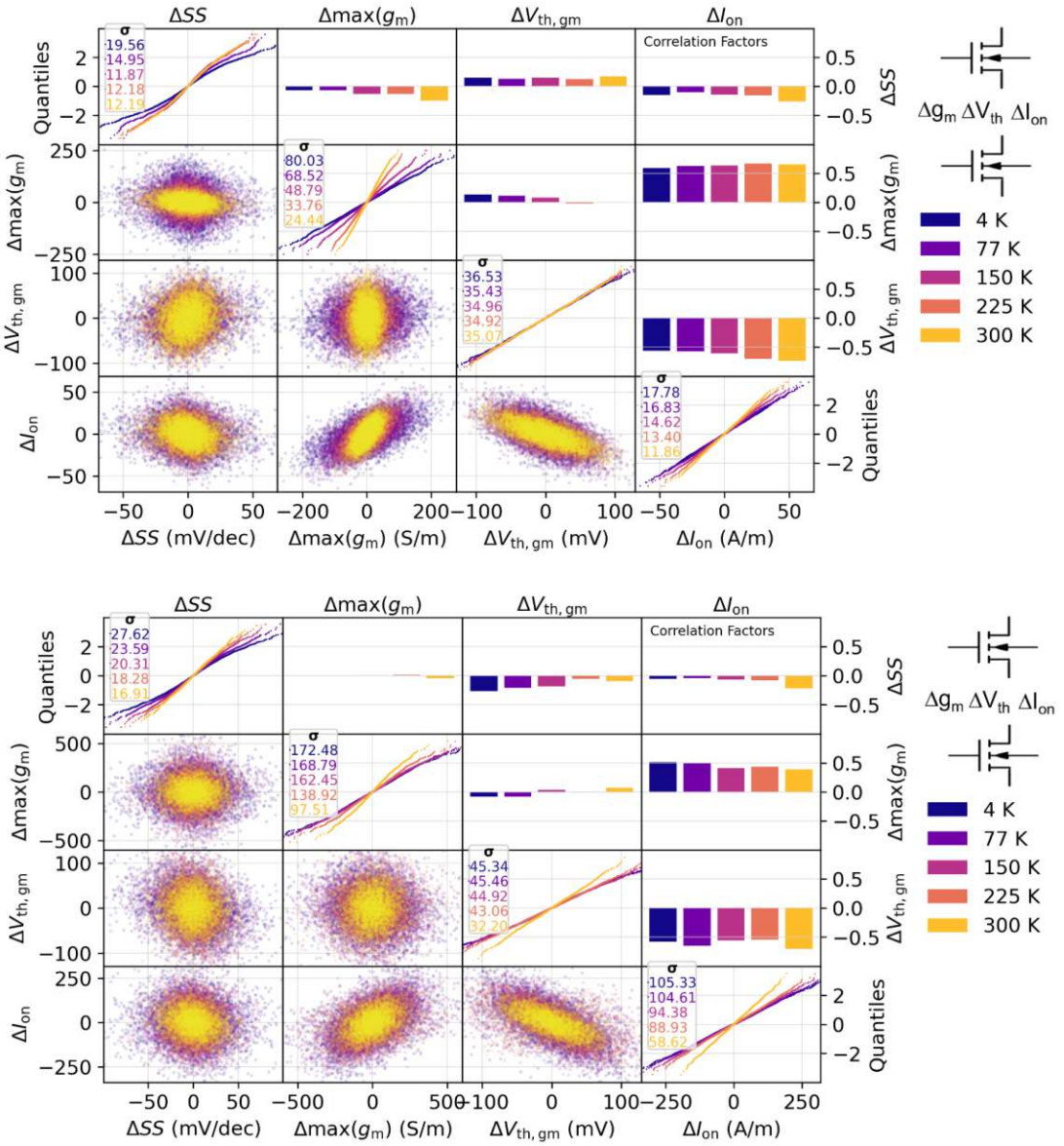

Figure 6.14. The same general trends can be seen in the mismatch plots for the linear (top) and the saturation (bottom) regime as before in the correlation plots Fig. 6.13. The variability of $\Delta \max(g_\mathrm{m})$ increases towards 4 K while it stays approximately constant for ΔSS, $\Delta V_{\mathrm{th},g_\mathrm{m}}$, and ΔI_ON. The histogram above the main diagonal strong negative correlation between ΔI_ON and $\Delta V_{\mathrm{th},g_\mathrm{m}}$ which can be explained by the definition of I_ON at $V_\mathrm{G} = 0.9\,\mathrm{V}$. The strong positive correlation between $\Delta \max(g_\mathrm{m})$ and ΔI_ON is caused by the fact that both depend on the increasing mobility towards 4 K. The damaged devices used at 300 K can not be seen in the mismatch plots. Figures taken from [MJC5].

temperatures. The maximum variability is in the same gate voltage region as $\max(g_\mathrm{m})$. Moreover, it can be seen that the variability increases towards larger device dimensions. The asymmetry in the transfer curves propagates to the transconductance, leading to a g_m in nMOS devices which is approximately by a factor 0.5 larger than in pMOS devices. The point of the maximum transconductance of the devices with $L \geq 135\,\mathrm{nm}$ is in the

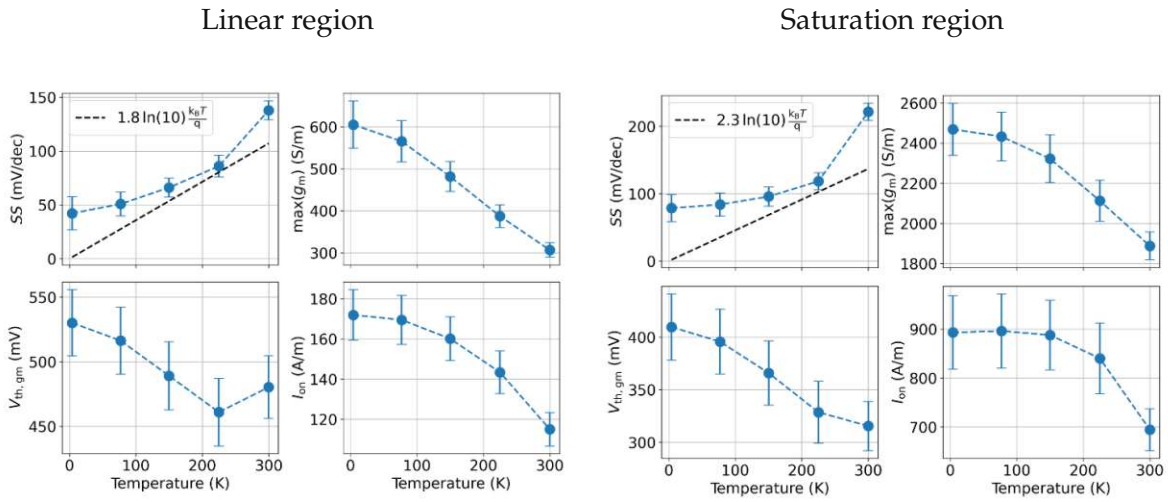

Figure 6.15. The mean values of SS, $\max(g_m)$, V_{th,g_m}, and I_{ON} in the linear and in the saturation region show the same behavior as large area devices. SS shows a typical saturation behavior, V_{th,g_m}, $\max(g_m)$, and I_{ON} increase towards 4 K. The confidence interval of one-sigma shows that the variability of $\max(g_m)$ increases while it stays approximately constant for all other parameters. Figures taken from [MJC5].

saturation region very close to the edge of the measurement window, or even outside of the window. This makes a further analysis of derived parameters difficult.

As it has been done before for *SmartArray A*, the recorded set of transition curves allows to analyze the correlation and mismatch of SS, $\max(g_m)$, V_{th,g_m}, and I_{ON} for every used dimension. Exemplary, correlation and mismatch plots for nMOS and pMOS in the linear region are shown in Fig. 6.18 for the smallest geometry $W \times L = 100\,\text{nm} \times 70\,\text{nm}$. The variability for both nMOS in Fig. 6.18 (top) and pMOS in Fig. 6.18 (bottom) behave similar as for *SmartArray A*. The strongest variability increase towards cryogenic temperatures can be seen in $\max(g_m)$. However, there is also a variability increase in SS and in I_{ON}, as can be seen in the quantile plots in the main diagonal. As observed before for *SmartArray A*, there is a strong negative correlation between I_{ON} and V_{th,g_m}, which can again be explained by the definition of a constant $V_G = 0.9\,\text{V}$. The strong positive correlation between I_{ON} and $\max(g_m)$ is caused by the strong dependence on the mobility of both parameters. Towards cryogenic temperatures the correlations decrease slightly, which may be explained by relatively stronger contact resistance. The same observations can be made for the variability of the pMOS devices with $W \times L = 100\,\text{nm} \times 70\,\text{nm}$.

The same overall trends can be observed in the mismatch plots for nMOS in Fig. 6.19 (top) and pMOS in Fig. 6.19 (bottom). The parameter mismatch, which is defined as the difference in a certain parameter of neighbored transistors shows the strongest variability increase in $\Delta\max(g_m)$. There is also a strong variability for the lowest measured temperature $T = 4.2\,\text{K}$ in ΔSS. The strong correlations between ΔI_{ON} and $\Delta\max(g_m)$ and between ΔI_{ON} and $\Delta V_{th,g_m}$ have been observed before in the correlation plots and are caused by the same mechanisms.

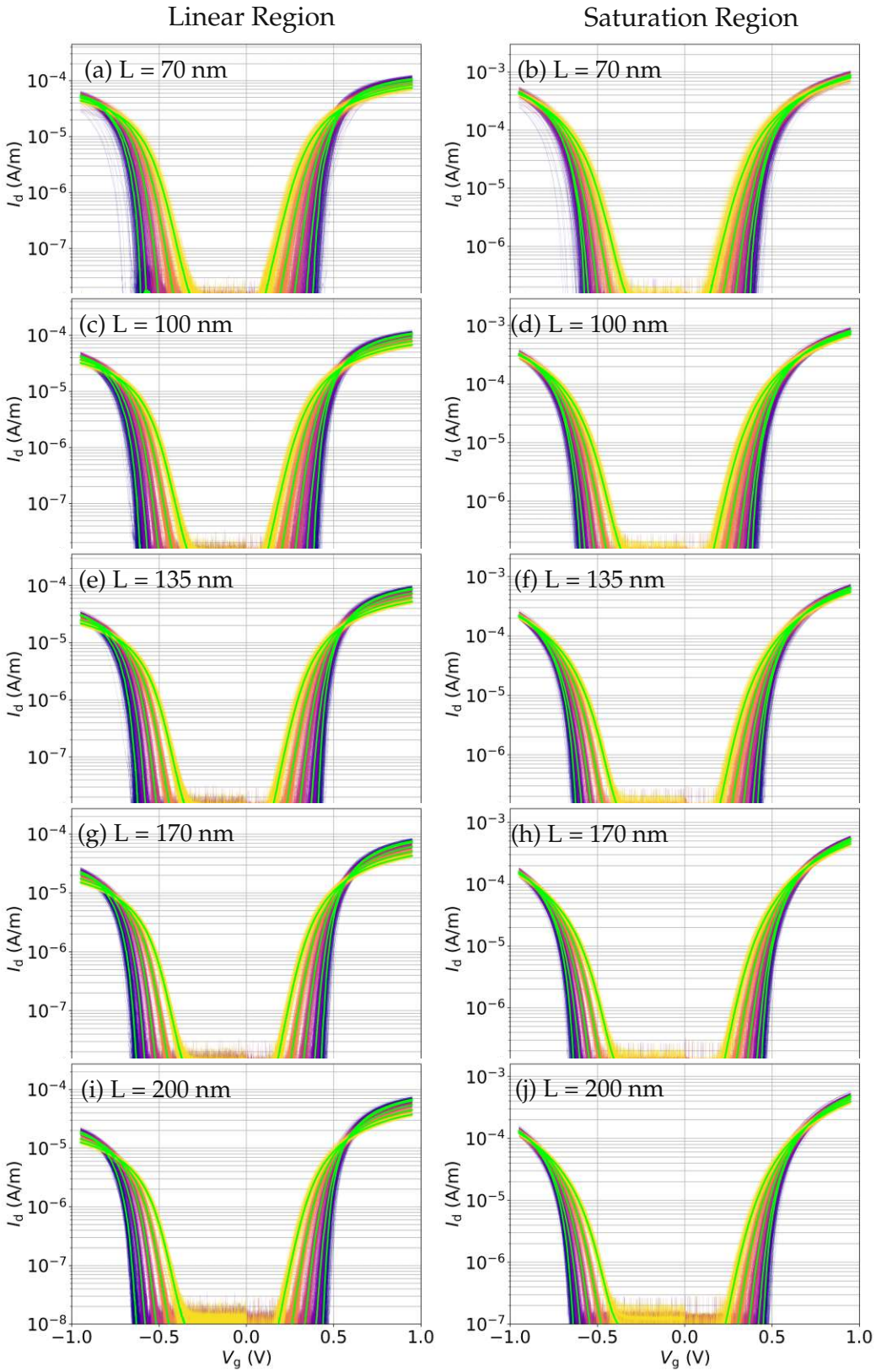

Figure 6.16. The measured transition curves on *SmartArray B* for $W = 100\,\text{nm}$ and increasing lengths in the linear (a,c,e,g,i) and saturation (b,d,f,h,j) region from $L = 70\,\text{nm}, 100\,\text{nm}, 135\,\text{nm}, 170\,\text{nm}$ up to $200\,\text{nm}$. With increasing L the absolute V_{th} shifts to higher voltages. The temperature dependence of the set of transition curves behaves in the same way as for large area devices in Section 6.1.

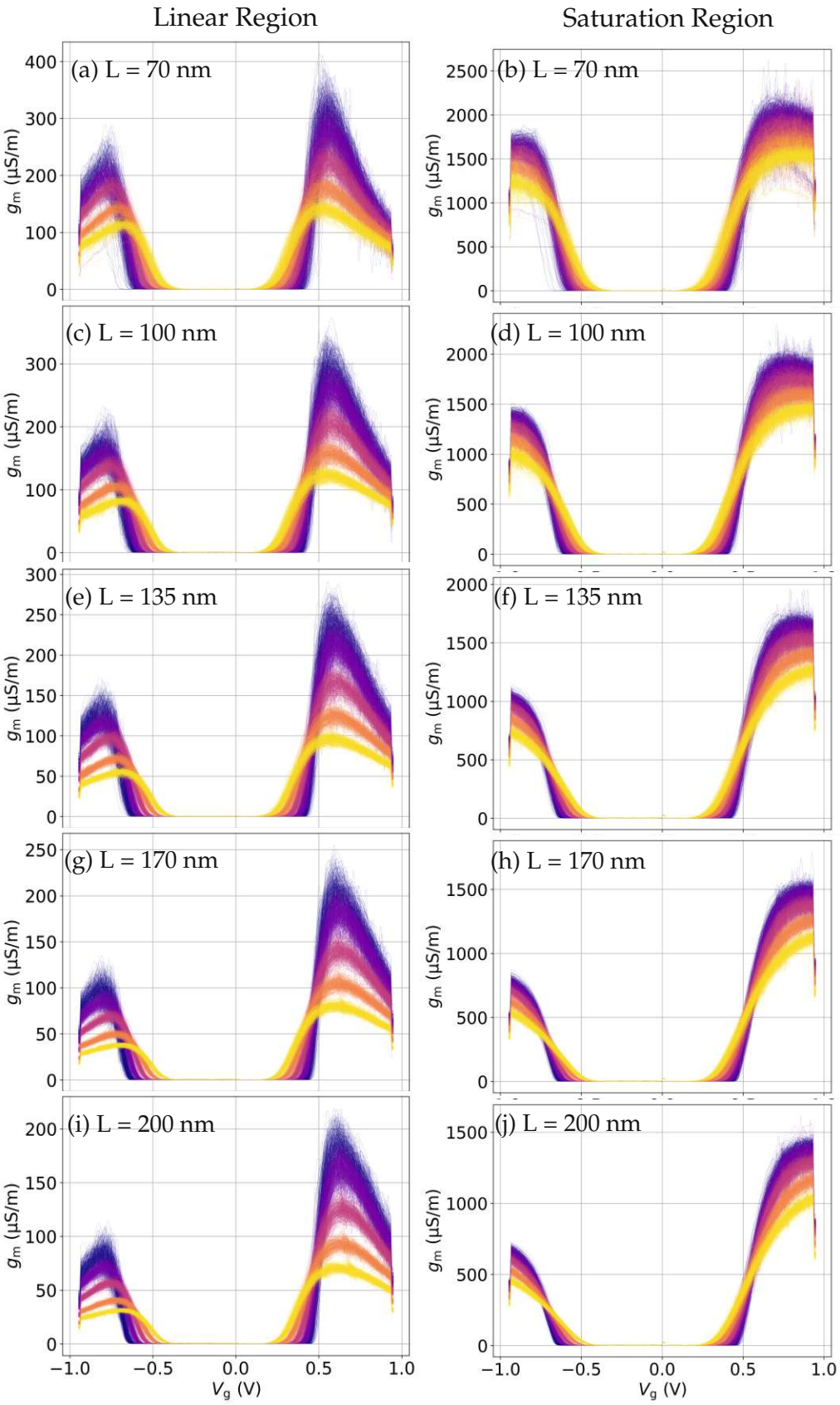

Figure 6.17. The transconductance for various drawn dimensions and temperatures in the linear and saturation region can be derived from the recorded transition curved in Fig. 6.16. g_m clearly shows an increasing variability towards 4.2 K and towards smaller gate lengths.

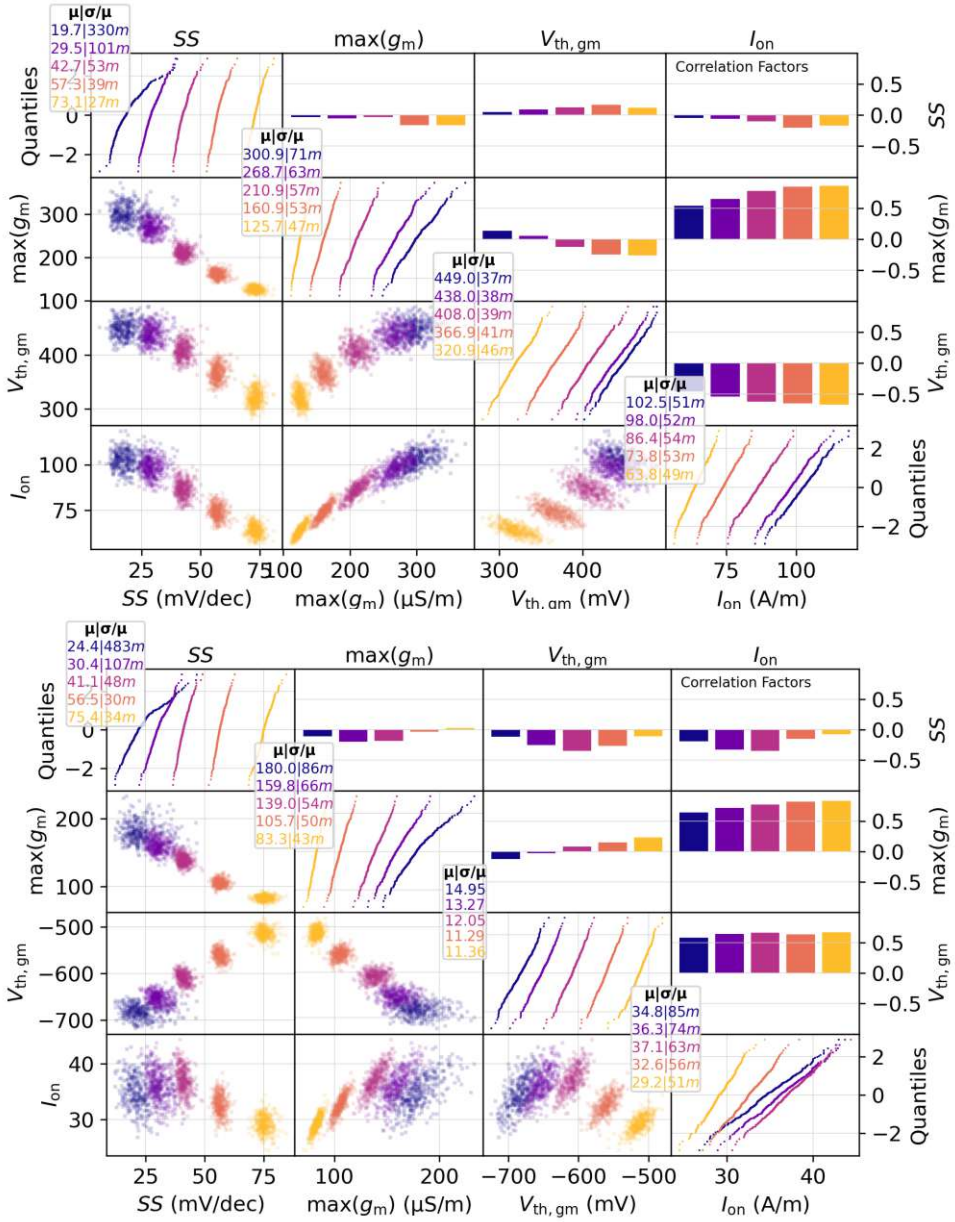

Figure 6.18. The correlation plots for nMOS (top) and pMOS (bottom) with drawn dimensions $W \times L = 100\,\text{nm} \times 70\,\text{nm}$ show an variability increase in $\max(g_\text{m})$ and in SS towards cryogenic temperatures. A strong negative correlation between I_ON and V_{th,g_m} is caused by the definition of I_ON at a constant V_G. The strong positive correlation between I_ON and $\max(g_\text{m})$ occurs because both depend directly on the mobility. The correlations decrease towards 4.2 K because the relative impact of the contact resistance increases.

The recorded transition curves in Fig. 6.16 allow not only the analysis of the variability but also of the impact of dimension and temperature variation on the mean SS, $\max(g_\text{m})$, V_{th,g_m}, and I_ON as it can be seen in Fig. 6.20 and 6.21 for nMOS and pMOS, respectively. In both, the linear region and the saturation region the parameters listed earlier have been extracted for transition curves on all 500 available transistors per geometry on *SmartArray B*. This has been done for $T = 4.2\,\text{K}, 77\,\text{K}, 150\,\text{K}, 225\,\text{K}$ and 300 K.

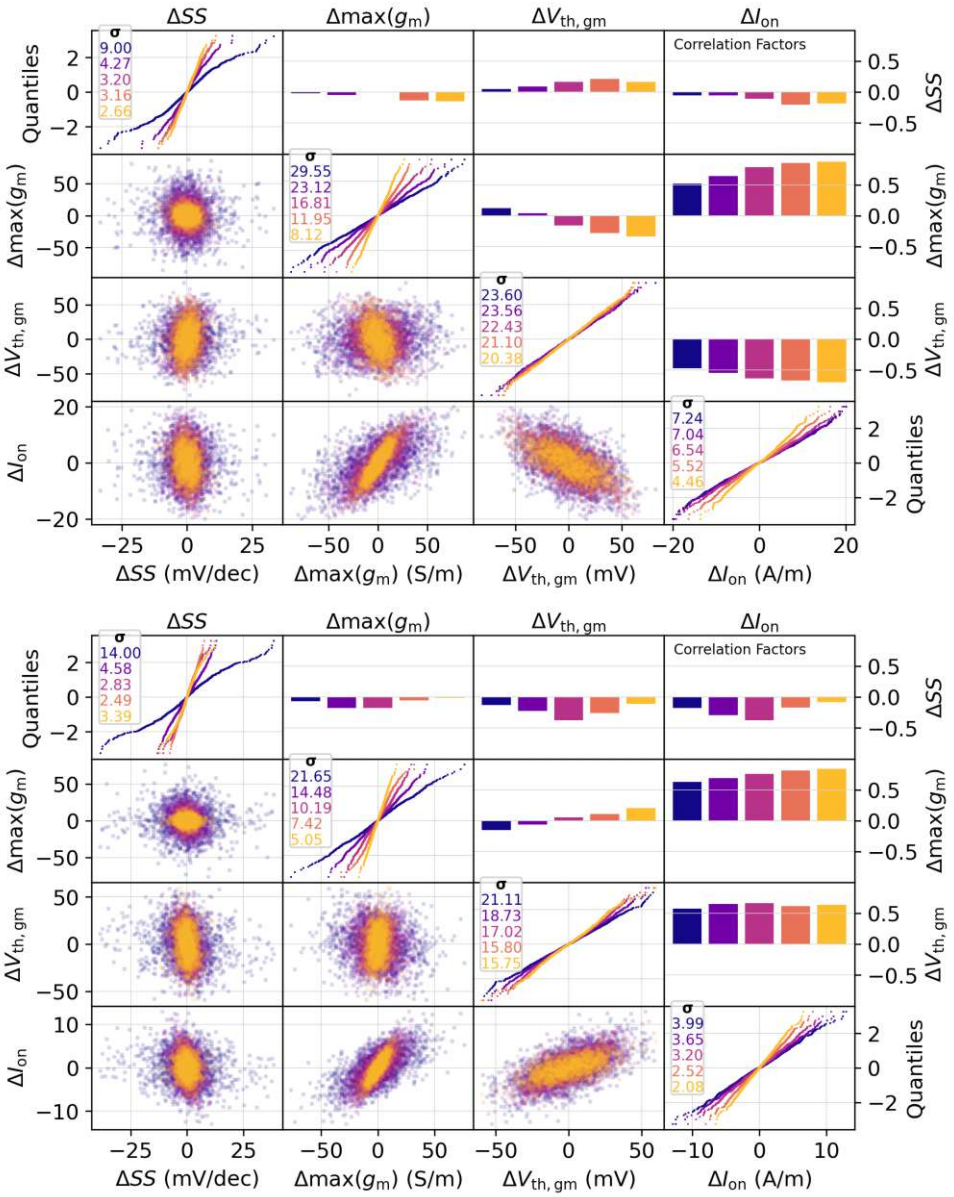

Figure 6.19. The corresponding mismatch plots to the correlation plots in Fig. 6.18 show the same overall trends for nMOS (top) and pMOS (bottom). There is a strong variability increase of max(g_m) towards cryogenic temperatures and strong correlation betweens I_{ON} and V_{th,g_m} and between I_{ON} and max(g_m) which decrease towards 4.2 K.

The extracted mean values are represented with a confidence interval of one-sigma. The increasing confidence intervals towards cryogenic temperatures for max(g_m) are in agreement with the correlation plots in Fig. 6.18.

The trends of the mean values are in agreement with the trends measured on large area devices in Section 6.1 and on *SmartArray A*. As expected, mean max(g_m) and mean I_{ON} increase for nMOS and decrease on pMOS towards cryogenic temperatures due to the increasing mobility. max(g_m) shows a saturation in the saturation region while I_{ON} shows a saturation for both V_D conditions. Smaller gate lengths result in a larger max(g_m) and I_{ON} as expected from basic MOSFET theory. The well studied

SS-saturation caused by band tail states is more distinct for the pMOS devices than for the nMOS devices, specially in the linear region. Conspicuously, *SS* shows a small dependence on the geometry on pMOS unlike the nMOS devices, on which the drawn dimensions have almost no influence. $V_{\mathrm{th,g_m}}$ increases on nMOS and decreases on pMOS towards 4.2 K. The saturation effect which is specially observed on the nMOS devices can be explained by band tail states [111]. As expected from theory, shorter channels lead to a smaller threshold voltage [205].

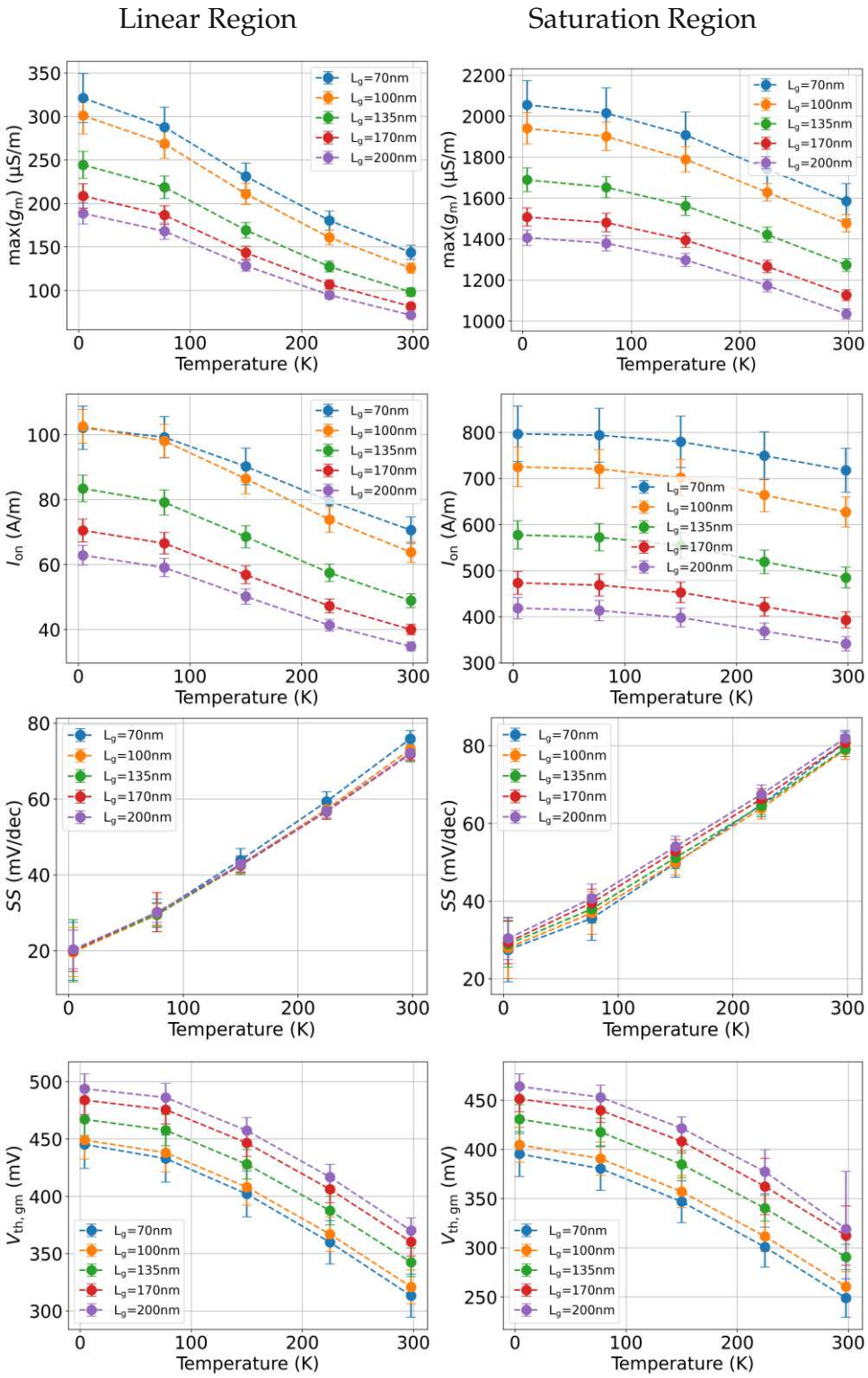

Figure 6.20. From the transfer curves from Fig. 6.16 the extracted max(g_m), I_{ON}, SS and V_{th,g_m} for nMOS in both the linear (left) and saturation region (right) are shown here. The confidence interval of one-sigma shows a large variability increase in max(g_m) towards 4.2 K. Mean parameters show the same behavior known from large area devices discussed in Section 6.1: $|\langle \max(g_m) \rangle|$, $|\langle I_{ON} \rangle|$, and $|\langle V_{th,g_m} \rangle|$ increase towards cryogenic temperatures due to increasing mobility and shifting Fermi level and $\langle SS \rangle$ shows the typical saturation caused by band tail states [111].

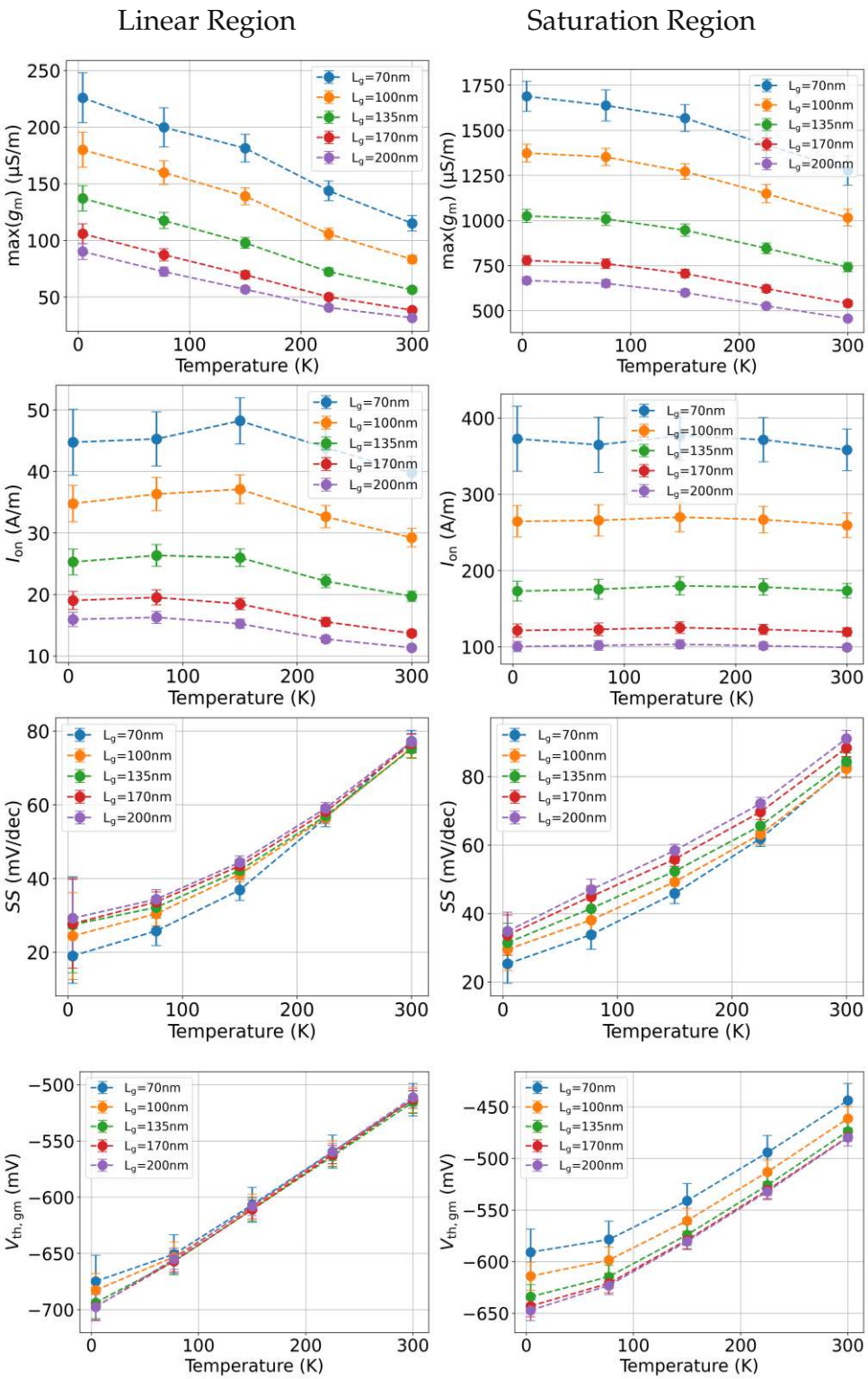

Figure 6.21. Mean $\max(g_\mathrm{m})$, I_ON, SS and $V_{\mathrm{th},g_\mathrm{m}}$ for pMOS in both the linear (left) and saturation (right) region are shown. The mean parameters show the same trends which are well known from large area devices and have been discussed in detail in Section 6.1.

Chapter 7

Charge Noise Characterization

Charge noise is inevitable during device operation and thus can also be seen in electrical characterization. It occurs in different forms on both large area devices and scaled devices. On scaled devices, there are discrete steps in the measured drain-source current I_D, which correspond to charge capture and emission events of single defects. On large area devices, thousands of such charge capture and emission events take place simultaneously. Therefore in large devices discrete steps are not visible anymore, however, the superposition of the defects can be observed as 1/f noise.

In the following first the experimental characterization and the theoretical models describing charge noise in the frequency domain are introduced. Afterwards, algorithms for detecting and extracting the step heights and capture and emission time distributions are explained before measured RTN at cryogenic temperatures will be analyzed for different technologies.

7.1 Experimental Characterization

Experimentally, charge noise characterization on a MOSFET is straight-forward: A constant voltage is applied at the gate of the DUT, and the drain-source current is recorded for a defined period using an equidistant sampling scheme. Using an initially measured $I_D(V_G)$ curve, the current signal can be mapped to an equivalent ΔV_{th} signal. Active charge traps can capture and emit charges which can be seen as discrete steps in the recorded current and the corresponding ΔV_{th} curve. This is shown on four exemplary traces in Fig. 7.1 (left).

Such traces can be recorded for various gate voltages and temperature conditions. This allows to determine the dependence of different parameters such as mean charge capture and emission times or mean step heights on V_G and T. By varying the sampling frequency from high to low frequencies, defects with high and low capture and emission rates can be accessed.

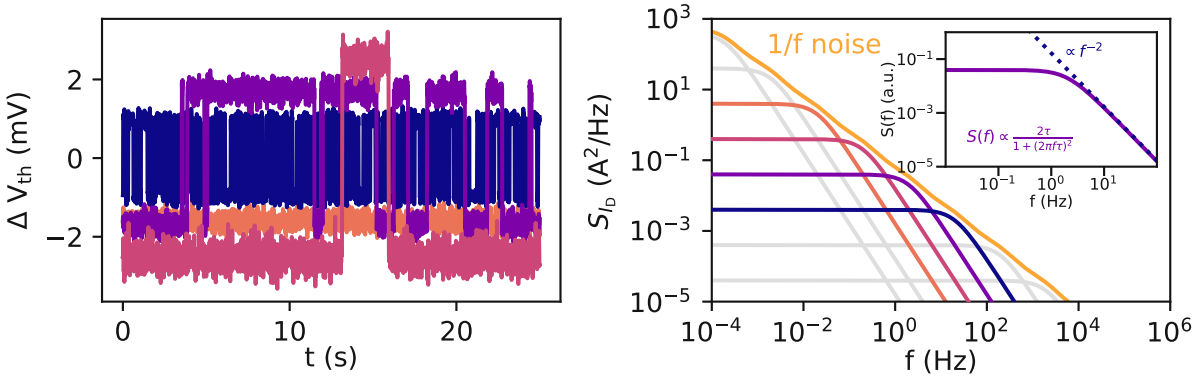

Figure 7.1. Four exemplary RTN signals show different mean charge capture and emission times and step heights. From these RTN signals the power spectral density (represented by the same color as the corresponding RTN signal) can be obtained, which follows a Lorentzian distribution. The corner frequency of the Lorentzian corresponds to the mean capture/emission time τ of the RTN signal. The superposition of uniformly distributed Lorentzian signals follows a $1/f$ behavior, which is well known from large area devices.

7.2 1/f Noise

Based on the measured drain-source signals or the corresponding ΔV_{th} traces in the time domain, the power spectral densities (PSDs) in the frequency domain can be obtained using for example the Welch algorithm [230] or Bartlett's method [231]. The PSD of a two-state RTN signal follows a Lorentzian, as can be seen in Fig. 7.1 (right). The Lorentzian of defect i can be described using

$$S_i(f) = \frac{(2d_i\tau_{0,i})^2}{(\tau_{\text{c},i} + \tau_{\text{e},i})(1 + (2\pi f\tau_{0,i})^2)}, \tag{7.1}$$

where d_i is the step height and $\tau_{0,i} = 1/(1/\tau_{\text{c},i} + 1/\tau_{\text{e},i})$. The superposition of multiple electrically active defects results in the superposition of the Lorentzian spectra

$$S(f) = \sum_i S_i(f) = \sum_i \frac{(2d_i\tau_{0,i})^2}{(\tau_{\text{c},i} + \tau_{\text{e},i})(1 + (2\pi f\tau_{0,i})^2)}. \tag{7.2}$$

As shown schematically in Fig. 7.1, on large area devices the Lorentzian spectra are uniformly distributed in $\log(\tau)$. When considering only a subset of defects in a small time interval, the approximation $\tau_{\text{e}} = \tau_{\text{c}} = \tau_i$ holds. Using the equivalent step height $d_{\text{e}} = d_i$ for all defects i, allows to simplify the expression to

$$S(f) \approx \frac{d_{\text{e}}^2}{2} \sum_i \frac{\tau_i}{1 + (\pi f\tau_i)^2}. \tag{7.3}$$

Due to the characteristic corners of the Lorentzian PSDs, a single defect always dominates at a certain frequency. For the time constant of this dominating defect $\tau_i = 1/(\pi f)$ holds at the corner frequency πf. This allows considering only the dominant defect,

leading to

$$S(f) \approx \frac{d_e^2}{4\pi f}. \tag{7.4}$$

This final result shows how the superposition of Lorentzian PSDs stemming from single defects results in a 1/f signal which is observed on large area devices.

These general considerations about connections between RTN signals caused by single defects and 1/f noise in large area devices can be related to the empirical relation between the spectral density of the drain-source current S_{I_D} and the number of charge carriers N in the channel proposed by Hooge [232]

$$\frac{S_{I_D}(f)}{I_D^2} = \frac{\alpha_H}{fN} \tag{7.5}$$

with α_H being the empirical Hooge parameter. Based on this empirical model, a unified flicker noise model [233] for MOSFETs was developed which describes the drain current PSD for a device with dimensions W and L as

$$S_{I_D}(f) = \frac{k_B T I_D^2}{\gamma f W L}\left(\frac{1}{N} + \alpha\mu\right)^2 N_t(E_{F,n}) \tag{7.6}$$

with γ being a WKB tunneling factor for the interface traps following a trap distribution $N_t(E)$, which is evaluated at the channel quasi Fermi level $E_{F,n}$. N is the charge carrier density, μ the carrier mobility and $\alpha = \pi m_e q^3/16\varepsilon^2 h k_B T$ with the average dielectric constant $\varepsilon = (\varepsilon_{Si} + \varepsilon_{Ox})/2$. From this equation the corresponding gate voltage power spectral density can be derived using

$$S_{V_G}(f) = \frac{S_{I_D}(f)}{g_m^2}. \tag{7.7}$$

resulting in

$$S_{V_G}(f) = \frac{k_B T q^2}{\gamma f W L C_{ox}^2}\left(1 + \alpha\mu N\right)^2 N_t(E_{F,n}). \tag{7.8}$$

These models are well established at room temperature, however, the linear relation in T, classical mobility models for μ, and the simplified evaluation of the trap density at the quasi Fermi level fail at cryogenic temperatures and cannot explain the increasing noise levels at 4.2 K compared to room temperature as shown in Fig. 7.2.

While the spectral noise density of I_D increases towards 4.2 K, the spectral noise density of V_G is rather temperature independent. However, both show the occurrence of Lorentzian features at cryogenic temperatures which can be attributed to single defects. This shows the importance of understanding the role of single defects at cryogenic

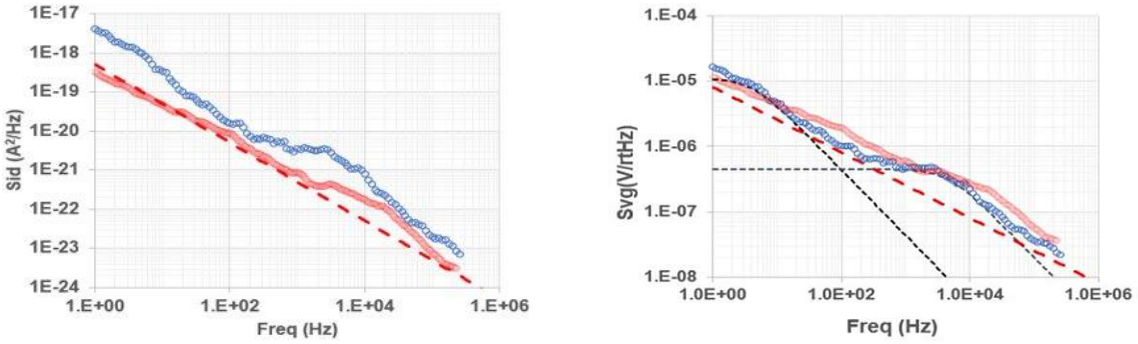

Figure 7.2. The spectral noise density of I_D (left) at 4.2 K (blue circles) increases compared to room temperature (red circles), while the spectral noise density of V_G (right) stays constant. Both noise spectra show the occurrence of Lorentzian corners at cryogenic temperatures, which can be attributed to single defects. The red dashed line indicates 1/f behavior and the black dashed lines show Lorentzian corners in the spectrum. Figures taken from [234].

temperatures and the necessity of a charge transition model being valid in the cryogenic regime.

7.3 Random Telegraph Noise

On small area devices single charge capture and emission events are observed in the measured drain-source current or the corresponding V_{th} curve as discrete steps. Every defect can therefore be identified by $\{\eta, \tau_c, \tau_e\}$, where η is the mean step height (either in the current or the equivalent ΔV_{th} representation) and τ_c and τ_e are the mean capture and emission times, respectively. This set of parameters can be found for various gate voltage and temperature conditions $\{\eta, \tau_c, \tau_e\}(V_G, T)$. It has to be noted that the step heights and the charge transition times are stochastic quantities. The step height is normally distributed around a mean value. As the charge transition times can be described as a Markov process, it can be shown using the Master equation that capture and emission times are exponentially distributed [174], as can be seen for the exemplary trace in Fig. 7.3.

For the extraction of the distributions of step heights, time constants and the corresponding set of mean parameters $\{\eta, \tau_c, \tau_e\}$ which identifies a defect, a precise detection of the position of the discrete steps is crucial. Since there should be ideally tens or hundreds of discrete steps in a single RTN trace, and typically multiple traces for different temperatures and gate voltage conditions are needed, an automatic or semi-automatic detection algorithm of the position and heights of the steps is required. In this work, two different algorithms were used: the *Canny edge detector* and *Otsu's method* which will be presented in the following sections.

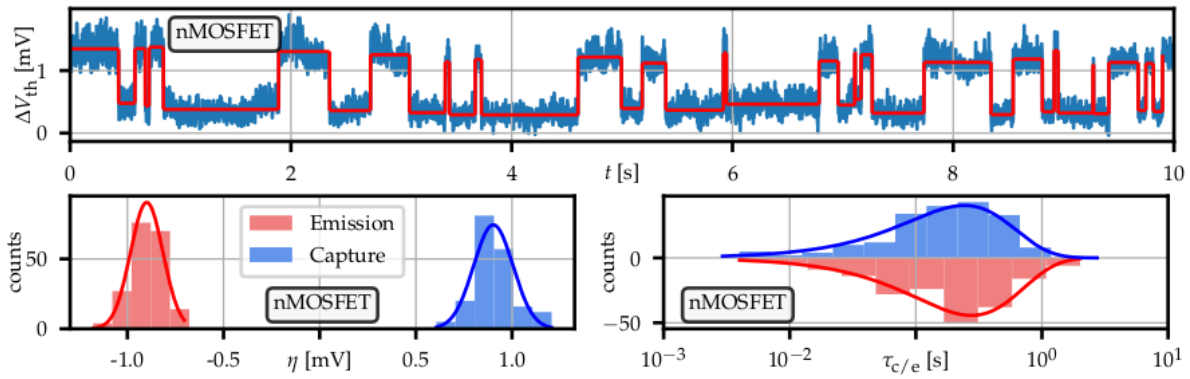

Figure 7.3. After detecting the steps of an arbitrary RTN signal (top) various parameters such as the step heights or charge capture and emission times can be extracted. Based on this it can be seen that step heights of capture and emission are normally distributed around the same absolute value. Capture and emission times are exponentially distributed shown here using a logarithmic x-axis.

7.3.1 Canny Edge Detector

The Canny edge detector was developed by John F. Canny in 1986 and was initially designed to detect edges in images [235]. While the original algorithm was developed for 2-dimensional applications, the step detection in RTN signals is a 1-dimensional problem which simplifies the algorithm considerably. Analogous to the deviation in [140], the algorithm can be broken down into the following steps, shown in Fig. 7.4:

1. The intensity of the measured RTN signal S increases when convoluting it with a discretized first derivative of a Gaussian filter G

$$R[n] = (S * G)[n] \sum_{k=-W}^{W} S[k]G'[n-k], \tag{7.9}$$

where W is the width of the truncation and

$$G'[t] = -\frac{t}{\sqrt{2\pi}\sigma^3} \mathrm{e}^{-\frac{t^2}{2\sigma^2}} \tag{7.10}$$

with σ being the width of the Gaussian filter.

2. The convuluted signal $R[n]$ can be simplified by applying a non-maximum suppression

$$R_{\mathrm{m}}[n] = \begin{cases} R[n], & (|R[n]| > R[n+1] \wedge |R[n]| > R[n-1]) \\ 0, & \text{else} \end{cases} . \tag{7.11}$$

3. A threshold value R_{th} is chosen, which suppresses responses from noise:

$$R_{\text{t}}[n] = \begin{cases} R[n], & (|R_m[n]| > R_{\text{th}}) \\ 0, & \text{else} \end{cases} . \tag{7.12}$$

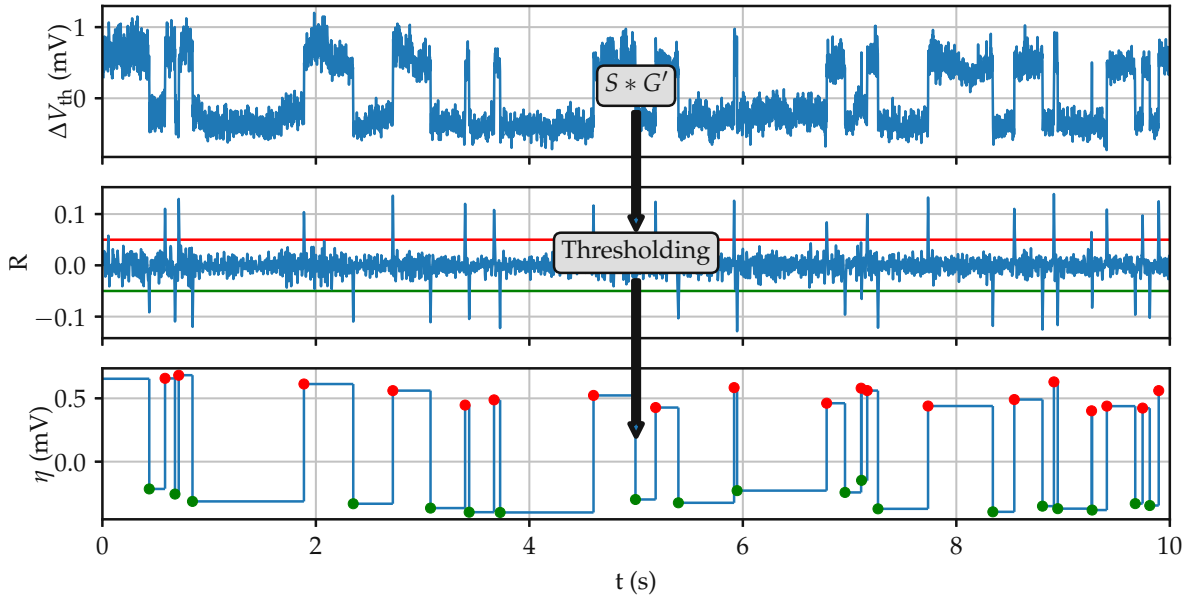

Figure 7.4. The Canny edge detector can be broken down into the following steps: An initial RTN signal S (top panel) is folded with the derivative of a Gaussian filter G'. The resulting signal R (central panel) shows distinctive peaks. The signal R is simplified by a non-maximum suppression before setting a threshold value (red/green). Peaks with an absolute value above the threshold are detected as steps (bottom panel), the sign of the step allows to differ between up-steps and down-steps which correspond to charge capture and emission events.

The sign of $R_{\text{t}}[n]$ determines whether an upwards or downwards step is detected, and the height of the step can be extracted from the original signal. This algorithm has a low error rate and the localization of the edges is very precise. If there are two active defects in a single signal, the detected steps can be clustered by their step heights using for example a K-means algorithm [236]. However, the algorithm has the disadvantage that the width of the Gaussian filter σ and the threshold value R_{th} are input parameters which have be to varied to detect different active defects. Thus the detection quality depends on the selected parameters. However, when dealing with a large set of RTN signals, a fully automated algorithm for step detection is necessary, like for example Otsu's method which is presented in the next section.

7.3.2 Otsu's Method

Otsu's method is an algorithm for automatic thresholding and was developed by Nobuyuki Otsu in 1979 [237]. Just like the Canny edge detector, Otsu's method

was initially designed for image thresholding, thus for separating pixels into two classes: foreground and background. The algorithm searches for a threshold value which maximizes the inter-class variance of the two classes 0 and 1 equivalent to minimizing the inter-class variance. This process is performed in four steps:

1. The normalized histogram and the probability $p(i)$ for every bin i of L bins in total is computed. The normalized histogram for the signal in Fig. 7.5 (left) can be seen in Fig. 7.5 (right).

2. The class probabilities $\omega_{0,1}(t=0)$ and class means $\mu_{0,1}(t=0)$ are initialized for $t=0$.

3. Next, the algorithm steps through all thresholds $t=1,\ldots,L-1$ and computes the class probabilities $\omega_{0,1}(t)$, the class means $\mu_{0,1}(t)$ and the inter-class variances $\sigma_b^2(t)$. Here, the class probabilities are defined as

$$\begin{aligned}\omega_0(t) &= \sum_{i=0}^{t-1} p(i) \text{ and} \\ \omega_1(t) &= \sum_{i=t}^{L-1} p(i).\end{aligned} \tag{7.13}$$

   Using the class probabilities, the class means $\mu_{0,1}(t)$ are obtain as

$$\begin{aligned}\mu_0(t) &= \frac{1}{\omega_0(t)} \sum_{i=0}^{t-1} ip(i) \\ \mu_1(t) &= \frac{1}{\omega_1(t)} \sum_{i=t}^{L-1} ip(i)\end{aligned} \tag{7.14}$$

   which allows to compute the inter-class variance

$$\sigma_b^2(t) = \omega_0(t)\omega_1(t)[\mu_0(t) - \mu_1(t)]^2. \tag{7.15}$$

   In Fig. 7.5 (right) the inter-class variance is plotted for every threshold value.

4. Finally, the optimum threshold value is determined as the threshold t for which the respective inter-class variance is maximized.

After obtaining the threshold t with Otsu's method, a step in the signal S at timestep i is given at those points, where

$$\Big(S(i-1) < t \wedge S(i) > t\Big) \vee \Big(S(i-1) > t \wedge S(i) < t\Big) \tag{7.16}$$

holds.

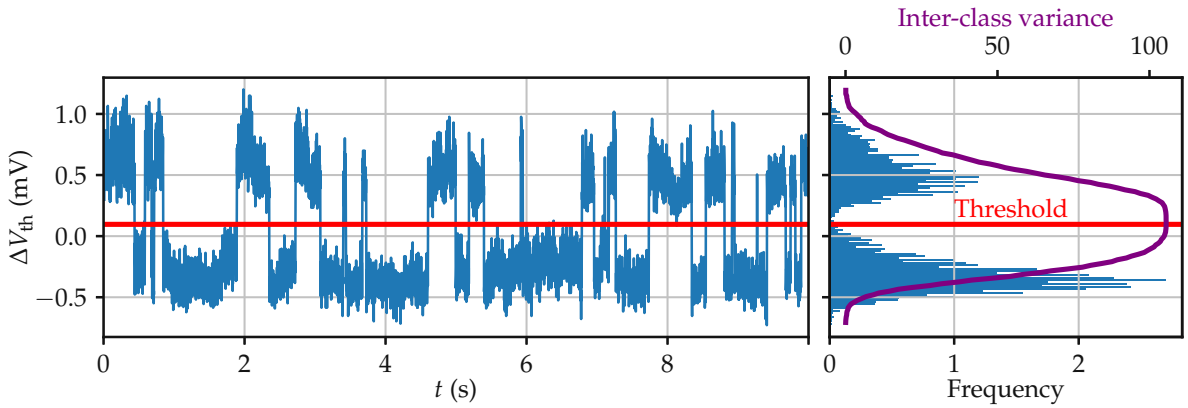

Figure 7.5. For a RTN signal (left) a normalized histogram (right) with L bins is computed. For every bin, the inter-class variance σ_b^2 can be calculated and the bin with the maximum σ_b^2 gives the best threshold value. The positions where the signal switches from above to below the threshold value can be extracted and corresponding step heights and time constants are obtained.

Compared to the Canny edge detector, Otsu's method has the advantage that the thresholding is completely automated. In general, it can be beneficial to apply a filter before the automated thresholding to suppress measurement noise. While this has the disadvantage that very short time constants get lost, it renders the thresholding extremely robust and even allows the detection of small step heights. In this work, the denoising algorithm proposed by *Chambolle* was used [238].

Otsu's method can only be applied on measurement data including a single active defect. In case of the detection of multiple active defects (multi-level RTN) typically only the most prominent defect signal will be extracted. To avoid incorrect detection additional sanity checks are applied, e.g. a Kolmogorov-Smirnov test for testing if capture and emission times follow an exponential distribution. For a detailed analysis of multi-level RTN more advanced techniques as a factorial hidden Markov model analysis are needed [239].

7.4 RTN in SiON Devices

In the following section RTN measurements on *Tech. B* are presented. While for the time zero characterization large area devices with $W \times L = 10 \times 10\,\mu m^2$ have been used, scaled devices with $W \times L = 0.22 \times 0.18\,\mu m^2$ are used for the presented noise analysis. A set of both nMOS and pMOS devices with this geometry has been scanned to find transistors with active RTN defects. In total 3 nMOS devices (defect A-C) and 1 pMOS device (defect D) with RTN signals which could be followed over a range of gate voltages and temperatures have been found. The analysis of the measured signals was performed using the Canny edge detector discussed in Section 7.3.1.

7.4.1 Temperature and Gate Voltage Variation

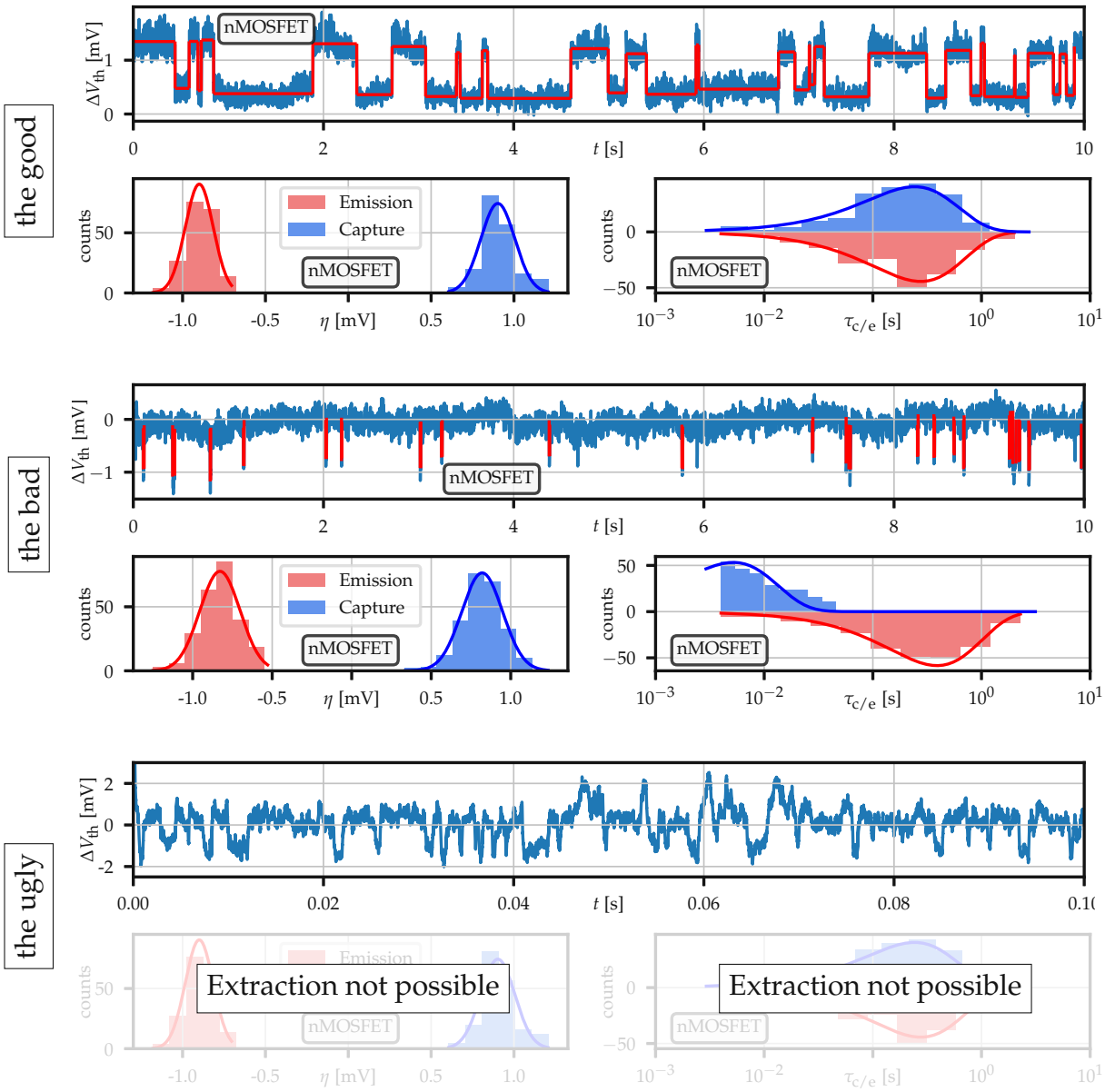

Figure 7.6. The recorded RTN signal at $T = 9\,\mathrm{K}$ and $V_G = 557.5\,\mathrm{mV}$ (top, *the good*) allows a step extraction using the Canny edge detector and the extraction of the normally distributed step heights and the exponentially distributed capture and emission times. Moving to a higher $V_G = 582.5\,\mathrm{mV}$ (center, *the bad*) results in very small capture times. The extraction of the step heights and the time constants is still possible, however, the distribution of τ_c is truncated due to the limitation of the measurement window. At a higher temperature $T = 44\,\mathrm{K}$ (bottom, *the ugly*) the remains of the RTN signal can still be observed, however, a step detection is not possible anymore. Figures taken from [MJC4].

An ideal measurement can be seen in Fig. 7.6 (*the good*) for a nMOS device at $T = 9\,\mathrm{K}$ and $V_G = 557.5\,\mathrm{mV}$. The steps can be easily detected with the Canny edge detector, as indicated with the red lines. The extraction of the step heights shows two Gaussian distributions ($\overline{\eta} = 0.9\,\mathrm{mV}$, $\sigma_\eta = 0.1\,\mathrm{mV}$) and the capture and emission times

are exponentially distributed. However, these ideal conditions are only observed for a rather small measurement window. When sweeping over a large set of gate voltages, for higher and lower voltages the charge transition times become very small or very large. This has the effect that the signal shows extremely narrow peaks as can be seen in Fig. 7.6 (*the bad*), again at $T = 9\,\mathrm{K}$ but at a higher gate voltage of $V_\mathrm{G} = 582.5\,\mathrm{mV}$. These sharp peaks still allow a reasonable extraction of the step heights, as can be seen in the Gaussian distributions ($\overline{\eta} = 0.9\,\mathrm{mV}$, $\sigma_\eta = 0.15\,\mathrm{mV}$). However, the extraction of the capture times τ_c is limited by the measurement resolution and thus the exponential distribution shows a cut off. While these truncated signals still allow to extract useful parameters, this becomes difficult at higher temperatures. At $T = 44\,\mathrm{K}$ in Fig. 7.6 (*the ugly*) one can clearly see the remains of the clear RTN signal seen before, however, a parameter extraction has become impossible. Nevertheless, across a certain temperature and voltage range it is possible to follow the defects and to extract step heights, capture and emission times.

The scan across temperatures and gate voltages can be seen in Fig. 7.7. At every set of (T, V_G) it is possible to extract a capture time τ_c (red) and an emission time τ_e (blue), as can be seen for $T = 4, 9, 14, 19, 24, 29\,\mathrm{K}$ for V_G between 0.54 V and 0.58 V. The capture time shows a strong gate voltage dependence, whereas the emission time is almost V_G independent, as known from other technologies [114], however, both can be fitted linearly on the log-scale. These linear fits for τ_c and τ_e intersect at the gate voltage where $\tau_\mathrm{c} = \tau_\mathrm{e}$. At this V_G the trap level is aligned with the semiconductor Fermi level and it will be used later in this section to study the T-dependence of the corresponding transition times.

Extraction of Step Heights

As shown in the previous section, the distribution of step heights can be extracted for every signal at a certain T and V_G. By extracting the means of these typically Gaussian distributions it is possible to find the T and V_G dependence of the mean step heights $\overline{\eta}$. The gate voltage dependence of $\overline{\eta}$ for different temperatures can be seen for defect A and defect D in Fig. 7.8. For nMOS, the mean step heights decrease with increasing V_G. This can be attributed to the fact that the channel uniformity increases with increasing V_G and thus the impact of a single charge on the drain-source current decreases [240]. This mechanism is temperature independent, as can be seen in the figure. The same observation holds for pMOS, where with decreasing V_G the step heights decrease, as can be seen for defect D in Fig. 7.8.

Using the extracted τ_c, τ_e for different V_G in Fig. 7.7, it is possible to compute the interpolated $\overline{\eta}$ at the point where $\tau_\mathrm{c} = \tau_\mathrm{e}$. This can be done for a range of temperatures to obtain the T dependence of $\overline{\eta}$. Again, it can be seen that the step height is nearly temperature independent in this cryogenic temperature range.

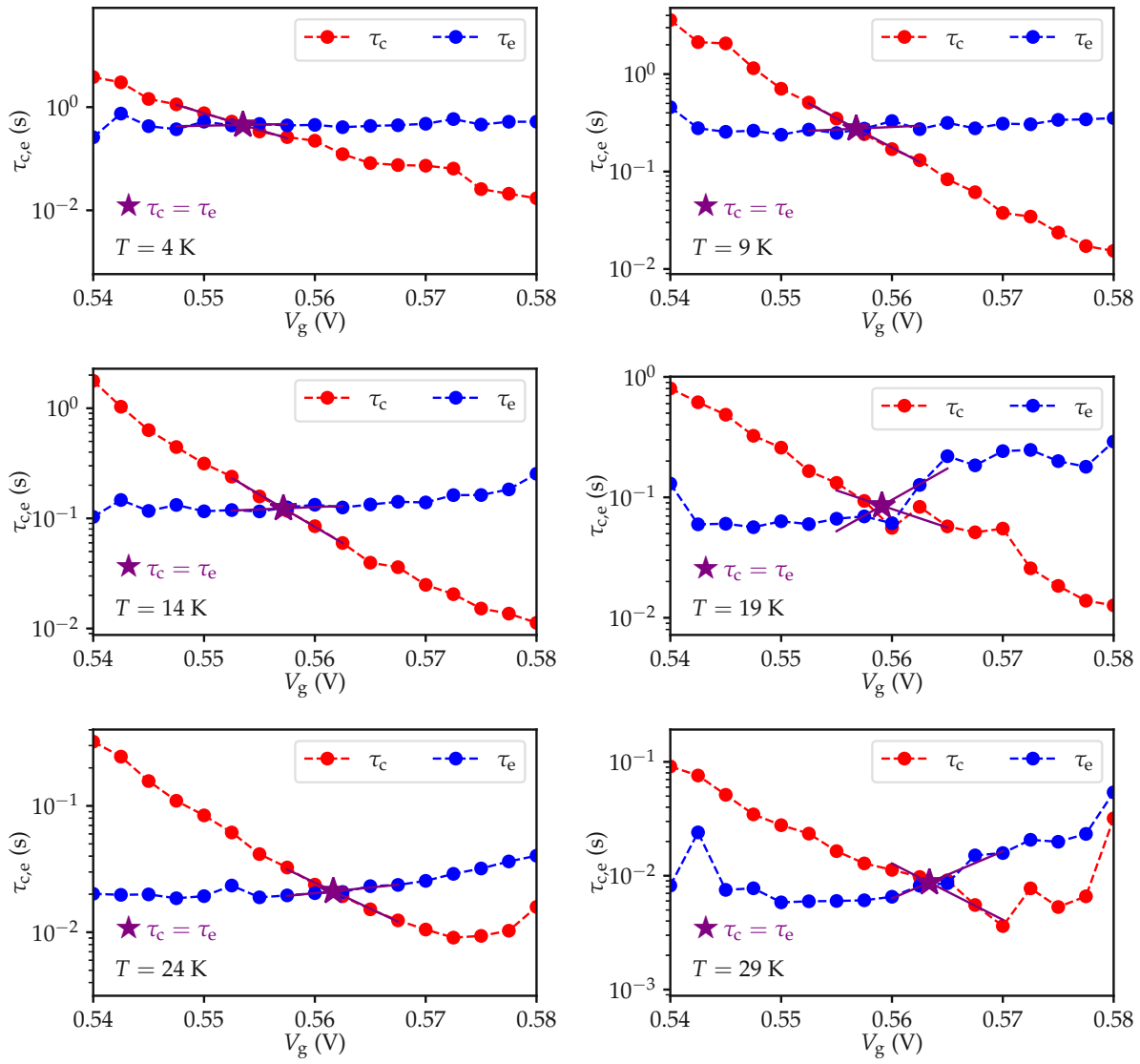

Figure 7.7. Mean capture and emission time extracted for $T = 4\,\mathrm{K}, 9\,\mathrm{K}, 14\,\mathrm{K}, 19\,\mathrm{K}, 24\,\mathrm{K}$ and 29 K and gate voltages between $V_G = 540\,\mathrm{mV}$ and 580 mV. While τ_e is almost V_G independent, τ_c shows a strong, exponential V_G dependence. Both τ_c and τ_e can be fitted linearly (on the log-scale) and the intersection point $\tau_c = \tau_e$ can be found (purple stars) for every temperature. At this point, the Fermi level is aligned with the trap level.

Extraction of Average Capture and Emission Times

For the analysis of the T and V_G dependence of the average capture and emission times τ_c and τ_e, again the extractions as shown in Fig. 7.7 are used. Now, in Fig. 7.9 the average times are shown in a more compact way with all measured temperatures in a single plot for defect A and D. Within this single plot it can be clearly seen that the measured τ_c and τ_e curves shift towards faster capture and emission times with increasing temperatures.

The marked stars in Fig. 7.9 where $\tau_e = \tau_c$ holds can be plotted in an Arrhenius-plot, see Fig. 7.10. It can be seen that the charge transition times become temperature inde-

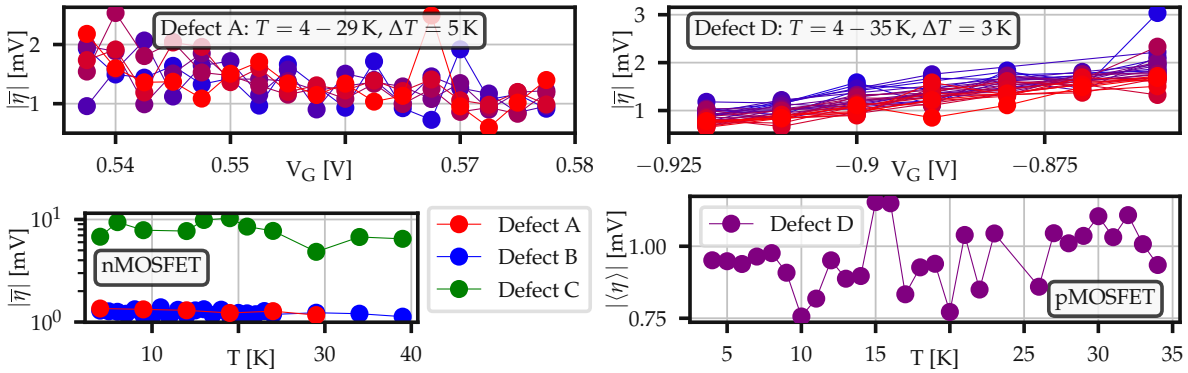

Figure 7.8. The step heights of defects A-D show a weak dependence on the applied V_G and T. Defect A on nMOS (upper left) shows that with increasing V_G the step heights decrease slightly due to the higher channel uniformity and the stronger impact of the drain-source current on a single charge. On pMOS devices the same effect appears with decreasing V_G (upper right). While V_G affects $\overline{\eta}$, the average step heights are almost unaffected by the temperature change, as can be seen in the lower panels. Figures taken from [MJC4].

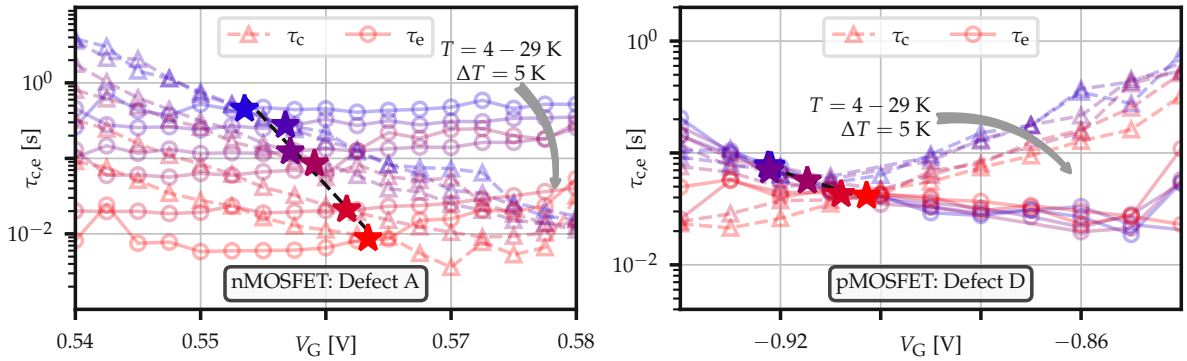

Figure 7.9. Capture and emission times for defect A and D can be analyzed in the same way as shown in Fig. 7.7 but presented in a more compact way with the whole temperature set in a single figure. This representation shows that mean capture and emission times get smaller with increasing temperatures, leading to higher charge transition rates. Figures taken from [MJC4].

pendent at cryogenic temperatures in both nMOS and pMOS devices. This temperature independent regime can be modeled using *Comphy* with the implemented quantum mechanical transition rates, as discussed in detail in Chapter 3. The solid lines, representing the simulation, show that nuclear tunneling explains the temperature independent behavior of charge trapping at cryogenic temperatures, whereas the classical model (dashed lines) freezes out completely. Nuclear tunneling at cryogenic temperatures is dominated by the overlap of the ground vibrational states of the initial and final states. These ground states are occupied, even at 4 K, and enable charge transitions.

The simulation allows to find a pre-existing defect with a parameter set of trap level E_T, relaxation energy E_R, configuration coordinate offset ΔQ, and position in the oxide

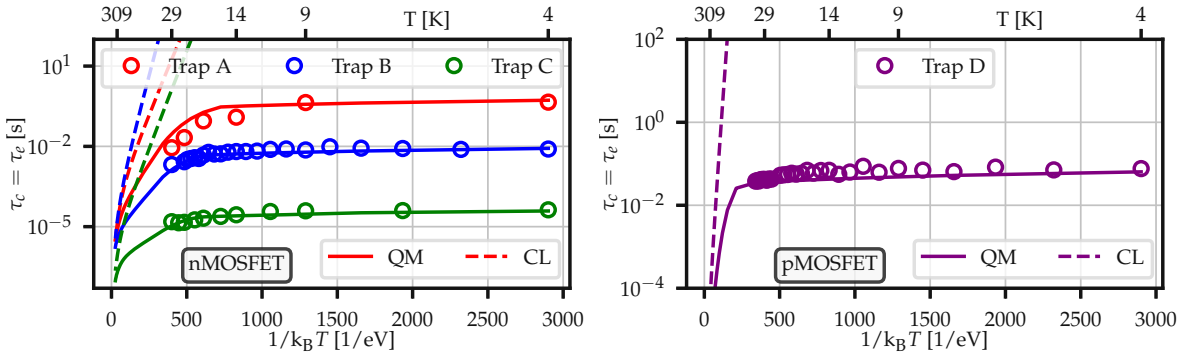

Figure 7.10. The extracted charge transition times at the point where $\tau_c = \tau_e$ becomes temperature independent at cryogenic temperatures for all characterized defects instead of an Arrhenius-like behavior. This can be explained by nuclear tunneling, which is dominated by the overlap of the vibrational ground states, which are the only occupied states at cryogenic temperatures. This can be accurately modeled using the full quantum mechanical 2-state NMP theory (solid lines), while the classical approximation (dashed lines) freezes out completely. Figures taken from [MJC4].

x_T which describes the measured temperature dependence of the charge transition times of the RTN signal. The extracted parameters for every defect are listed in Tab. 7.1.

Defect	Device	E_T [eV]	E_R [meV]	x_T [nm]	ΔQ [$\sqrt{u}$Å]	E_B [meV]
Defect A	nMOS	0.44	110	0.2	3.5	28
Defect B	nMOS	0.48	95	0.3	3.0	24
Defect C	nMOS	0.46	60	0.05	2.6	15
Defect D	pMOS	-0.48	100	0.25	3.0	25

Table 7.1. The extracted temperature dependence of the charge transition times allows to extract trap levels, relaxation energies, spacial positions and configuration coordinate displacements of defect candidates for the measured RTN signals.

The extracted spatial and energetic position are shown in Fig. 7.11. As can be seen, the electrically active defects on nMOS (left) are close to the conduction band edge. The active defect in the pMOS device, on the other hand, is close to the valence band edge of the substrate. This energetic alignment guarantees charge transition times with $\tau_c = \tau_e$. The spatial position of all defects is less than 0.5 nm from the Si/SiO_2 interface. This, together with the trap level alignment strongly indicates that the defects responsible for RTN in cryogenic environments are interface defects.

This idea can further be supported by DFT calculations. The regions marked as *oxide defects* in Fig. 7.12 are extracted from [MJJ2]. The extracted E_T and E_R parameters are in good agreement for possible oxide defect candidates [MJJ2]. The ellipsoid formed region marked as *interface defects* within the Si bandgap is extracted from [142]. It shows that interface defects are expected to have low relaxation energies compared to oxide defects. The colored circles represent Defect A to D discussed in this section. As can be

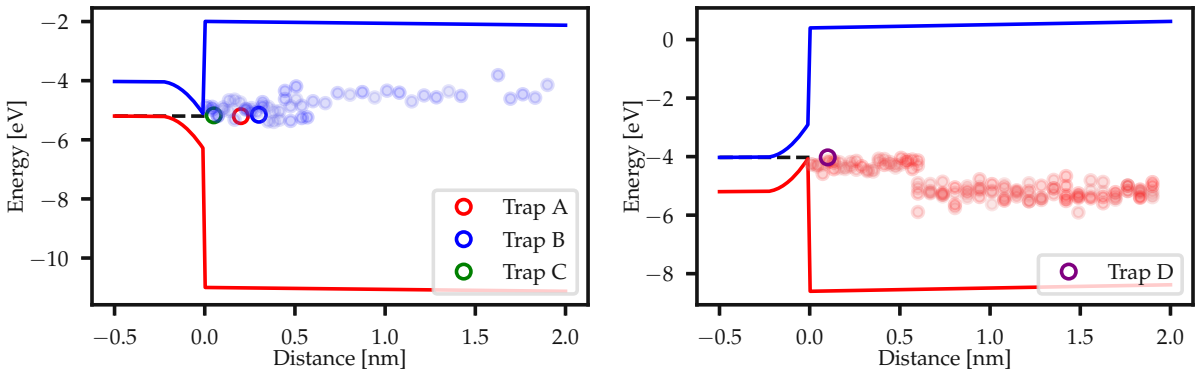

Figure 7.11. The band diagrams show the spatial and energetic position of the extracted RTN defects. As can be seen, the defects measured on nMOS are very close to the conduction band edge, the defect on pMOS close to the valence band edge. Spatially the defects are very close to the interface between the oxide and the Si substrate. This spatial position together with the extracted trap levels and relaxation energies strongly indicate that the defects responsible for RTN at cryogenic temperatures are interface defects. Figures taken from [MJC4].

seen, the extracted values are in good agreement with the DFT calculations for interface defects, which are highlighted in Fig. 7.12.

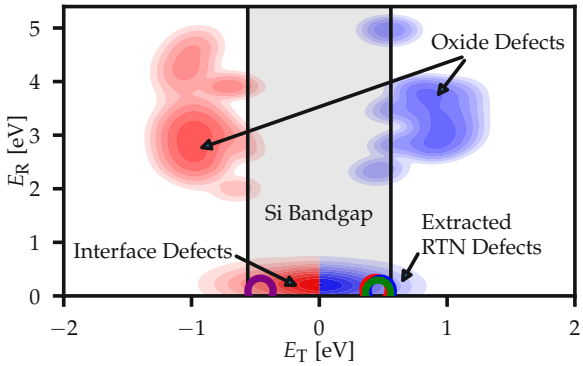

Figure 7.12. The (E_T, E_R)-heatmap shows the active regions for oxide defects, extracted in [MJJ2] and for interface defects, extracted in [142]. Interface defects show smaller relaxation energies compared to oxide defects and are energetically located close to the band edges or in the bandgap. These properties agree with the extracted defect parameters of the RTN defects A to D. Figure taken from [MJC4].

7.5 RTN in MoS_2 MOSFETs

The fact that RTN does not freeze out at cryogenic temperatures can be seen across various channel materials and device geometries. In the following section, RTN has been measured at cryogenic temperatures on *Tech. C* with dimensions $W \times L = 100\,\text{nm} \times 70\,\text{nm}$. For this, RTN signals have been measured within a range of gate voltages for a set of temperatures between 10 K and 75 K at $V_D = 1\,\text{V}$. For every measured RTN trace, the signal was analyzed using the Canny Edge Detector discussed in Section 7.3.1

for detecting charge capture and emission events and extracting the distribution of the corresponding capture and emission times and the step heights. This can be seen for a few examples in Fig. 7.13 for $T = 25\,\mathrm{V}$ at $V_\mathrm{G} = -1.51\,\mathrm{V}$. The left panel shows an RTN signal with two active traps, one having a mean step height of $\overline{\Delta I} = 25\,\mathrm{nA/\mu m}$ and the other of $\overline{\Delta I} = 12\,\mathrm{nA/\mu m}$. For every measured trace, the extracted step heights can be plotted against their capture and emission times, as can be seen in the central panel. The detected charge capture and emission events can be clustered by their step heights using a K-means algorithm [236]. The two clusters with two different mean step heights are visualized by two different colors, see Fig. 7.13 (central panel). This data can then be conglomerated to a histogram for every defect showing the exponential distribution of the time constants as can be seen in the right panel.

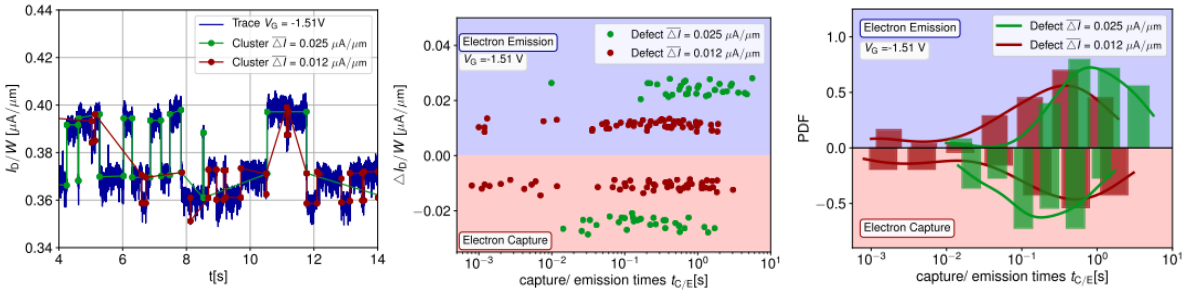

Figure 7.13. Steps on a 100 s RTN signal (left) recorded on *Tech. C* at $T = 25\,\mathrm{K}$ with $V_\mathrm{G} = -1.51\,\mathrm{V}$ and $V_\mathrm{D} = 1\,\mathrm{V}$ have been extracted using the Canny Edge Detector. Two active defects with mean step heights of $\overline{\Delta I} = 25\,\mathrm{nA/\mu m}$ and $\overline{\Delta I} = 12\,\mathrm{nA/\mu m}$ have been found. The step heights are shown over the capture and emission times in the central figure. Conglomerating this to a histogram shows the expected exponential distributions (right). Figures are reprinted from [241, MJC6].

Repeating this analysis for every bias and temperature condition allows to plot the mean capture and emission times over the gate voltage at every measured T. This can be seen in Fig. 7.14 (left) for $T = 25\,\mathrm{K}, 50\,\mathrm{K}$ and $75\,\mathrm{K}$. As this data set has been measured up to comparatively high temperatures, the intersection point where $\tau_\mathrm{e} = \tau_\mathrm{c}$ holds, and where the trap level is aligned with the Fermi level $E_\mathrm{T} = E_\mathrm{F}$, cannot be determined for the curves at higher temperatures. Thus, the temperature dependence is compared for a constant gate voltage.

As can be seen in Fig. 7.14 (right), the charge transition times at the point where $\tau_\mathrm{c} = \tau_\mathrm{e}$ holds show a clear non-Arrhenius like behavior and become completely temperature independent towards 10 K. This is the same behavior as measured on *Tech. B* in the previous section and which agrees well with the full quantum mechanical transition rate. At deep cryogenic temperatures, the only accessible vibrational state is the ground state. The overlap of the ground state vibrational wave function of the initial state with the vibrational wave function of the final state is temperature independent and so is the corresponding lineshape function leading to a vanishing temperature dependence of the transition rates and the corresponding capture and emission times.

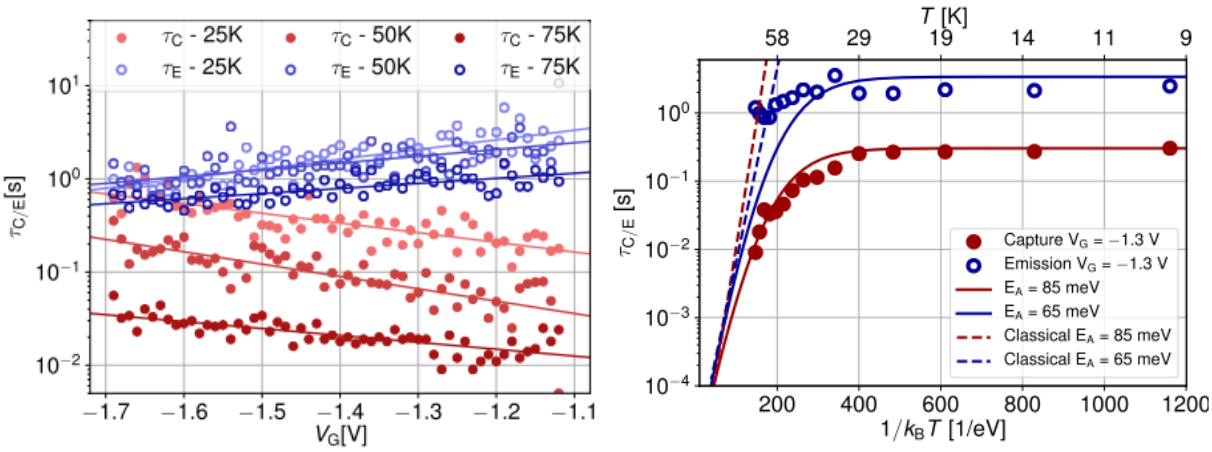

Figure 7.14. RTN signals were measured for a range of gate voltages and temperatures between $T = 10\,\mathrm{K}$ and $T = 80\,\mathrm{K}$. For every measurement, the mean capture and emission time was computed and plotted over V_G, as shown for $T = 25, 50$ and $75\,\mathrm{K}$ in the left panel. The capture and emission times can be plotted in an Arrhenius plot for a constant gate voltage, as it is shown in the right panel for $V_\mathrm{G} = 1.3\,\mathrm{V}$. The mean capture and emission times clearly behave non-Arrhenius like towards cryogenic temperatures. This can be explained with the full quantum mechanical transition rate and the fact that nuclear tunneling dominates in cryogenic charge transitions. Figures are reprinted from [241, MJC6].

7.6 RTN in the Resonant Tunneling Region

As discussed in Section 6.3, $I_\mathrm{D}(V_\mathrm{G})$-curves can show resonant tunneling at low temperatures and for small drain voltages, caused by defects, dopants, impurities, oxide variability or any other fluctuation in the potential which can lead to confinement and thus the formation of a quantum dot (QD). As can be seen in the $I_\mathrm{D}(V_\mathrm{G})$ curve in Fig. 6.9 recorded on *SmartArray B* using *Tech. A* with $W \times L = 100\,\mathrm{nm} \times 70\,\mathrm{nm}$ at 4.2 K, there are devices which show an active RTN defect in the same region where the resonance occurs, see Fig. 7.15. This allows to study the interaction between the active RTN defect and the resonance caused by a QD.

Between $V_\mathrm{G} = 0.3\,\mathrm{V}$ and $V_\mathrm{G} = 0.4\,\mathrm{V}$ RTN traces with lengths of 10 s have been recorded using a step size of $\Delta V_\mathrm{G} = 1\,\mathrm{mV}$. This was done for two drain voltage conditions $V_\mathrm{D} = 10\,\mathrm{mV}$ and $V_\mathrm{D} = 25\,\mathrm{mV}$. The steps have been detected using Otsu's method for every recorded trace and the distribution of the capture time τ_c, the emission time τ_e and the step heights η were extracted, as can be seen for a few examples in Fig. 7.16 for $V_\mathrm{G} = 0.366\,\mathrm{V}$ with $V_\mathrm{D} = 25\,\mathrm{mV}$. The extracted distributions allow to calculate the mean values of τ_c, τ_e and η for every V_G in the measured range.

The calculated mean capture and emission times across the characterized V_G region can be seen in Fig. 7.17. With increasing V_G, τ_c decreases while τ_e is constant. This asymmetry between τ_c and τ_e is well known from various technologies and operational conditions and does not contribute to the resonance in the $I_\mathrm{D}(V_\mathrm{G})$ curve [114]. Also the exponential behavior of the time constants on the gate voltage was shown for various

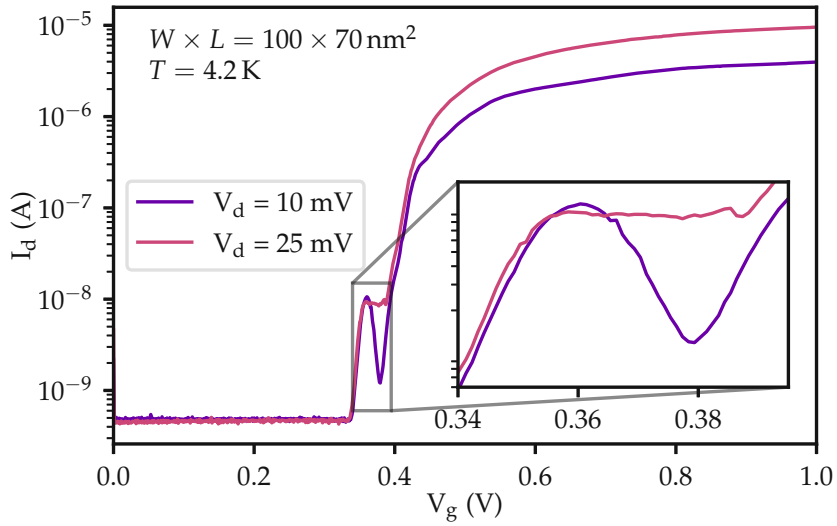

Figure 7.15. Resonant tunneling measured at cryogenic temperatures on an nMOS of *Tech. A* with dimensions $W \times L = 100\,\mathrm{nm} \times 70\,\mathrm{nm}$. For the two selected drain voltages $V_D = 10\,\mathrm{mV}$ and $V_D = 25\,\mathrm{mV}$ the impact of the resonance on an RTN signal characterized in the same voltage region is studied.

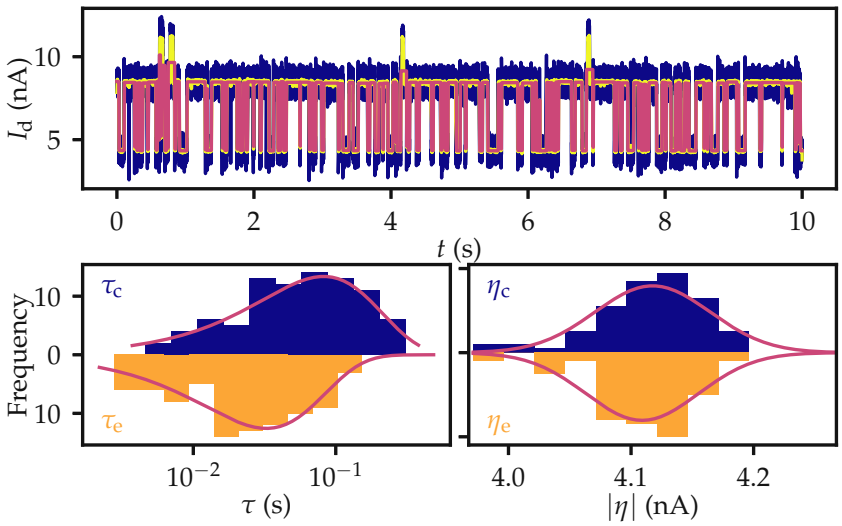

Figure 7.16. An exemplary 10 s RTN trace recorded on *Tech. A* at $V_G = 0.336\,\mathrm{V}$ with $V_D = 25\,\mathrm{mV}$. Using Otsu's method, steps can be detected which allow to extract the distributions of τ_c , τ_e and η. Thus, a mean step height and mean capture and emission times for measurement conditions (V_G, V_D, T) can be computed.

technologies [114, 174, MJC4] under various measurement conditions. Contrary to previous reports [224], the exponential dependence of the charge transition times is not affected by the resonance, which occurs in the same gate voltage region.

This exponential dependence of the charge transition times on the gate voltage agrees with the models derived in NMP theory which are discussed in detail Chapter 3. According to the classical limit of NMP theory, the charge transition rates depend exponentially on the classical transition barrier between neutral and charged state of a trap. When applying V_G to the device, the trap level is shifted. These shifts result

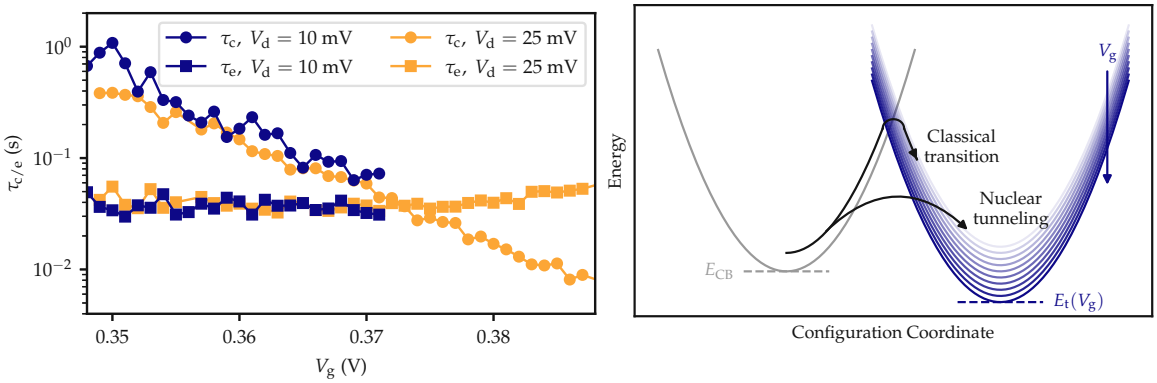

Figure 7.17. The charge transition times of an RTN signal in the resonant tunneling region show an exponential dependence on V_G for both $V_D = 10\,\text{mV}$ and $V_D = 25\,\text{mV}$ (left). This can be explained within NMP theory by the linear dependence of the classical barrier on V_G as can be seen in the right figure.

in a proportional shift of the transition barrier and an exponential V_G-dependence. Describing the V_G-dependence with the quantum mechanical picture is more complex because it involves the vibrational wavefunctions. A shift in V_G leads to a change in the overlap of the vibrational wavefunctions resulting in a complex relation between the transition rate and V_G. However, according to a simplified idea of nuclear tunneling, a quantum mechanical tunneling rate also depends exponentially on the classical barrier. And this classical barrier shows a linear dependence on V_G, resulting in an exponential dependence on V_G. Additionally, Fig. 3.9 shows the dependence of the lineshape function on the energetic difference between the initial and the final state. This energetic offset is proportional to V_G and looks qualitatively the same at cryogenic temperatures as for room temperature, which additionally indicates an exponential dependence of the transition rates on V_G, even at cryogenic temperatures.

Based on NMP theory, the transition rates and the corresponding transition times depend on the lineshape function (which depends on the relative positions of the trap levels) and on the carrier density of states D_n. While the energy levels are not affected by a QD, the D_n would be dramatically different inside of a QD and therefore directly affect the charge transition times. Since this is not the case in the measured mean capture and emission times, it can be concluded that the defect is located outside of the QD as indicated in Fig. 7.18.

While the exponential dependence of the transition times on V_G can be explained within NMP theory, the relation between the step heights η and the QD is more complex. As can be seen in Fig. 7.19 for $V_D = 10\,\text{mV}$ and $V_D = 25\,\text{mV}$ the step height η (blue) is clearly affected by the QD and is thus related to the resonance in the $I_D(V_G)$ curve (red). Often, η is normalized to I_D, as can be seen in the central plots. These normalized step heights look similar to data in previous studies [240], however, a direct comparison is not possible because I_D does not increase monotonically with V_G. Using the initially measured $I_D(V_G)$ curves, η can be mapped from the current domain to an equivalent

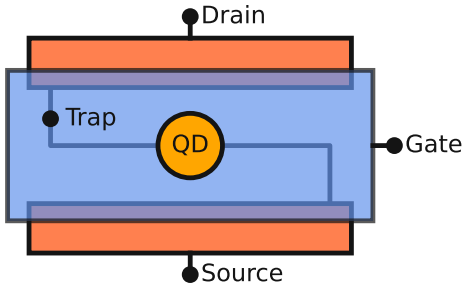

Figure 7.18. Local disorders in the potential cause the formation of a quantum dot (QD) which is responsible for the resonance in the $I_D(V_G)$ curve. A trap can be located anywhere in the oxide. If it is located close to the QD, the changing charge carrier density is expected to have a strong impact on the trap. Charge transition times of a trap located far away from the QD as indicated in this figure are not expected to be affected by the QD.

threshold voltage ΔV_{th} as can be seen in Fig. 7.19 (lower panels). This representation shows that η does not increase or decrease directly proportionally with I_D, which renders a simple explanation of the relation between η and V_G impossible.

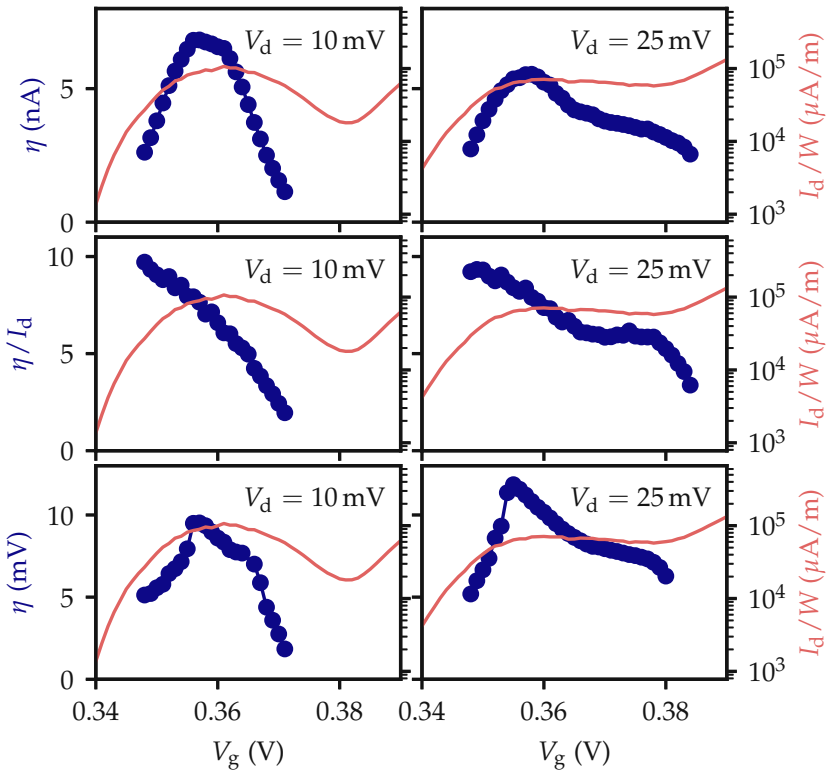

Figure 7.19. The extracted step heights η in the measured currents for both $V_D = 10\,mV$ and $V_D = 25\,mV$ (upper panels) are strongly impacted by the QD, which causes resonant tunneling (red lines). While this dependence is less clear when normalized to I_D as is done frequently in literature [240] (central panels) the strong coupling to the QD can be seen in the mapped equivalent voltage shift step height.

A strong impact of the QD on η has been reported before in [224, 242]. In these studies a uniform potential fluctuation model was used for describing the dependence of η on V_G

$$\eta = \frac{\partial I_D}{\partial \phi}\Delta\phi = \frac{\partial I_D}{\partial V_G}\Delta V_G = g_m \Delta V_G. \tag{7.17}$$

According to this simple model, RTN should vanish at the turning points where $g_m = 0\,S$, which is in contradiction to our measurements. Therefore, a more complex model is needed to describe the relation between source-drain current and η. For this, a precise description of the device electrostatics and transport phenomena is needed, including confinement and the impact of single dopants or other possible sources of the resonance. It might even be necessary to include specific percolation paths in the drain-source current to obtain a conclusive model of the dependence of η on V_G. Developing such a model, however, exceeds the scope of this thesis but might be investigated in future studies.

Chapter 8

Bias Temperature Instability

The effect of bias temperature instability (BTI) occurs, whenever a bias is applied on the gate site of a transistor. The gate bias can reduce or increase the effective energy barrier for a charge transfer reaction between the charge reservoirs and pre-existing or created traps. This process can be accelerated by increasing the temperature, which increases the phonon energy, but also by increasing the applied gate bias. The higher phonon energy leads then to an accelerated energy dissipation of the electron excess energy. This leads to higher charge transition rates via a multiphonon process. BTI degradation has been known since the 1960s [243–245] and has become one of the main reliability concerns ever since. Historically, NBTI on pMOS devices was the largest concern, however, modern nMOS technologies are eventually affected by PBTI [180].

The charges that become trapped during a stress phase of a BTI experiment have an impact on the device electrostatics, leading to a change of V_{th}. This perturbation can be approximated using the charge sheet approximation together with an estimate for the active number of traps [199]. When the stress bias is lowered, the defects eventually release the trapped charge. The threshold voltage drift, which has been accumulated during stress by ΔV_{th} can recover back to the initial $V_{\text{th},0}$ if it is fully recoverable or otherwise to a constant permanent offset added to $V_{\text{th},0}$. However, separating such a permanent and recoverable part experimentally is difficult which makes the analysis and theoretical description complicated. It is widely accepted that the recoverable component is caused by charge captured at defects in the bulk oxide which can be described with the NMP model [174]. The permanent component can be modeled with an empirical double-well model or, as suggested recently, by considering the depassivation of Si-H bonds which leads to the creation of P_{b}-centers at the interface [246, 247].

Experimentally, a common approach for the characterization of BTI is using elevated gate bias and temperature conditions for characterizing the shift of the threshold voltage ΔV_{th}. One widely applied method is the extended Measure-Stress-Measure (eMSM) [248] scheme presented in the following section.

8.1 Extended Measure-Stress-Measure Scheme

For this work, the extended Measure-Stress-Measure (eMSM) scheme [248] is the method of choice for the experimental characterization of BTI. The method consists of three phases and is shown in Fig. 8.1:

- **Phase 1:** First, an $I_D(V_G)$ curve is measured on a pristine device. This $I_D(V_G)$ curve is used as reference for further ΔV_{th} conversion from the measured drain current. For that, it is necessary to identify the V_{th} of the transistor which can be done in different ways as discussed in detail in Section 6.1.3. For the characterization of large area devices, typically a constant current criteria is used for the definition of V_{th}, e.g. $V_{th} = V_G(I_D = 100\,\text{nA} \times W/L)$. For scaled devices on *SmartArrays*, which show a comparable large variability, a constant current criteria is not practical. Instead, V_{th} is defined using $\max(g_m)$. The first phase is shown in Fig. 8.1 marked by $I_D(V_G)$.

- **Phase 2:** After measuring the initial $I_D(V_G)$ curve, the device is stressed at a well defined condition given by a stress gate voltage $V_{G,stress}$ and T for a stress time t_s. These phases are called *stress phases*, and are shown in Fig. 8.1 in the red sections. During stress, traps can capture a charge resulting in a decreased drain-source current and an increasing, corresponding ΔV_{th}, as can be seen in the lower part in Fig. 8.1.

- **Phase 3:** The third phase is called *recovery phase*. It is defined by applying a recovery voltage $V_{G,recovery}$ for a defined recovery period t_r. During that period, typically the source current I_S is measured, because it is more stable against oxide degradation than the drain current I_D [248]. The traps, which have captured a charge during the stress phase, can now emit their charge to the reservoir, thus the absolute value of the recovery voltage decreases, as indicated in Fig. 8.1. Typically, I_D is mapped to ΔV_{th} using the initially measured $I_D(V_G)$. If the device degradation is fully recoverable, then ΔI_D and ΔV_{th} decreases to zero. However, typically a permanent component, which does not decay anymore, is built up and potentially caused by P_b centers [114].

Phase 2 and phase 3 are repeated subsequently, which is done typically with increasing stress and recovery times. On large area devices, the measured recovery curves are smooth as they result from a large number of defects which emit their captured charges continuously, as can be seen in Fig. 8.2 (left) for a commercial 180 nm SiON technology node with drawn dimensions $W \times L = 10 \times 10\,\mu\text{m}^2$ [MJJ2]. On scaled devices it is possible to obtain single capture and emission events as can be seen in Fig. 8.2 (right) on the same technology with drawn dimensions $W \times L = 400\,\text{nm} \times 180\,\text{nm}$ [MJC1]. The measurements on scaled devices build the basis for the time-depend defect spectroscopy (TDDS) [249]. The TDDS is based on the observation that capture and emission times are

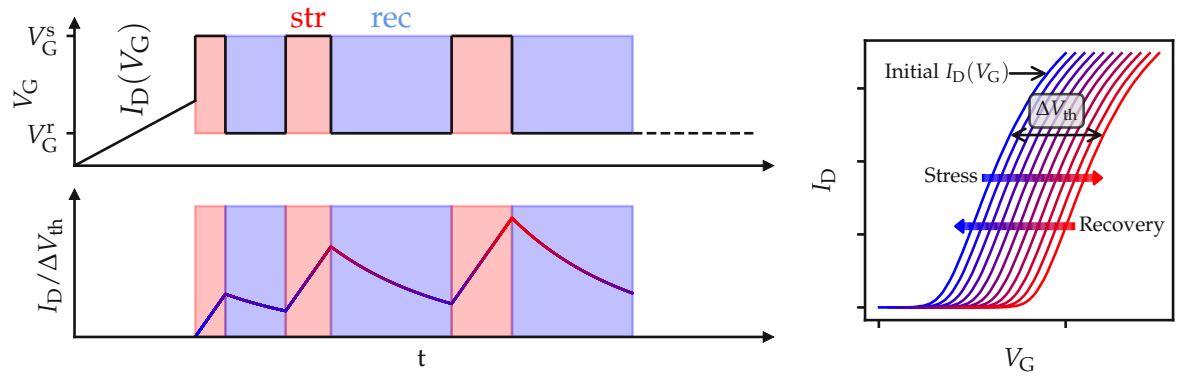

Figure 8.1. The extended Measure-Stress-Measure (eMSM) scheme [248] is used for characterizing BTI. After an initial $I_D(V_G)$, alternating stress (red) and recovery (blue) phases are applied with defined applied gate voltages $V_{G,s/r}$ for a defined period. During recovery, the drain currend I_D is measured and mapped to ΔV_{th} using the initial $I_D(V_G)$. As indicated in the lower plot and the right figure, during stress ΔV_{th} increases, because traps can capture a charge. During recovery, the trapped charge can be released and ΔV_{th} partially recovers back to the initially measured value.

statistically distributed. By measuring a set of recovery curves with the same $V_{G,r}$ and t_r, the extracted set of step heights and associated emission times can be plotted in a scatter map resulting in a so-called spectral map. The resulting clusters correspond to different defects which allow further investigations of different properties, as discussed in detail in Section 8.3 and in [249].

8.2 Temperature Dependence of BTI Measurements

BTI characterization has been performed by applying the eMSM scheme on *Tech. A* and *Tech. B* for a range of temperatures and stress voltages. Following measurement parameters have been used:

- *Tech. A*: For $T \in (4\,\text{K}, 77\,\text{K}, 150\,\text{K}, 225\,\text{K}, 300\,\text{K})$ BTI measurements have been conducted on $W \times L = 10 \times 1\,\mu\text{m}^2$ devices using a stress voltage of $V_{G,s} = 2\,\text{V}$ for nMOS and $V_{G,s} = -2\,\text{V}$ for pMOS devices. Phase 2 and phase 3 in the eMSM measurement scheme have been repeated 5 times using the following stress and recovery times: $(t_s, t_r) \in [(10\,\mu\text{s}, 10\,\text{s}), (1\,\text{ms}, 100\,\text{s}), (100\,\text{ms}, 100\,\text{s}), (10\,\text{s}, 1\,\text{ks}), (1\,\text{ks}, 10\,\text{ks})]$.

- *Tech. B*: For $T \in (4\,\text{K}, 150\,\text{K}, 300\,\text{K})$ BTI measurements have been conducted on $W \times L = 10 \times 10\,\mu\text{m}^2$ devices using a stress voltage of $V_{G,s} \in (\pm 4.51\,\text{V}, \pm 5.1\,\text{V})$ where the positive sign refers to measurements on nMOS and the negative sign to measurements on pMOS devices. Phase 2 and phase 3 in the eMSM measurement scheme have been repeated 5 times using following stress and recovery times: $(t_s, t_r) \in [(10\,\mu\text{s}, 10\,\text{s}), (1\,\text{ms}, 100\,\text{s}), (100\,\text{ms}, 100\,\text{s}), (10\,\text{s}, 1\,\text{ks}), (1\,\text{ks}, 10\,\text{ks})]$.

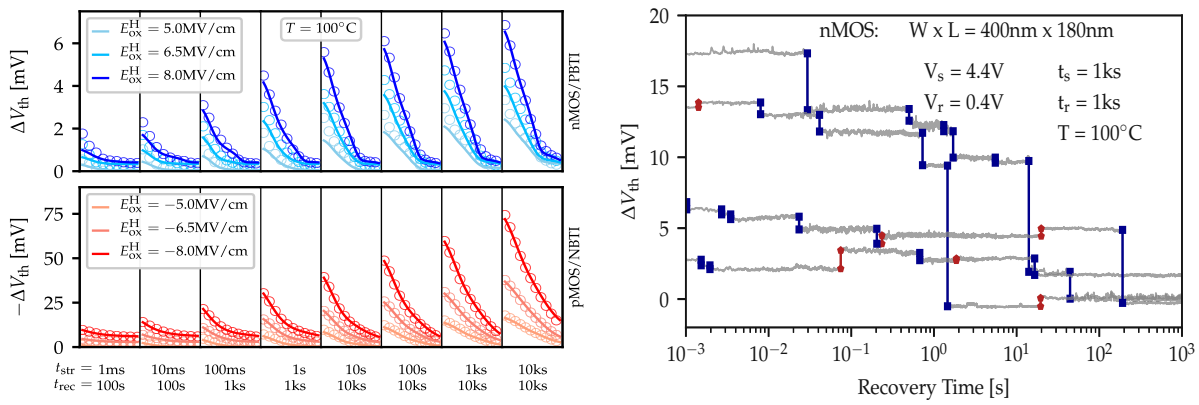

Figure 8.2. The measured drain current I_D is continuous on large area devices, because a large number of defects emit previously captured charges simultaneously and the impact of a single defect on the device behavior is too small to be resolved within measurements, as can be seen in the left panel for a $W \times L = 10 \times 10\,\mu m^2$ SiON device. On a scaled device of the same technology with drawn dimensions $W \times L = 400 \times 180\,nm^2$ discrete steps can be seen in the measured current resulting from single charge emission events as shown in the right panel. Figures taken from [MJJ2] and [MJC1].

The extracted recovery curves for the measurements of *Tech. B* are shown in Fig. 8.3 and have been published in [MJC4]. At room temperature, immediately after stress a ΔV_{th} shift of about $\pm 100\,mV$ has been measured for the longest stress time, where the positive sign refers to PBTI on nMOS and the negative sign to NBTI on pMOS. Towards cryogenic temperatures, both NBTI and PBTI freezes out and only a few mV of threshold voltage shift can be detected, even when the stress time is very long (note the log-scale for ΔV_{th}).

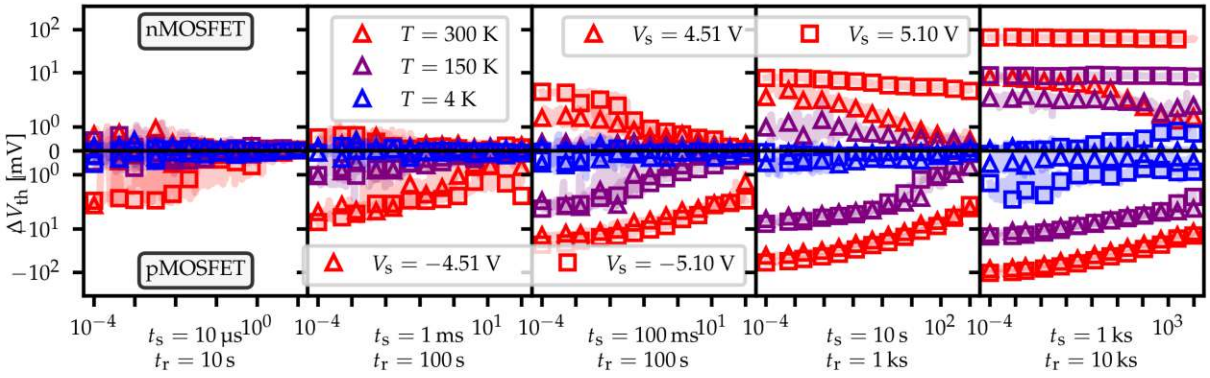

Figure 8.3. Using the eMSM method, PBTI and NBTI trends have been measured on nMOS and pMOS devices for $T = 4, 150, 300\,K$ and two stress voltage conditions $V_{G,s} = \pm 4.51\,V$ and $V_{G,s} = \pm 5.10\,V$. While at room temperature ΔV_{th} is around 100 mV for the longest stress times, BTI freezes out completely towards cryogenic temperatures. Figure taken from [MJC4].

The trap bands for this technology have been extracted by using measurement data obtained at room temperature and above [MJJ2]. The MSM curves shown in Fig. 8.2 (left) have been measured at $T = 100\,°C$ for three different bias conditions on nMOS

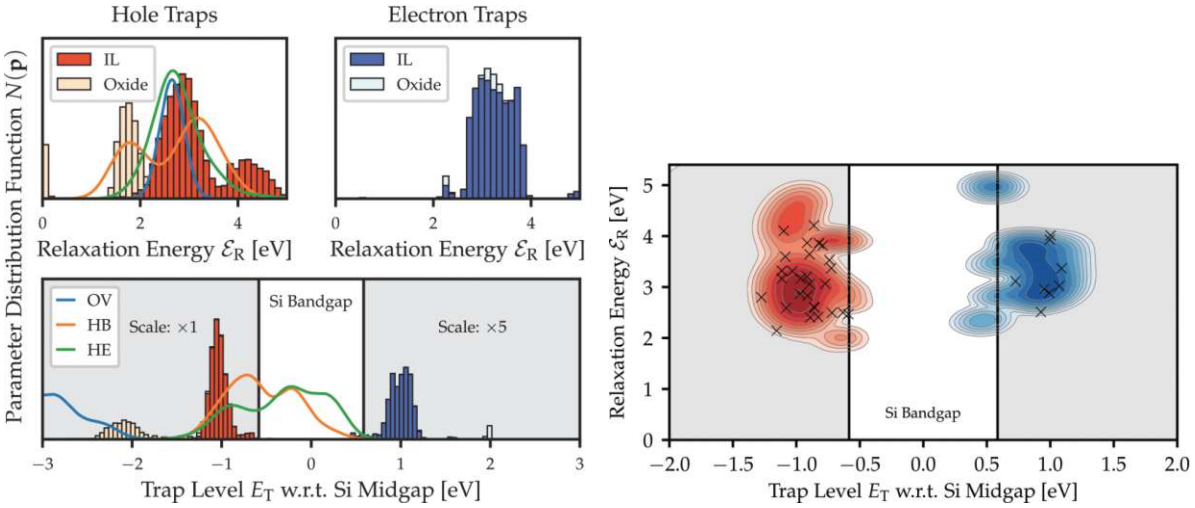

Figure 8.4. The measured ΔV_{th} curves shown in Fig. 8.2 (left) allow the extraction of a interface trap layer and an oxide trap band using the ESiD algorithm. The extracted trap levels and relaxation energies (left) of the IL defects are in good agreement with DFT calculations (solid lines). The (E_T, E_R) heat map (right) for the IL trap bands shows that the extracted parameters are in good agreement with extracted parameters from single defect measurements [250, 251] represented by the black crosses. Figures taken from [MJJ2].

and pMOS and were used to extract the trap bands with *Comphy* employing the ESiD algorithm presented in Section 4.2. By using two independent search regions for the algorithm, two different trap bands representing the oxide and a 0.6 nm thick interface layer (IL) are extracted and shown in Fig. 8.4 (left). Most defects are located within the IL and exhibit trap levels approximately 1 eV below and above the midgap of the Si substrate. The distribution of relaxation energy and trap level is in good agreement with DFT calculations, as can be seen for the hydrogen bridge (HB), hydroxyl-E' (HE) center and oxygen vacancy (OV) represented by the lines in Fig. 8.4. Using this extracted trap bands, the measured recovery behavior in Fig. 8.2 (left) can be replicated by the models as shown by the solid lines. By plotting the relaxation energy against the trap level as in Fig. 8.4 (right), it can be seen that results from earlier single-defect measurements [250, 251] shown by black crosses are in excellent agreement with the extracted measurements.

While the recovery curves in Fig. 8.3 show, that *Tech. B* freezes out completely towards cryogenic temperatures, this is not the case for *Tech. A*. As can be seen in Fig. 8.5 (left), there is a high asymmetry between nMOS and pMOS devices. The points, representing measurement data show that across all stress time conditions at room temperature ΔV_{th} on nMOS is almost twice as large as observed for pMOS. This can also be seen in Fig. 8.5 (right) for a higher temperature of $T = 300$ K. While BTI within the pMOS device freezes out completely towards cryogenic temperatures, the nMOS device still shows a large ΔV_{th}, even at 4 K. At cryogenic temperatures, there is still a ΔV_{th} of more than 20 mV for the longest t_s.

The simulations shown in Fig. 8.5 (left) as solid lines have been perforemed using *Comphy* in combination with the ESiD optimization algorithm and the WKB-based

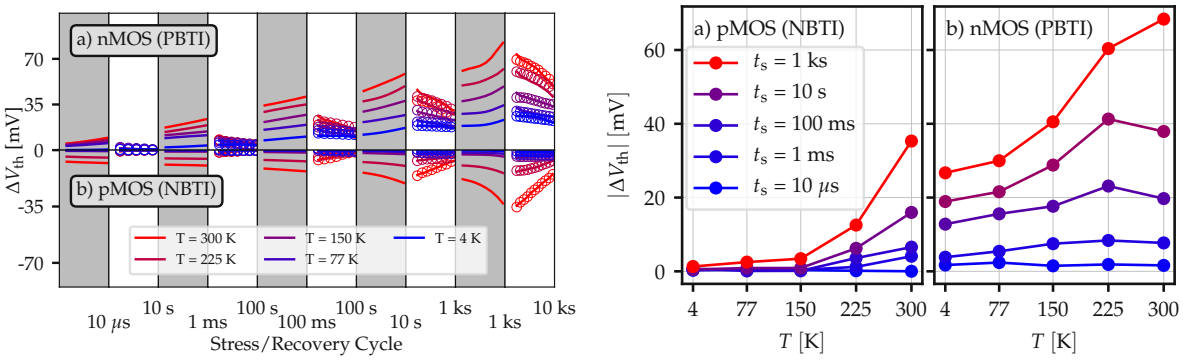

Figure 8.5. *Tech. A* shows a very asymmetric behavior of ΔV_{th} across all measured temperatures. At room temperature ΔV_{th} in the PBTI case on nMOS is twice as large as NBTI on pMOS. The latter freezes out completely below 150 K, while on nMOS there is still a ΔV_{th} of around 40 mV. Even at 4 K nMOS clearly does not freeze out but shows a relatively large ΔV_{th} shift of more than 20 mV. Figures taken from [MJJ5].

approximation of the full quantum mechanical 2-state NMP transition model. For this, a grid of relaxation energies E_R and trap levels E_T has been spanned. This has been done for two independent trap bands for both, nMOS and pMOS (in total 4 independent bands), one trap band represents defects at the interface layer in the SiO_2 and one trap band represents defects in the HfO_2. Additionally, the grid evolves across the spatial position x_T and the configuration coordinate offset ΔQ. For these two grids, the average response was computed while the least square fit has been done across E_T and E_R. The full set of parameters can be seen in Tab. 8.1.

Layer	E_T range	ΔE_T	E_R range	ΔE_R
SiO_2	−2 eV to 2 eV	0.1 eV	0.1 eV to 5.0 eV	0.2 eV
HfO_2	−2 eV to 2 eV	0.1 eV	0.1 eV to 5.0 eV	0.2 eV
Layer	x_T range	Δx_T	ΔQ range	ΔQ
SiO_2	0 nm to 1.0 nm	0.1 nm	2.0 $\sqrt{\text{u}}$Å to 8.0 $\sqrt{\text{u}}$Å	1.0 $\sqrt{\text{u}}$Å
HfO_2	1.0 nm to 3.1 nm	0.1 nm	2.0 $\sqrt{\text{u}}$Å to 8.0 $\sqrt{\text{u}}$Å	1.0 $\sqrt{\text{u}}$Å

Table 8.1. The initial grid parameters of the ESiD optimization for a shallow and a deep trapband in both, the SiO_2-layer and the HfO_2-layer. The mean response of the configuration coordinate offset ΔQ and the spatial position x_T has been computed and used for the NNLS optimization of E_T and E_R.

Using this modeling approach allows to accurately and efficiently rebuild the measured data within the simulation, as can be seen from the lines, as can be seen in the simulated lines in Fig. 8.5 (left). The extracted trap band parameters from the simulation are shown in Fig. 8.6 (left). As was the case for *Tech. B*, the interface trap bands in the SiO_2 layer are roughly 1 eV above and below the Si midgap. The deep trap bands, mainly responsible for NBTI in pMOS are concentrated very close to the valence bandedge, whereas the shallow trap bands, responsible for PBTI in nMOS are located around 300 meV above the bandedge. This can be seen also in the (E_T,E_R)-heatmap in Fig. 8.6

(right). There is a peak at small relaxation energies in the shallow SiO_2 trapband which typically does not result in observable ΔV_{th} at room temperature, because these defects exhibit too small charge transition times and therefore do not contribute to ΔV_{th} within the measurement window. However, at cryogenic temperatures these defects contribute to the measured threshold voltage shift.

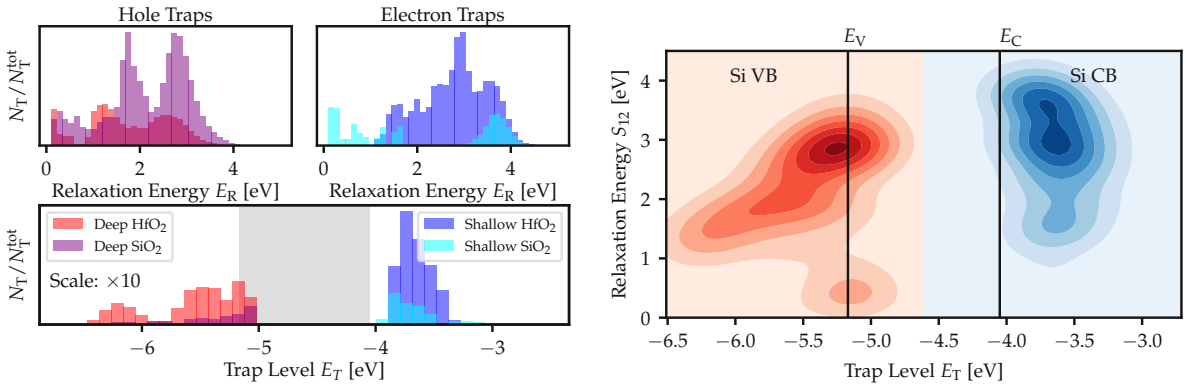

Figure 8.6. The eMSM measurements in Fig. 8.5 (left) allow the extraction of trap bands responsible for causing the ΔV_{th} shift using *Comphy* and the ESiD method. The extracted distribution of trap levels and relaxation energies (left) of defects in the SiO_2 interface layer and the high-κ layer show mean trap levels of approximately 1 eV above and below Si midgap, which is in good agreement with DFT calculations. The corresponding (E_T, E_R)-heatmap (right) shows peaks at the relaxation energies around 3 eV. Figures taken from [MJJ5].

The impact of the different trap bands at different temperatures can be seen in Fig. 8.7. The total ΔV_{th} (solid lines) is composed by the ΔV_{th} caused by defects in the SiO_2 layer (dotted lines) and by such causing the ΔV_{th} of the HfO_2 layer (dashed lines). For nMOS at 300 K the main contribution of the total ΔV_{th} shift comes from defects in the HfO_2-layer which recover slowly. At 4 K, the impact of the HfO_2 trap band is reduced. Additionally, there is a number of very fast defects in the SiO_2 interface layer. The majority of SiO_2 defects recovers slower and gives thus a relevant contribution to the total ΔV_{th} shift.

The described mechanisms are illustrated in Fig. 8.8 with different colors indicating the threshold voltage shift of every defect after a stress of $V_{G,s} = \pm 2\,V$ for $t_s = 1\,ks$ (blue and red defects give a larger contribution to ΔV_{th} than grey defects). Defects which recover immediately after removing the stress are represented by stars. As can be seen, at 300 K a contribution from very fast defects arises, but the main contribution stems from slower defects in the HfO_2 layer. At cryogenic temperatures, the charge transfer kinetics get slower and many defects in the HfO_2 freeze out completely, thus, defects in the SiO_2 IL and in the HfO_2 layer deliver an approximately equal contribution to the total ΔV_{th} shift. NBTI on pMOS in Fig. 8.8 (right) is dominated by defects in the SiO_2 IL. The majority of these defects freezes out and only a less significant subset of defects contributes to ΔV_{th} at cryogenic temperatures.

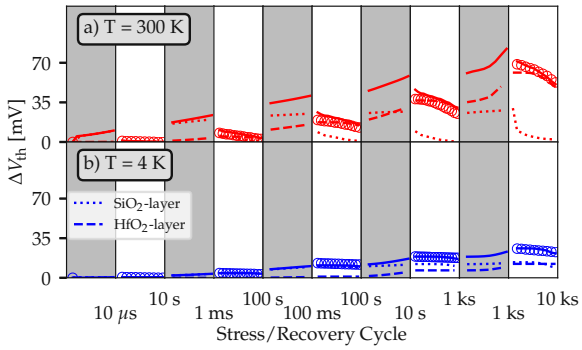

Figure 8.7. The total PBTI threshold voltage shift is composed of the ΔV_{th} caused by the trap band in the SiO_2 layer (dotted) and the trap band in the HfO_2 layer (dashed). At 300 K (red) the main contribution comes from defects in the HfO_2 layer. Very fast interface defects in the SiO_2 layer give a large ΔV_{th} shift, however, this ΔV_{th} recovers almost immediately as soon as the stress is removed. At cryogenic temperatures (blue) the impact defects in the HfO_2, which are far away from the substrate interface, is reduced and the SiO_2 defects recover slower and thus give a relevant contribution to the total ΔV_{th}. Figures taken from [MJJ5].

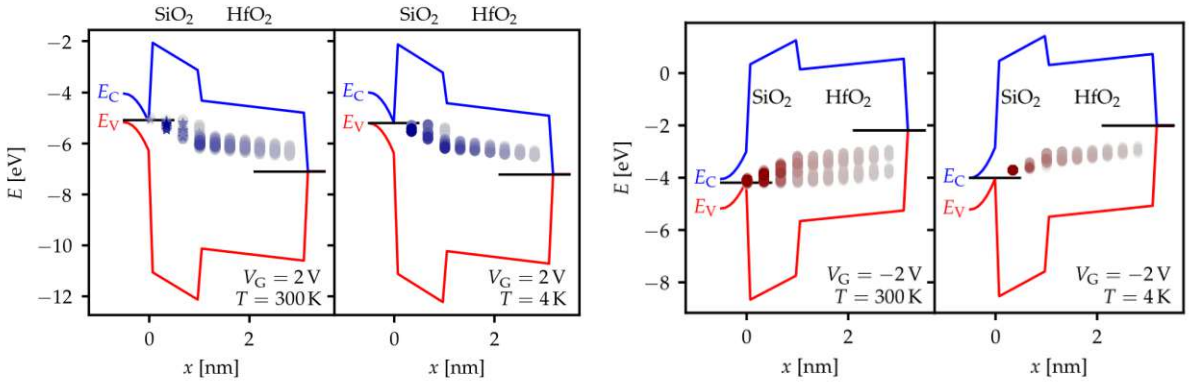

Figure 8.8. Illustration of defects contributing to the total ΔV_{th}. PBTI on nMOS (left) after $t_s = 1$ ks of stress is dominated by slow defects in the HfO_2 layer and very fast defects in the SiO_2 IL, which are marked by stars. At $T = 4$ K many defects in the HfO_2 layer freeze out and the charge transition times constants of the defects in the IL get slower, resulting in equal contributions of the SiO_2 and the HfO_2 defect band to ΔV_{th}. NBTI on pMOS (right) is dominated by defects in the SiO_2 layer. These defects freeze out towards 4 K and only a small, not significant subset of defects contributes to ΔV_{th} at cryogenic temperatures. Figures taken from [MJJ5].

The asymmetry in the ΔV_{th} curves results from the different trap levels and relaxation energies in the trap band distributions. This effect is already known from other technologies, e.g. NBTI on pMOS in [136] where the deep trap bands responsible for NBTI is located closer to the valence band edge than the shallow trap bands leading to an asymmetry of the quantitative exhibition between NBTI and PBTI. The same effect can be seen in *Tech. A* as shown for two exemplary, active defects in Fig. 8.9. Compared to the donor-like defect (right), the acceptor-like defect (left) exhibits a lower classical

barrier (in both cases, stress and recovery), which in turn leads to a larger overlap of the ground state with the corresponding wave function at the same energy level of the final state. Therefore, the donor-like defect delivers a larger contribution to the total ΔV_{th} compared to the acceptor-like defect which will have smaller charge transition times.

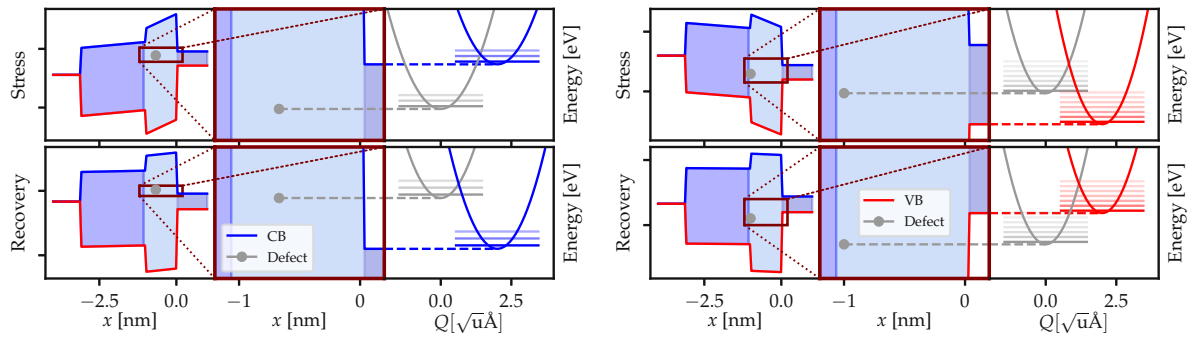

Figure 8.9. Acceptor-like defects responsible for PBTI on nMOS (left) show lower classical barriers compared to donor-like defects responsible for NBTI on pMOS (right). This has the effect that the overlap between the vibrational wavefunction of the initial and the final state for the acceptor-like defect is larger than on comparable donor-like defects leading to smaller charge transition times at cryogenic temperatures. Figures taken from [MJJ5].

The charge transition times corresponding to the selected defects in Fig. 8.9 are highlighted in Fig. 8.10 for $T = 4, 77, 150, 225, 300$ K. The majority of the charge transition times of defects in the nMOS device is smaller than those of the defects in the pMOS device. Defects with charge transition times that are accessible in the measurement window at 300 K are shifted out of it in the pMOS device at 4 K, as can be seen in the lower right corner, while on nMOS there is still a large number of defects with charge transition times below 10 ks.

The characterization of NBTI and PBTI on large area devices shows that different technologies show a different freeze-out behavior. While on *Tech. B* both NBTI and PBTI freeze out completely towards cryogenic temperatures, *Tech. A* shows significant PBTI on nMOS even at 4 K while NBTI on pMOS freezes out. This asymmetry can be modeled with the reliability simulator *Comphy* using defect bands which show an asymmetry in the offset of the thermodynamic trap level to the band edge and in the relaxation energies.

The measurements on large area devices characterized by employing the eMSM method in this section show continuous recovery traces. If stress/recovery measurements are performed on scaled devices, discrete steps occur, which correspond to charge transitions involving single defects, as discussed in more detail in the following section.

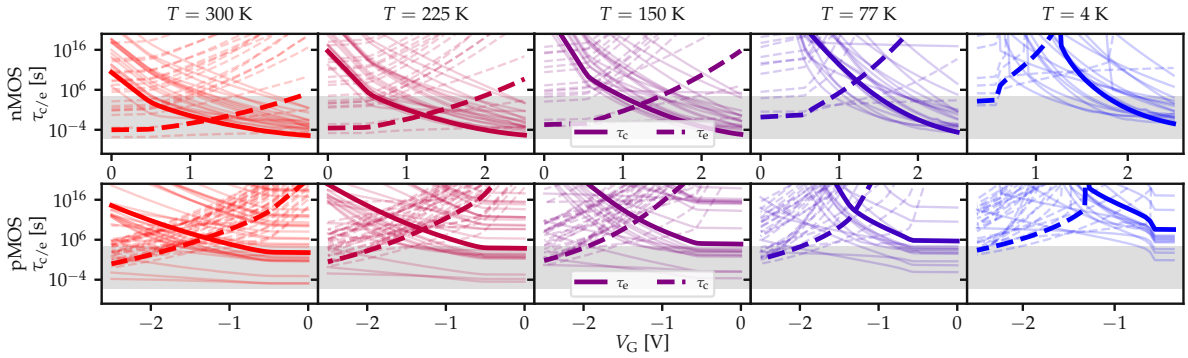

Figure 8.10. The highlighted charge transition times for the nMOS (upper) and the pMOS (lower) figure for different gate voltages at $T = 4, 77, 150, 225, 300\,\mathrm{K}$ correspond to defects shown in Fig. 8.9. Most charge transition times shown for the nMOS device are faster than for the pMOS device leading to more active defects in the grey marked measurement window. While almost all defects of the pMOS are shifted out of the measurement window, there is still a significant number of nMOS defects within the window at the low temperatures of 77 K and 4 K. Figure taken from [MJJ5].

8.3 TDDS Measurements at Cryogenic Temperatures

The time-dependent defect spectroscopy (TDDS), next to RTN, is the second electrical characterization method presented in this work, to characterize single defects. TDDS measurements have the advantage that the charging of a trap can be enforced by applying a high gate bias compared to RTN characterization, and has been used for characterizing a device of *Tech. C* with drawn dimensions $W \times L = 70 \times 70\,\mathrm{nm}^2$. After applying a stress bias of $V_{\mathrm{G,s}} = 16\,\mathrm{V}$ for $t_{\mathrm{s}} = 100\,\mathrm{s}$, a recovery trace at $V_{\mathrm{G,r}} = 8\,\mathrm{V}$ has been recorded. This has been repeated 25 times at $T = 20, 40, 60, 80$ and $100\,\mathrm{K}$. From the set of recovery traces recorded at one temperature, e.g. Fig. 8.11 (left) for $T = 20\,\mathrm{K}$, it is possible to extract a heatmap of the step heights and the emission times as it can be seen in Fig. 8.11 (center). The heatmap shows that the emission times are exponentially distributed with a mean of $\overline{\tau_{\mathrm{e}}} = 8.5\,\mathrm{s}$ and a mean step height of $\overline{\Delta I} = 3.7\,\mathrm{nA/\mu m}$. Note that often multiple steps with significantly different mean step heights and time constants can be found in a single recovery trace. These defects are then located in clearly separated clusters in the heatmap.

The extracted mean emission times for each measured temperature can be plotted in an Arrhenius plot in Fig. 8.11 (right). As can be seen, the mean emission times do not exhibit an Arrhenius-like temperature activation towards 20 K. As discussed in the previous chapters, this can be explained by the full quantum mechanical version of the NMP theory. At cryogenic temperatures the vibrational groundstates of the initial and final state of a charge transition are still occupied and overlapping, which allows nuclear tunneling between the atomistic configurations. Therefore, the transition rates become constant towards cryogenic temperatures allowing charge transitions even at low temperatures.

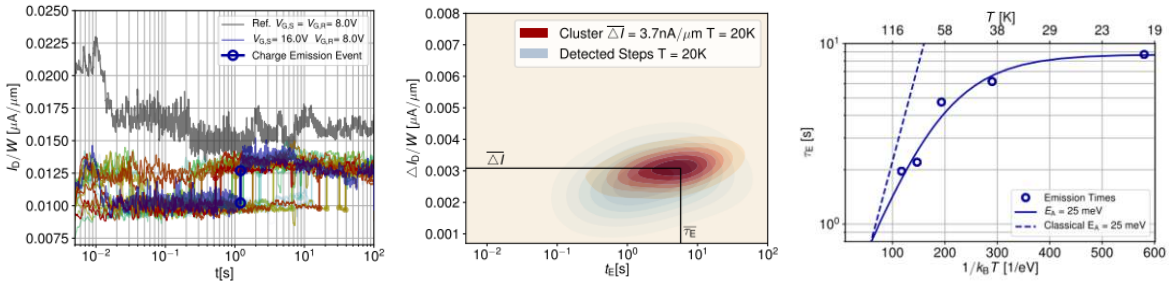

Figure 8.11. A set of recovery traces (left) measured at $V_{G,r} = 8\,V$ and $T = 20\,K$ shows a charge emission event after a stress of $V_{G,s} = 16\,V$ for $t_s = 100\,s$. From the set of recovery traces the mean emission time $\overline{\tau_e}$ and the mean step height $\overline{\Delta I}/W$ can be extracted and the distribution of parameters can be presented in a heatmap (center). $\overline{\tau_e}$ for a set of temperatures shows that the emission time behaves not Arrhenius-like but gets dominated by nuclear tunneling towards cryogenic temperatures (right). The Figures taken from [241, MJC6].

To summarize this chapter, BTI measurements on large area devices and TDDS measurements on down-scaled devices show that the freeze out behavior of the degradation shows a strong technology dependence. While BTI on SiON devices freezes out completely, there is still PBTI degradation on nMOS HKMG devices. This can be modeled with the reliability simulator *Comphy* by using the 2-state NMP charge transition model. The TDDS measurements on *Tech. C* confirm single defect emission events in the limit of cryogenic temperatures also for devices made from 2D materials.

Chapter 9

Conclusion and Outlook

9.1 Conclusion

Cryo-CMOS can be used in various emerging technologies such as in quantum computing as control interface for qubits located at a 4 K stage or as a host technology for a monolithic co-integration of qubits and control instances. It can also be used for high performance computing to increase computational power and to save energy consumption. These applications often require very narrow margins of errors, therefore it is crucial to understand the temperature dependence of MOSFET characteristics to guarantee a high reliability and a low variability of the devices.

All the named applications make use of the changes in the MOSFET $I_D(V_G)$ transfer curves towards cryogenic temperatures. Due to the reduced scattering, the charge carrier mobility increases which leads to an increasing on-state current, an increasing transconductance and a steeper subthreshold slope. The temperature dependence of the Fermi levels, the carrier distributions, the band gap and other device properties lead to an increasing threshold voltage. However, in principle the threshold voltage can be reduced by changing design properties, e.g. the processed gate metal or the channel material which allows in combination with the steep subthreshold slope low-V_{DD} operation at cryogenic temperatures. This then enables to increase logic switching rates and thus the performance of the MOSFET. However, low-V_{DD} applications are extremely sensitive to device-to-device variability. In this work, the variability of a commercial planar 28 nm CMOS high-κ metal gate technology is studied on *SmartArrays* with thousands of devices which can be addressed digitally. This allows to show that the variability of multiple device parameters increases towards 4 K which must be considered in future circuit designs.

In addition, various reliability issues which can arise from electrical stress such as bias temperature instability (BTI), hot carrier degradation (HCD) or random telegraph noise (RTN) have been linked to oxide defects and defects at the interface between insulator and substrate. These defects can capture and emit charges during operation and thus alter the transfer characteristics. The charge transfer kinetics can be described

precisely within the nonradiative multiphonon (NMP) model. According to the NMP model, charge transitions between a defect and a charge reservoir are enabled by phonons. While the classical limit of the theory would predict a freeze out of charge transitions, the full quantum mechanical picture allows transitions between atomistic configurations due to nuclear tunneling even at cryogenic temperatures. Since the calculation of the quantum mechanical transition rate is computationally too expensive for reliability simulations, for which typically thousands of defects must be sampled, a WKB-based approximation has been developed which is computationally superior. This model has been implemented in the reliability simulator *Comphy* which is further used for modeling reliability measurement data.

It has to be pointed out that different reliability issues are dominant across various bias conditions. In this work, the focus has been set on bias temperature instability (BTI) and random telegraph noise (RTN), which has been characterized on various technologies, among others the before mentioned high-κ metal gate technology and a commercial planar SiON technology of a 180 nm technology node. BTI characterization has been performed between room temperature and 4 K for various stress conditions. While on SiON BTI freezes out completely towards 4 K, there is considerably positive BTI on nMOSFETs but no negative BTI on pMOSFETs in the HKMG technology at cryogenic temperatures. This can be modeled using the WKB-based charge transition rate and *Comphy* by sampling defect bands with different dominant energy levels and relaxation energy distributions for nMOS and pMOS. The occurring BTI degradation at 4 K can be a major reliability concern, specially for low-V_{DD} applications which allow only small margins of threshold voltage shifts. Therefore, BTI shifts must be considered in optimized cryo PDKs and circuits.

While BTI is characterized under elevated stress bias conditions and might be weaker during operation, the impact of RTN cannot be avoided. Charge trapping occurs even at deep cryogenic temperatures due to nuclear tunneling. Towards 4 K nuclear tunneling gets temperature independent, resulting in non-Arrhenius-like transition rates at cryogenic temperatures. By using *Comphy*, defects at the interface between oxide and substrate have been identified to be responsible for RTN in the commercial SiON technology at cryogenic temperatures. This may be a major concern in solid state qubits and future monolithic designs, because qubits are extremely sensitive to decoherence. Charge noise could cause dephasing and lower the fidelity of the qubits, therefore a deep understanding of the trap kinetics and how to reduce charge noise to a minimum is essential.

The different reliability and variability studies show that charge trapping between atomistic configurations caused by nuclear tunneling is a concern for the stable operation of devices, even at cryogenic temperatures. This is a special concern for potential low-V_{DD} applications which are currently under investigation to reduce power consumption in the cryogenic environment which is essential for quantum computing and to increase logic switching rates for high performance computing. Knowledge gained from the

studies presented in this work can further be applied to related fields as the design of solid state qubits, where charge noise caused by oxide defects is crucial for a high fidelity.

9.2 Outlook

In the field of reliability physics, many tools and characterization methods have been developed for getting a deep understanding of the ongoing degradation mechanisms and the responsible defect physics. In this work, the focus is set on BTI degradation and RTN, however, other degradation mechanisms such as HCD, stress induced leakage current or time-dependent gate oxide breakdown have been left out of the picture. Characterizing these degradation mechanisms between 4 K and room temperature would be an enormous step towards a complete overview of the overall trends of reliability issues at cryogenic temperatures.

BTI has been studied in this work mainly on large area devices. By using time-dependent defect spectroscopy, the extension of the characterization to scaled devices could give deep insights in the ongoing charge trapping kinetics at cryogenic temperatures. Charge noise, on the other hand, was mainly characterized in the form of RTN on scaled devices. Since this is quite time-consuming, studying $1/f$ noise on large area devices between room temperature and 4 K could give a clearer understanding of the overall charge noise trends and its dependence on temperature and bias conditions.

The studies in these work have been done mainly on a commercial planar 28 nm high-κ metal gate technology and a commercial planar 180 nm SiON technology. Repeating these studies on other technology nodes, specially on back-gated FD-SOI technologies, which are in the focus of many research groups, could deliver a valuable input regarding the impact of the device design on the reliability. In the case of back-gated FD-SOI devices the impact of the back-gate bias on BTI degradation, charge noise and device-to-device variability could be studied.

The access to experimental low-V_{DD} MOSFETs, which are specially optimized for cryogenic applications and do not work at room temperature, would allow almost infinite possibilities for reliability studies. The performance in terms of reliability of such structures is unknown yet. However, due to the minimized V_{DD} the structures might be very sensitive to variability and threshold voltage shifts, therefore detailed studies of the known degradation mechanisms are of a very high interest.

The knowledge of defect physics in the context of MOSFETs at cryogenic temperatures could also be applied to qubit structures. Since qubits are complicated structures with multiple oxide layers, charge trapping kinetics of defects in these oxide layers could be very similar to the kinetics in MOSFETs. Studying the technology used for qubits could give important insights how to reduce charge trapping, which is a main concern for stable operation, because it can lead to decoherence and dephasing of the qubits.

Appendix A

Appendix

A.1 WKB Approximation

The WKB approximation was developed in 1926 independently by Gregor Wentzel, Hendrik Kramers and Léon Brillouin [178]. It can be used for solving quantum mechanical eigen-value problems iteratively. Thus, it can also be used for approximating the time independent Schrödinger equation

$$-\frac{\hbar^2}{2m}\frac{\mathrm{d}^2\psi(x)}{\mathrm{d}x^2} + V(x)\psi(x) = E\psi(x), \tag{A.1}$$

which can be rewritten as

$$\frac{\mathrm{d}^2\psi(x)}{\mathrm{d}x^2} + k^2(x)\psi(x) = 0 \tag{A.2}$$

with $k^2(x) = \frac{2m}{\hbar^2}(E - V(x))$. Solving this equation can be done using the ansatz

$$\psi(x) = c\mathrm{e}^{\frac{i}{\hbar}W(x)}. \tag{A.3}$$

Plugging the ansatz (A.3) into the Schrödinger equation (A.2) it can be transfered to an inhomogeneous non-linear differential equation of second order for W which results in

$$(W'(x))^2 - i\hbar W''(x) = \hbar^2 k^2(x). \tag{A.4}$$

Based on the principle that quantum mechanical relations should correspond to classical relations when $\hbar \to 0$ the phase $W(x)$ can be expanded in $\hbar$

$$W(x) = \sum_{n=0}^{\infty} (i\hbar)^n W_n(x). \tag{A.5}$$

Plugging (A.5) into (A.4) and sorting the result by the order in $\hbar$ we get

$$(W_0'^2 - \hbar^2 k^2) + i\hbar(2W_0'W_1' - W_0'') + \mathcal{O}(\hbar^2) = 0 \tag{A.6}$$

where $\hbar^2 k^2$ is of order zero in $\hbar$ because $k^2 \propto 1/\hbar^2$. From the zeroth order term the differential equation

$$W_0'(x) = \pm\hbar k(x) \tag{A.7}$$

can be extracted and solved:

$$W_0(x) = \pm\hbar \int^x k(x')\mathrm{d}x'. \tag{A.8}$$

From the first order $\hbar$ term the corresponding differential equation

$$W_1'(x) = \frac{1}{2}\frac{W_0''(x)}{W_0'(x)} \tag{A.9}$$

can be solved

$$W_1(x) = \ln\sqrt{k(x)}. \tag{A.10}$$

Using these results in the ansatz (A.3) the WKB-wavefunction has the from

$$\psi(x) = \frac{c_+}{\sqrt{k(x)}} \mathrm{e}^{i\int^x \sqrt{k(x')}\mathrm{d}x'} + \frac{c_-}{\sqrt{k(x)}} \mathrm{e}^{-i\int^x \sqrt{k(x')}\mathrm{d}x'}. \tag{A.11}$$

A.2 Saddlepoint Method

The stationary phase approximation or saddlepoint method was developed in the 18$^{\text{th}}$ century by Pierre-Simon Laplace [252]. The aim is to find an approximation for the integral

$$I = \int_{-\infty}^{\infty} \mathrm{e}^{-f(x)}\mathrm{d}x \tag{A.12}$$

for analytic functions $f(x)$ with a global minimum at x_0, which implies that

$$f'(x_0) = 0 \tag{A.13}$$

and

$$f''(x_0) > 0 \tag{A.14}$$

holds. Expanding $f(x)$ in a Taylor series at the minimum x_0 and simplifying the expression using $f'(x_0) = 0$

$$\begin{aligned} f(x) &\approx f(x_0) + f'(x_0)(x - x_0) + \frac{1}{2} f''(x_0)(x - x_0)^2 \\ &= f(x_0) + \frac{1}{2} f''(x_0)(x - x_0)^2, \end{aligned} \tag{A.15}$$

we get

$$\begin{aligned} I &= \int_{-\infty}^{\infty} \mathrm{e}^{-f(x)} \mathrm{d}x \\ &\approx \int_{-\infty}^{\infty} \mathrm{e}^{-f(x_0) - \frac{1}{2} f''(x_0)(x - x_0)^2} \mathrm{d}x \\ &= \mathrm{e}^{-f(x_0)} \int_{-\infty}^{\infty} \mathrm{e}^{-\frac{1}{2} f''(x_0)(x - x_0)^2} \mathrm{d}x. \end{aligned} \tag{A.16}$$

Using the solution of the Gaussian integral

$$\int_{-\infty}^{\infty} \mathrm{e}^{-a(x+b)^2} \mathrm{d}x = \sqrt{\frac{\pi}{a}} \tag{A.17}$$

we can evaluate (A.16) to get the final result

$$I = \mathrm{e}^{-f(x_0)} \sqrt{\frac{2\pi}{f''(x_0)}}. \tag{A.18}$$

References

[1] A. G. J. MacFarlane, J. P. Dowling, and G. J. Milburn. "Quantum Technology: The Second Quantum Revolution". In: *Philosophical Transactions of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences* 361.1809 (2003), pp. 1655–1674. DOI: 10.1098/rsta.2003.1227.

[2] J. E. Lilienfeld. *Device for Controlling Electric Current*. US1900018A. 1928.

[3] O. Heil. *Improvements in or Relating to Electrical Amplifiers and other Control Arrangements and Devices*. GB439457. 1935.

[4] C.-T. Sah. "Evolution of the MOS Transistor–From Conception to VLSI". In: *Proceedings of the IEEE* 76.10 (1988), pp. 1280–1326. DOI: 10.1109/5.16328.

[5] J. Bardeen and W. H. Brattain. *Three-Electrode Circuit Element Utilizing Semiconductive materials*. US2524035A. Oct. 1950.

[6] D. Kahng. *Electric Field Controlled Semiconductor Device*. US3102230. Aug. 1963.

[7] M. M. Atalla. *Semiconductor Devices Having Dielectric Coatings*. US3206670. Sept. 1965.

[8] B. Deal. "Standardized Terminology for Oxide Charges Associated with Thermally Oxidized Silicon". In: *IEEE Transactions on Electron Devices* 27.3 (1980), pp. 606–608. DOI: 10.1109/T-ED.1980.19908.

[9] F. M. Wanlass and C. T. Sah. "Nanowatt Logic Using Field-Effect Metal-Oxide Semiconductor Triodes". In: *Semiconductor Devices: Pioneering Papers*, pp. 637–638. DOI: 10.1142/9789814503464_0081.

[10] F. M. Wanlass. *Low Stand-By Power Complementary Field Effect Circuitry*. US3356858. Dec. 1967.

[11] G. E. Moore. *Cramming More Components onto Integrated Circuits, Reprinted from Electronics, volume 38, number 8, April 19, 1965, pp.114 ff.* 2006. DOI: 10.1109/N-SSC.2006.4785860.

[12] *International Technology Roadmap for Semiconductors*. 2020. URL: http://www.itrs2.net/ (visited on 05/29/2022).

[13] M. M. Waldrop. "The Chips are Down for Moore's Law". In: *Nature* 530 (2016), pp. 144–147. DOI: 10.1038/530144a.

[14] *IEEE International Roadmap for Devices and Systems*. 2022. URL: https://irds.ieee.org/ (visited on 05/29/2022).

[15] H. L. Chiang, T. C. Chen, J. F. Wang, S. Mukhopadhyay, W. K. Lee, C. L. Chen, W. S. Khwa, B. Pulicherla, P. J. Liao, K. W. Su, K. F. Yu, T. Wang, H. S. P. Wong, C. H. Diaz, and J. Cai. "Cold CMOS as a Power-Performance-Reliability Booster for Advanced FinFETs". In: *2020 IEEE Symposium on VLSI Technology*. 2020, pp. 1–2. DOI: 10.1109/VLSITechnology18217.2020.9265065.

[16] S. S. T. Nibhanupudi, S. R. Sundara Raman, M. Cassé, L. Hutin, and J. P. Kulkarni. "Ultra-Low-Voltage UTBB-SOI-Based, Pseudo-Static Storage Circuits for Cryogenic CMOS Applications". In: *IEEE Journal on Exploratory Solid-State Computational Devices and Circuits* 7.2 (2021), pp. 201–208. DOI: 10.1109/JXCDC.2021.3130839.

[17] R. Saligram, S. Datta, and A. Raychowdhury. "CryoMem: A 4K-300K 1.3GHz eDRAM Macro with Hybrid 2T-Gain-Cell in a 28nm Logic Process for Cryogenic Applications". In: *2021 IEEE Custom Integrated Circuits Conference (CICC)*. 2021, pp. 1–2. DOI: 10.1109/CICC51472.2021.9431527.

[18] F. Tavernier, A. Gatti, and C. Barretto. "Chip Design for Future Gravitational Wave Detectors". In: *2020 IEEE International Electron Devices Meeting (IEDM)*. 2020, pp. 25.4.1–25.4.4. DOI: 10.1109/IEDM13553.2020.9372071.

[19] *Outline of the People's Republic of China 14th Five-Year Plan for National Economic and Social Development and Long-Range Objectives for 2035*. 2021. URL: https://cset.georgetown.edu/wp-content/uploads/t0284_14th_Five_Year_Plan_EN.pdf (visited on 05/29/2022).

[20] *Quantum Flagship*. 2021. URL: https://qt.eu/ (visited on 05/29/2022).

[21] F. Arute et al. "Quantum Supremacy Using a Programmable Superconducting processor". In: *Nature* 574.7779 (2019), pp. 505–510. ISSN: 1476-4687. DOI: 10.1038/s41586-019-1666-5.

[22] M. Steffen, D. P. DiVincenzo, J. M. Chow, T. N. Theis, and M. B. Ketchen. "Quantum Computing: An IBM Perspective". In: *IBM Journal of Research and Development* 55.5 (2011), 13:1–13:11. DOI: 10.1147/JRD.2011.2165678.

[23] T. Last, N. Samkharadze, P. Eendebak, R. Versluis, X. Xue, A. Sammak, D. Brousse, K. Loh, H. Polinder, G. Scappucci, M. Veldhorst, L. Vandersypen, K. Maturová, J. Veltin, and G. Alberts. "Quantum Inspire: QuTech's Platform for Co-Development and Collaboration in Quantum Computing". In: *Novel Patterning Technologies for Semiconductors, MEMS/NEMS and MOEMS 2020*. Ed. by M. I. Sanchez and E. M. Panning. Vol. 11324. International Society for Optics and Photonics. SPIE, 2020, pp. 49–59. DOI: 10.1117/12.2551853.

[24] J. A. Jones, M. Mosca, and R. H. Hansen. "Implementation of a Quantum Search Algorithm on a Quantum Computer". In: *Nature* 393.6683 (May 1998), pp. 344–346. ISSN: 1476-4687. DOI: 10.1038/30687.

[25] I. L. Chuang, N. Gershenfeld, and M. Kubinec. "Experimental Implementation of Fast Quantum Searching". In: *Phys. Rev. Lett.* 80 (15 Apr. 1998), pp. 3408–3411. DOI: 10.1103/PhysRevLett.80.3408.

[26] I. L. Chuang, L. M. K. Vandersypen, X. Zhou, D. W. Leung, and S. Lloyd. "Experimental Realization of a Quantum Algorithm". In: *Nature* 393.6681 (May 1998), pp. 143–146. ISSN: 1476-4687. DOI: 10.1038/30181.

[27] L. M. K. Vandersypen, M. Steffen, G. Breyta, C. S. Yannoni, M. H. Sherwood, and I. L. Chuang. "Shor's Quantum Factoring Algorithm Using Nuclear Magnetic Resonance". In: *Nature* 414.6866 (Dec. 2001), pp. 883–887. ISSN: 1476-4687. DOI: 10.1038/414883a.

[28] J. Gambetta. *IBM's Roadmap for Scaling Quantum Technology*. Sept. 2020. URL: https://research.ibm.com/blog/ibm-quantum-roadmap (visited on 05/29/2022).

[29] P. J. Karalekas, N. A. Tezak, E. C. Peterson, C. A. Ryan, M. P. da Silva, and R. S. Smith. "A Quantum-Classical Cloud Platform Optimized for Variational Hybrid Algorithms". In: *Quantum Science and Technology* 5.2 (Apr. 2020), p. 024003. DOI: 10.1088/2058-9565/ab7559.

[30] R. Izsak, C. Riplinger, N. S. Blunt, B. de Souza, N. Holzmann, O. Crawford, J. Camps, F. Neese, and P. Schopf. *Quantum Computing in Pharma: A Multilayer Embedding Approach for Near Future Applications*. 2022. DOI: 10.48550/ARXIV.2202.04460.

[31] J. Hsu. "CES 2018: Intel's 49-Qubit Chip Shoots for Quantum Supremacy". In: *IEEE Spectrum Tech Talk* (2018), pp. 1–6.

[32] *Alibaba Cloud and CAS Launch One of the World's Most Powerful Public Quantum Computing Services*. Mar. 2018. URL: https://www.alibabacloud.com/press-room/alibaba-cloud-and-cas-launch-one-of-the-worlds-most?spm=a3c0i.8288105.651515.3.61356dc28LK6Rw (visited on 01/28/2022).

[33] A. K. Fedorov, A. V. Akimov, J. D. Biamonte, A. V. Kavokin, F. Y. Khalili, E. O. Kiktenko, N. N. Kolachevsky, Y. V. Kurochkin, A. I. Lvovsky, A. N. Rubtsov, G. V. Shlyapnikov, S. S. Straupe, A. V. Ustinov, and A. M. Zheltikov. "Quantum Technologies in Russia". In: *Quantum Science and Technology* 4.4 (Oct. 2019), p. 040501. DOI: 10.1088/2058-9565/ab4472.

[34] R. Barends et al. "Digitized Adiabatic Quantum Computing with a Superconducting Circuit". In: *Nature* 534.7606 (June 2016), pp. 222–226. ISSN: 1476-4687. DOI: 10.1038/nature17658.

[35] J. Kelly. *A Preview of Bristlecone, Google's New Quantum Processor*. Mar. 2018. URL: https://ai.googleblog.com/2018/03/a-preview-of-bristlecone-googles-new.html (visited on 05/29/2022).

[36] *Scalable Quantum Systems Built from the Chip up to Power Practical Applications*. Nov. 2021. URL: https://www.rigetti.com/what-we-build (visited on 05/29/2022).

[37] *Quantum Inspire Spin-2 QPU*. URL: https://www.quantum-inspire.com/backends/spin-2/ (visited on 05/29/2022).

[38] *Quantum Inspire Starmon-5*. URL: https://www.quantum-inspire.com/backends/starmon-5/ (visited on 05/29/2022).

[39] R. Blümel, N. Grzesiak, N. H. Nguyen, A. M. Green, M. Li, A. Maksymov, N. M. Linke, and Y. Nam. "Efficient Stabilized Two-Qubit Gates on a Trapped-Ion Quantum Computer". In: *Phys. Rev. Lett.* 126 (22 June 2021), p. 220503. DOI: 10.1103/PhysRevLett.126.220503.

[40] P. Chapman. *Scaling IonQ's Quantum Computers: The Roadmap*. Dec. 2022. URL: https://ionq.com/posts/december-09-2020-scaling-quantum-computer-roadmap (visited on 05/29/2022).

[41] *Demonstrating Benefits of Quantum Upgradable Design Strategy: System Model H1-2 First to Prove 2,048 Quantum Volume*. Dec. 2021. URL: https://www.quantinuum.com/pressrelease/demonstrating-benefits-of-quantum-upgradable-design-strategy-system-model-h1-2-first-to-prove-2-048-quantum-volume (visited on 05/29/2022).

[42] *Quantum Milestone: How We Quadrupled Performance*. 2021. URL: https://www.honeywell.com/us/en/news/2021/03/quantum-milestone-how-we-quadrupled-performance (visited on 05/29/2022).

[43] M. Sheffer, D. Azses, and E. G. D. Torre. "Playing Quantum Nonlocal Games with Six Noisy Qubits on the Cloud". In: *Advanced Quantum Technologies* 5.3 (Jan. 2022), p. 2100081. DOI: 10.1002/qute.202100081.

[44] J. M. Arrazola et al. "Quantum Circuits with many Photons on a Programmable Nanophotonic Chip". In: *Nature* 591.7848 (Mar. 2021), pp. 54–60. ISSN: 1476-4687. DOI: 10.1038/s41586-021-03202-1.

[45] *Xanadu Quantum Cloud*. 2022. URL: https://www.xanadu.ai/cloud (visited on 05/29/2022).

[46] X.-L. Wang, Y.-H. Luo, H.-L. Huang, M.-C. Chen, Z.-E. Su, C. Liu, C. Chen, W. Li, Y.-Q. Fang, X. Jiang, J. Zhang, L. Li, N.-L. Liu, C.-Y. Lu, and J.-W. Pan. "18-Qubit Entanglement with Six Photons' Three Degrees of Freedom". In: *Phys. Rev. Lett.* 120 (June 2018), p. 260502. DOI: 10.1103/PhysRevLett.120.260502.

[47] H.-S. Zhong et al. "Quantum Computational Advantage Using Photons". In: *Science* 370.6523 (2020), pp. 1460–1463. DOI: 10.1126/science.abe8770.

[48] R. Harris et al. "Experimental Demonstration of a Robust and Scalable Flux Qubit". In: *Phys. Rev. B* 81 (13 Apr. 2010), p. 134510. DOI: 10.1103/PhysRevB.81.134510.

[49] S. Boixo, T. F. Rønnow, S. Isakov, Z. Wang, D. Wecker, D. Lidar, J. Martinis, and M. Troyer. "Quantum Annealing with More Than One Hundred Qubits". In: *Nat Phys* 10 (Apr. 2013). DOI: 10.1038/nphys2900.

[50] J. King, S. Yarkoni, M. M. Nevisi, J. P. Hilton, and C. C. McGeoch. *Benchmarking a Quantum Annealing Processor with the Time-To-Target Metric*. 2015. DOI: 10.48550/ARXIV.1508.05087.

[51] N. Dattani, S. Szalay, and N. Chancellor. *Pegasus: The Second Connectivity Graph for Large-Scale Quantum Annealing Hardware*. 2019. DOI: 10.48550/ARXIV.1901.07636.

[52] M. A. Nielsen and I. L. Chuang. *Quantum Computation and Quantum Information*. Cambridge University Press, 2000.

[53] P. W. Shor. "Polynomial-Time Algorithms for Prime Factorization and Discrete Logarithms on a Quantum Computer". In: *SIAM Review* 41.2 (1999), pp. 303–332. DOI: 10.1137/S0036144598347011.

[54] J. Biamonte, P. Wittek, N. Pancotti, P. Rebentrost, N. Wiebe, and S. Lloyd. "Quantum Machine Learning". In: *Nature* 549.7671 (Sept. 2017), pp. 195–202. ISSN: 1476-4687. DOI: 10.1038/nature23474.

[55] A. Montanaro. "Quantum Algorithms: An Overview". In: *npj Quantum Information* 2.1 (Jan. 2016), p. 15023. ISSN: 2056-6387. DOI: 10.1038/npjqi.2015.23.

[56] B. P. Lanyon, J. D. Whitfield, G. G. Gillett, M. E. Goggin, M. P. Almeida, I. Kassal, J. D. Biamonte, M. Mohseni, B. J. Powell, M. Barbieri, A. Aspuru-Guzik, and A. G. White. "Towards Quantum Chemistry on a Quantum Computer". In: *Nature Chemistry* 2.2 (Feb. 2010), pp. 106–111. ISSN: 1755-4349. DOI: 10.1038/nchem.483.

[57] S. Woerner and D. J. Egger. "Quantum Risk Analysis". In: *npj Quantum Information* 5.1 (Feb. 2019), p. 15. ISSN: 2056-6387. DOI: 10.1038/s41534-019-0130-6.

[58] E. Pednault, J. Gunnels, D. Maslov, and J. Gambetta. *On "Quantum Supremacy"*. Oct. 2019. URL: https://www.ibm.com/blogs/research/2019/10/on-quantum-supremacy/ (visited on 01/14/2022).

[59] L. K. Grover. "Quantum Mechanics Helps in Searching for a Needle in a Haystack". In: *Phys. Rev. Lett.* 79 (2 July 1997), pp. 325–328. DOI: 10.1103/PhysRevLett.79.325.

[60] *IBM Research*. 2022. URL: https://research.ibm.com/ (visited on 04/14/2022).

[61] K. R. Brown, A. C. Wilson, Y. Colombe, C. Ospelkaus, A. M. Meier, E. Knill, D. Leibfried, and D. J. Wineland. "Single-Qubit-Gate Error Below 10^{-4} in a Trapped Ion". In: *Phys. Rev. A* 84 (Sept. 2011), p. 030303. DOI: 10.1103/PhysRevA.84.030303.

[62] G. Waldherr, Y. Wang, S. Zaiser, M. Jamali, T. Schulte-Herbrüggen, H. Abe, T. Ohshima, J. Isoya, J. F. Du, P. Neumann, and J. Wrachtrup. "Quantum Error Correction in a Solid-State Hybrid Spin Register". In: *Nature* 506.7487 (Feb. 2014), pp. 204–207. ISSN: 1476-4687. DOI: 10.1038/nature12919.

[63] F. Dolde, V. Bergholm, Y. Wang, I. Jakobi, B. Naydenov, S. Pezzagna, J. Meijer, F. Jelezko, P. Neumann, T. Schulte-Herbrüggen, J. Biamonte, and J. Wrachtrup. "Spin Entanglement Using Optimal Control". In: *Nature Communications* 5.1 (Feb. 2014), p. 3371. ISSN: 2041-1723. DOI: 10.1038/ncomms4371.

[64] P. Kok, W. J. Munro, K. Nemoto, T. C. Ralph, J. P. Dowling, and G. J. Milburn. "Linear Optical Quantum Computing with Photonic Qubits". In: *Rev. Mod. Phys.* 79 (Jan. 2007), pp. 135–174. DOI: 10.1103/RevModPhys.79.135.

[65] B. Govoreanu, S. Kubicek, J. Jussot, B. Chan, N. Dumoulin-Stuyck, F. Mohiyaddin, R. Li, G. Simion, T. Ivanov, D. Mocuta, J. Lee, and I. Radu. "Moving Spins From Lab to Fab: A Silicon-Based Platform for Quantum Computing Device Technologies". In: *2019 Silicon Nanoelectronics Workshop (SNW)*. 2019, pp. 1–2. DOI: 10.23919/SNW.2019.8782903.

[66] M. Veldhorst, C. H. Yang, J. C. C. Hwang, W. Huang, J. P. Dehollain, J. T. Muhonen, S. Simmons, A. Laucht, F. E. Hudson, K. M. Itoh, A. Morello, and A. S. Dzurak. "A Two-Qubit Logic Gate in Silicon". In: *Nature* 526.7573 (Oct. 2015), pp. 410–414. ISSN: 1476-4687. DOI: 10.1038/nature15263.

[67] I. P. Radu et al. "Solid State Qubits: How Learning from CMOS Fabrication can Speed-Up Progress in Quantum Computing". In: *2021 Symposium on VLSI Technology*. 2021, pp. 1–2. DOI: 10.23919/VLSICircuits52068.2021.9492397.

[68] M. Kjaergaard, M. E. Schwartz, J. Braumüller, P. Krantz, J. I.-J. Wang, S. Gustavsson, and W. D. Oliver. "Superconducting Qubits: Current State of Play". In: *Annual Review of Condensed Matter Physics* 11.1 (2020), pp. 369–395. DOI: 10.1146/annurev-conmatphys-031119-050605.

[69] P. Krantz, M. Kjaergaard, F. Yan, T. P. Orlando, S. Gustavsson, and W. D. Oliver. "A Quantum Engineer's Guide to Superconducting Qubits". In: *Applied Physics Reviews* 6.2 (2019), p. 021318. DOI: 10.1063/1.5089550.

[70] M. H. Devoret, A. Wallraff, and J. M. Martinis. *Superconducting Qubits: A Short Review*. 2004. DOI: 10.48550/ARXIV.COND-MAT/0411174.

[71] D. Stancil and G. Byrd. *Principles of Superconducting Quantum Computers*. Wiley, 2022. ISBN: 9781119750741.

[72] J. Verjauw, R. Acharya, J. Van Damme, T. Ivanov, D. P. Lozano, F. A. Mohiyaddin, D. Wan, J. Jussot, A. M. Vadiraj, M. Mongillo, M. Heyns, I. Radu, B. Govoreanu, and A. Potočnik. "Path Toward Manufacturable Superconducting Qubits With Relaxation Times Exceeding 0.1 ms". In: (2022). DOI: 10.48550/ARXIV.2202.10303.

[73] F. A. Mohiyaddin et al. "Multiphysics Simulation AMP; Design of Silicon Quantum Dot Qubit Devices". In: *2019 IEEE International Electron Devices Meeting (IEDM)*. 2019, pp. 39.5.1–39.5.4. DOI: 10.1109/IEDM19573.2019.8993541.

[74] F. N. M. Froning, L. C. Camenzind, O. A. H. van der Molen, A. Li, E. P. A. M. Bakkers, D. M. Zumbühl, and F. R. Braakman. "Ultrafast Hole Spin Qubit With Gate-Tunable Spin–Orbit Switch Functionality". In: *Nature Nanotechnology* 16.3 (Mar. 2021), pp. 308–312. ISSN: 1748-3395. DOI: 10.1038/s41565-020-00828-6.

[75] K. Ono, T. Mori, and S. Moriyama. "High-Temperature Operation of a Silicon Qubit". In: *Scientific Reports* 9.1 (Jan. 2019), p. 469. ISSN: 2045-2322. DOI: 10.1038/s41598-018-36476-z.

[76] R. Maurand, X. Jehl, D. Kotekar-Patil, A. Corna, H. Bohuslavskyi, R. Laviéville, L. Hutin, S. Barraud, M. Vinet, M. Sanquer, and S. De Franceschi. "A CMOS Silicon Spin Qubit". In: *Nature Communications* 7.1 (Nov. 2016), p. 13575. ISSN: 2041-1723. DOI: 10.1038/ncomms13575.

[77] R. Li, N. I. D. Stuyck, S. Kubicek, J. Jussot, B. T. Chan, F. A. Mohiyaddin, A. Elsayed, M. Shehata, G. Simion, C. Godfrin, Y. Canvel, T. Ivanov, L. Goux, B. Govoreanu, and I. P. Radu. "A Flexible 300 mm Integrated Si MOS Platform for Electron- and Hole-Spin Qubits Exploration". In: *2020 IEEE International Electron Devices Meeting (IEDM)*. 2020, pp. 38.3.1–38.3.4. DOI: 10.1109/IEDM13553.2020.9371956.

[78] J. Yoneda, K. Takeda, T. Otsuka, T. Nakajima, M. R. Delbecq, G. Allison, T. Honda, T. Kodera, S. Oda, Y. Hoshi, N. Usami, K. M. Itoh, and S. Tarucha. "A Quantum-Dot Spin Qubit with Coherence Limited by Charge Noise and Fidelity Higher Than 99.9%". In: *Nature Nanotechnology* 13.2 (Feb. 2018), pp. 102–106. ISSN: 1748-3395. DOI: 10.1038/s41565-017-0014-x.

[79] J. M. Elzerman, R. Hanson, L. H. Willems van Beveren, B. Witkamp, L. M. K. Vandersypen, and L. P. Kouwenhoven. "Single-Shot Read-Out of an Individual Electron Spin in a Quantum Dot". In: *Nature* 430.6998 (2004), pp. 431–435. ISSN: 1476-4687. DOI: 10.1038/nature02693.

[80] N. I. D. Stuyck, R. Li, C. Godfrin, A. Elsayed, S. Kubicek, J. Jussot, B. T. Chan, F. A. Mohiyaddin, M. Shehata, G. Simion, Y. Canvel, L. Goux, M. Heyns, B. Govoreanu, and I. P. Radu. "Uniform Spin Qubit Devices with Tunable Coupling in an All-Silicon 300 mm Integrated Process". In: *2021 Symposium*

on VLSI Circuits. 2021, pp. 1–2. DOI: 10.23919/VLSICircuits52068.2021.9492427.

[81] J. C. Bardin et al. "Design and Characterization of a 28-nm Bulk-CMOS Cryogenic Quantum Controller Dissipating Less Than 2 mW at 3 K". In: *IEEE Journal of Solid-State Circuits* 54.11 (2019), pp. 3043–3060. DOI: 10.1109/JSSC.2019.2937234.

[82] A. Ruffino, T.-Y. Yang, J. Michniewicz, Y. Peng, E. Charbon, and M. F. Gonzalez-Zalba. "A Cryo-CMOS Chip that Integrates Silicon Quantum Dots and multiplexed dispersive readout electronics". In: *Nature Electronics* 5.1 (Jan. 2022), pp. 53–59. ISSN: 2520-1131. DOI: 10.1038/s41928-021-00687-6.

[83] L. L. Guevel, G. Billiot, X. Jehl, S. De Franceschi, M. Zurita, Y. Thonnart, M. Vinet, M. Sanquer, R. Maurand, A. G. M. Jansen, and G. Pillonnet. "19.2 A 110mK 295µW 28nm FDSOI CMOS Quantum Integrated Circuit with a 2.8GHz Excitation and nA Current Sensing of an On-Chip Double Quantum Dot". In: *2020 IEEE International Solid- State Circuits Conference - (ISSCC)*. 2020, pp. 306–308. DOI: 10.1109/ISSCC19947.2020.9063090.

[84] R. Barends et al. "Superconducting Quantum Circuits at the Surface Code Threshold for Fault Tolerance". In: *Nature* 508.7497 (2014), pp. 500–503. ISSN: 1476-4687. DOI: 10.1038/nature13171.

[85] J. Benhelm, G. Kirchmair, C. F. Roos, and R. Blatt. "Towards Fault-Tolerant Quantum Computing with Trapped Ions". In: *Nature Physics* 4.6 (2008), pp. 463–466. ISSN: 1745-2481. DOI: 10.1038/nphys961.

[86] J. van Dijk, E. Kawakami, R. Schouten, M. Veldhorst, L. Vandersypen, M. Babaie, E. Charbon, and F. Sebastiano. "Impact of Classical Control Electronics on Qubit Fidelity". In: *Phys. Rev. Applied* 12 (4 Oct. 2019), p. 044054. DOI: 10.1103/PhysRevApplied.12.044054.

[87] A. G. Fowler, M. Mariantoni, J. M. Martinis, and A. N. Cleland. "Surface Codes: Towards Practical Large-Scale Quantum Computation". In: *Phys. Rev. A* 86 (2012), p. 032324. DOI: 10.1103/PhysRevA.86.032324.

[88] H. Homulle, F. Sebastiano, and E. Charbon. "Deep-Cryogenic Voltage References in 40-nm CMOS". In: *IEEE Solid-State Circuits Letters* 1.5 (2018), pp. 110–113. DOI: 10.1109/LSSC.2018.2875821.

[89] J. M. Hornibrook, J. I. Colless, I. D. Conway Lamb, S. J. Pauka, H. Lu, A. C. Gossard, J. D. Watson, G. C. Gardner, S. Fallahi, M. J. Manfra, and D. J. Reilly. "Cryogenic Control Architecture for Large-Scale Quantum Computing". In: *Phys. Rev. Applied* 3 (2 Feb. 2015), p. 024010. DOI: 10.1103/PhysRevApplied.3.024010.

[90] E. Charbon, F. Sebastiano, A. Vladimirescu, H. Homulle, S. Visser, L. Song, and R. M. Incandela. "Cryo-CMOS for Quantum Computing". In: *2016 IEEE International Electron Devices Meeting (IEDM)*. 2016, pp. 13.5.1–13.5.4. DOI: 10.1109/IEDM.2016.7838410.

[91] S. R. Ekanayake, T. Lehmann, A. S. Dzurak, R. G. Clark, and A. Brawley. "Characterization of SOS-CMOS FETs at Low Temperatures for the Design of Integrated Circuits for Quantum Bit Control and Readout". In: *IEEE Transactions on Electron Devices* 57.2 (2010), pp. 539–547. DOI: 10.1109/TED.2009.2037381.

[92] X. Xue et al. "CMOS-Based Cryogenic Control of Silicon Quantum Circuits". In: *Nature* 593.7858 (May 2021), pp. 205–210. ISSN: 1476-4687. DOI: 10.1038/s41586-021-03469-4.

[93] G. Batey, Á. J. Matthews, and M. Patton. "A New Ultra-Low-Temperature Cryogen-Free Experimental Platform". In: *Journal of Physics: Conference Series* 568.3 (2014), p. 032014. DOI: 10.1088/1742-6596/568/3/032014.

[94] M. A. Castellanos-Beltran and K. W. Lehnert. "Widely Tunable Parametric Amplifier Based on a Superconducting Quantum Interference Device Array Resonator". In: *Applied Physics Letters* 91.8 (2007), p. 083509. DOI: 10.1063/1.2773988.

[95] N. Bergeal, F. Schackert, M. Metcalfe, R. Vijay, V. E. Manucharyan, L. Frunzio, D. E. Prober, R. J. Schoelkopf, S. M. Girvin, and M. H. Devoret. "Phase-Preserving Amplification Near the Quantum Limit with a Josephson Ring Modulator". In: *Nature* 465.7294 (2010), pp. 64–68. ISSN: 1476-4687. DOI: 10.1038/nature09035.

[96] M. Hatridge, R. Vijay, D. H. Slichter, J. Clarke, and I. Siddiqi. "Dispersive Magnetometry with a Quantum Limited SQUID Parametric Amplifier". In: *Phys. Rev. B* 83 (13 Apr. 2011), p. 134501. DOI: 10.1103/PhysRevB.83.134501.

[97] K. Geerlings, S. Shankar, E. Edwards, L. Frunzio, R. J. Schoelkopf, and M. H. Devoret. "Improving the Quality Factor of Microwave Compact Resonators by Optimizing Their Geometrical Parameters". In: *Applied Physics Letters* 100.19 (2012), p. 192601. DOI: 10.1063/1.4710520.

[98] J. M. Hornibrook, E. E. Mitchell, and D. J. Reilly. "Suppressing Dissipative Paths in Superconducting Coplanar Waveguide Resonators". In: *IEEE Transactions on Applied Superconductivity* 23.3 (2013), pp. 1501604–1501604. DOI: 10.1109/TASC.2013.2251055.

[99] G. Viola and D. P. DiVincenzo. "Hall Effect Gyrators and Circulators". In: *Phys. Rev. X* 4 (May 2014), p. 021019. DOI: 10.1103/PhysRevX.4.021019.

[100] A. Potočnik, S. Brebels, J. Verjauw, R. Acharya, A. Grill, D. Wan, M. Mongillo, R. Li, T. Ivanov, S. V. Winckel, F. A. Mohiyaddin, B. Govoreanu, J. Craninckx, and I. P. Radu. "Millikelvin Temperature Cryo-CMOS Multiplexer for Scalable Quantum Device Characterisation". In: *Quantum Science and Technology* 7.1 (Oct. 2021), p. 015004. DOI: 10.1088/2058-9565/ac29a1.

[101] D. H. Slichter, O. Naaman, and I. Siddiqi. "Millikelvin Thermal and Electrical Performance of Lossy Transmission Line Filters". In: *Applied Physics Letters* 94.19 (2009), p. 192508. DOI: 10.1063/1.3133362.

[102] J. Wenner, M. Neeley, R. C. Bialczak, M. Lenander, E. Lucero, A. D. OĆonnell, D. Sank, H. Wang, M. Weides, A. N. Cleland, and J. M. Martinis. "Wirebond Crosstalk and Cavity Modes in Large Chip Mounts for Superconducting Qubits". In: *Superconductor Science and Technology* 24.6 (Mar. 2011), p. 065001. DOI: 10.1088/0953-2048/24/6/065001.

[103] M. Rahman and T. Lehmann. "A Cryogenic DAC Operating Down to 4.2K". In: *Cryogenics* 75 (2016), pp. 47–55. ISSN: 0011-2275. DOI: https://doi.org/10.1016/j.cryogenics.2016.02.003.

[104] M. E. P. V. Zurita, L. Le Guevel, G. Billiot, A. Morel, X. Jehl, A. G. M. Jansen, and G. Pillonnet. "Cryogenic Current Steering DAC With Mitigated Variability". In: *IEEE Solid-State Circuits Letters* 3 (2020), pp. 254–257. DOI: 10.1109/LSSC.2020.3013443.

[105] P. Vliex, C. Degenhardt, C. Grewing, A. Kruth, D. Nielinger, S. van Waasen, and S. Heinen. "Bias Voltage DAC Operating at Cryogenic Temperatures for Solid-State Qubit Applications". In: *IEEE Solid-State Circuits Letters* 3 (2020), pp. 218–221. DOI: 10.1109/LSSC.2020.3011576.

[106] B. Patra et al. "19.1 A Scalable Cryo-CMOS 2-to-20GHz Digitally Intensive Controller for 4×32 Frequency Multiplexed Spin Qubits/Transmons in 22nm FinFET Technology for Quantum Computers". In: *2020 IEEE International Solid- State Circuits Conference - (ISSCC)*. 2020, pp. 304–306. DOI: 10.1109/ISSCC19947.2020.9063109.

[107] R. Saligram, W. Chakraborty, N. Cao, Y. Cao, S. Datta, and A. Raychowdhury. "Power Performance Analysis of Digital Standard Cells for 28 nm Bulk CMOS at Cryogenic Temperature Using BSIM Models". In: *IEEE Journal on Exploratory Solid-State Computational Devices and Circuits* 7.2 (2021), pp. 193–200. DOI: 10.1109/JXCDC.2021.3131100.

[108] R. Wang, S. Guo, Z. Zhang, Q. Wang, D. Wu, J. Wang, and R. Huang. "Too Noisy at the Bottom? —Random Telegraph Noise (RTN) in Advanced Logic Devices and Circuits". In: *2018 IEEE International Electron Devices Meeting (IEDM)*. 2018, pp. 17.2.1–17.2.4. DOI: 10.1109/IEDM.2018.8614594.

[109] A. Beckers, F. Jazaeri, and C. Enz. "Characterization and Modeling of 28-nm Bulk CMOS Technology Down to 4.2 K". In: *IEEE Journal of the Electron Devices Society* 6 (2018), pp. 1007–1018. DOI: 10.1109/JEDS.2018.2817458.

[110] A. Beckers, F. Jazaeri, and C. Enz. "Theoretical Limit of Low Temperature Subthreshold Swing in Field-Effect Transistors". In: *IEEE Electron Device Letters* 41.2 (2020), pp. 276–279. DOI: 10.1109/LED.2019.2963379.

[111] A. Beckers, F. Jazaeri, A. Grill, S. Narasimhamoorthy, B. Parvais, and C. Enz. "Physical Model of Low-Temperature to Cryogenic Threshold Voltage in MOSFETs". In: *IEEE Journal of the Electron Devices Society* 8 (2020), pp. 780–788. DOI: 10.1109/JEDS.2020.2989629.

[112] A. Beckers, F. Jazaeri, A. Ruffino, C. Bruschini, A. Baschirotto, and C. Enz. "Cryogenic Characterization of 28 nm Bulk CMOS Technology for Quantum Computing". In: *2017 47th European Solid-State Device Research Conference (ESSDERC)*. 2017, pp. 62–65. DOI: 10.1109/ESSDERC.2017.8066592.

[113] T. Grasser, ed. *Bias Temperature Instability for Devices and Circuits*. Springer Science+Business Media New York, 2013. ISBN: 978-1-4614-7908-6. DOI: 10.1007/978-1-4614-7909-3.

[114] T. Grasser, B. Kaczer, W. Gös, H. Reisinger, T. Aichinger, P. P. Hehenberger, P.-J. Wagner, F. Schanovsky, J. Franco, M. Toledano-Luque, and M. Nelhiebel. "The Paradigm Shift in Understanding the Bias Temperature Instability: From Reaction-Diffusion to Switching Oxide Traps". In: *IEEE Transactions on Electron Devices* 58.11 (2011), pp. 3652–3666.

[115] J. H. Stathis, S. Mahapatra, and T. Grasser. "Controversial Issues in Negative Bias Temperature Instability". In: *Microelectronics Reliability* 81 (Dec. 2018), pp. 244–251. DOI: 10.1016/j.microrel.2017.12.035.

[116] T. Grasser, ed. *Noise in Nanoscale Semiconductor Devices*. Springer Science + Business Media New York, 2020. ISBN: 978-3-030-37499-0. DOI: 10.1007/978-3-030-37500-3.

[117] M. Kirton and M. Uren. "Noise in Solid-State Microstructures: A New Perspective on Individual Defects, Interface States and Low-Frequency ($1/f$) Noise". In: *Advances in Physics* 38.4 (1989), pp. 367–468. DOI: 10.1080/00018738900101122.

[118] M. J. Uren, M. J. Kirton, and S. Collins. "Anomalous Telegraph Noise in Small-Area Silicon Metal-Oxide-Semiconductor Field-Effect Transistors". In: *Phys. Rev. B* 37 (14 May 1988), pp. 8346–8350. DOI: 10.1103/PhysRevB.37.8346.

[119] S. Tyaginov and T. Grasser. "Modeling of Hot-Carrier Degradation: Physics and Controversial Issues". In: *2012 IEEE International Integrated Reliability Workshop Final Report*. 2012, pp. 206–215. DOI: 10.1109/IIRW.2012.6468962.

[120] T. Grasser, ed. *Hot Carrier Degradation in Semiconductor Devices*. Springer International Publishing, 2014. ISBN: 978-3-319-08993-5. DOI: 10.1007/978-3-319-08994-2.

[121] R. Huang, X. B. Jiang, S. F. Guo, P. P. Ren, P. Hao, Z. Q. Yu, Z. Zhang, Y. Y. Wang, and R. S. Wang. "Variability-and Reliability-Aware Design for 16/14 nm and Beyond Technology". In: *2017 IEEE International Electron Devices Meeting (IEDM)*. 2017, pp. 12.4.1–12.4.4. DOI: 10.1109/IEDM.2017.8268378.

[122] G. Wolfowicz, F. J. Heremans, C. P. Anderson, S. Kanai, H. Seo, A. Gali, G. Galli, and D. D. Awschalom. "Quantum Guidelines for Solid-State Spin Defects". In: *Nature Reviews Materials* 6.10 (Oct. 2021), pp. 906–925. ISSN: 2058-8437. DOI: 10.1038/s41578-021-00306-y.

[123] J. M. Martinis, S. Nam, J. Aumentado, K. M. Lang, and C. Urbina. "Decoherence of a Superconducting Qubit Due to Bias Noise". In: *Phys. Rev. B* 67 (9 Mar. 2003), p. 094510. DOI: 10.1103/PhysRevB.67.094510.

[124] D. J. Van Harlingen, T. L. Robertson, B. L. T. Plourde, P. A. Reichardt, T. A. Crane, and J. Clarke. "Decoherence in Josephson-Junction Qubits Due to Critical-Current Fluctuations". In: *Phys. Rev. B* 70 (6 Aug. 2004), p. 064517. DOI: 10.1103/PhysRevB.70.064517.

[125] R. McDermott. "Materials Origins of Decoherence in Superconducting Qubits". In: *IEEE Transactions on Applied Superconductivity* 19.1 (2009), pp. 2–13. DOI: 10.1109/TASC.2008.2012255.

[126] D. Culcer, X. Hu, and S. Das Sarma. "Dephasing of Si Spin Qubits Due to Charge Noise". In: *Applied Physics Letters* 95.7 (2009), p. 073102. DOI: 10.1063/1.3194778.

[127] Y.-C. Yang, S. N. Coppersmith, and M. Friesen. "High-Fidelity Single-Qubit Gates in a Strongly Driven Quantum-Dot Hybrid Qubit with $1/f$ Charge Noise". In: *Phys. Rev. A* 100 (2 Aug. 2019), p. 022337. DOI: 10.1103/PhysRevA.100.022337.

[128] E. J. Connors, J. Nelson, H. Qiao, L. F. Edge, and J. M. Nichol. "Low-Frequency Charge Noise in Si/SiGe Quantum Dots". In: *Phys. Rev. B* 100 (Oct. 2019), p. 165305. DOI: 10.1103/PhysRevB.100.165305.

[129] L. Petit, J. M. Boter, H. G. J. Eenink, G. Droulers, M. L. V. Tagliaferri, R. Li, D. P. Franke, K. J. Singh, J. S. Clarke, R. N. Schouten, V. V. Dobrovitski, L. M. K. Vandersypen, and M. Veldhorst. "Spin Lifetime and Charge Noise in Hot Silicon Quantum Dot Qubits". In: *Phys. Rev. Lett.* 121 (7 Aug. 2018), p. 076801. DOI: 10.1103/PhysRevLett.121.076801.

[130] P. Huang, N. M. Zimmerman, and G. W. Bryant. "Spin Decoherence in a Two-Qubit CPHASE Gate: the Critical Role of Tunneling Noise". In: *npj Quantum Information* 4.1 (Nov. 2018), p. 62. ISSN: 2056-6387. DOI: 10.1038/s41534-018-0112-0.

[131] C. H. Yang, K. W. Chan, R. Harper, W. Huang, T. Evans, J. C. C. Hwang, B. Hensen, A. Laucht, T. Tanttu, F. E. Hudson, S. T. Flammia, K. M. Itoh, A. Morello, S. D. Bartlett, and A. S. Dzurak. "Silicon Qubit Fidelities Approaching Incoherent Noise Limits via Pulse Engineering". In: *Nature Electronics* 2.4 (Apr. 2019), pp. 151–158. ISSN: 2520-1131. DOI: 10.1038/s41928-019-0234-1.

[132] N. E. Penthorn, J. S. Schoenfield, J. D. Rooney, and H. W. Jiang. "Reduction of Charge Impurities in a Silicon Metal-Oxide-Semiconductor Quantum Dot Qubit Device Patterned with Nano-Imprint Lithography". In: *Nanotechnology* 30.46 (Sept. 2019), p. 465302. DOI: 10.1088/1361-6528/ab3cb9.

[133] K. Ghosh, H. Ma, M. Onizhuk, V. Gavini, and G. Galli. "Spin–Spin Interactions in Defects in Solids from Mixed All-Electron and Pseudopotential First-Principles Calculations". In: *npj Computational Materials* 7.1 (July 2021), p. 123. ISSN: 2057-3960. DOI: 10.1038/s41524-021-00590-w.

[134] J. Xu, A. Habib, S. Kumar, F. Wu, R. Sundararaman, and Y. Ping. "Spin-Phonon Relaxation from a Universal Ab Initio Density-Matrix Approach". In: *Nature Communications* 11.1 (June 2020), p. 2780. ISSN: 2041-1723. DOI: 10.1038/s41467-020-16063-5.

[135] J. Xu, A. Habib, R. Sundararaman, and Y. Ping. "Ab Initio Ultrafast Spin Dynamics in Solids". In: *Phys. Rev. B* 104 (Nov. 2021), p. 184418. DOI: 10.1103/PhysRevB.104.184418.

[136] G. Rzepa, J. Franco, B. O'Sullivan, A. Subirats, M. Simicic, G. Hellings, P. Weckx, M. Jech, T. Knobloch, M. Waltl, P. Roussel, D. Linten, B. Kaczer, and T. Grasser. "Comphy — A Compact-Physics Framework for Unified Modeling of BTI". In: *Microelectronics Reliability* 85 (2018), pp. 49–65. ISSN: 0026-2714. DOI: https://doi.org/10.1016/j.microrel.2018.04.002.

[137] M. L. Reed and J. D. Plummer. "Chemistry of Si-SiO_2 Interface Trap Annealing". In: *Journal of Applied Physics* 63.12 (1988), pp. 5776–5793. DOI: 10.1063/1.340317.

[138] Y. Nishi. "Study of Silicon-Silicon Dioxide Structure by Electron Spin Resonance I". In: *Japanese Journal of Applied Physics* 10.1 (Jan. 1971), pp. 52–62. DOI: 10.1143/jjap.10.52.

[139] N. H. Thoan, K. Keunen, V. V. Afanas'ev, and A. Stesmans. "Interface State Energy Distribution and P_b Defects at Si(110)/SiO_2 Interfaces: Comparison to (111) and (100) Silicon Orientations". In: *Journal of Applied Physics* 109.1 (2011), p. 013710. DOI: 10.1063/1.3527909.

[140] B. Stampfer. "Advanced Electrical Characterization of Charge Trapping in MOS Transistors". PhD thesis. Institut für Mikroelektronik, 2020. DOI: 10.34726/hss.2020.86423.

[141] J. P. Campbell and P. M. Lenahan. "Density of States of P_{b1} Si/SiO_2 Interface Trap Centers". In: *Applied Physics Letters* 80.11 (2002), pp. 1945–1947. DOI: 10.1063/1.1461053.

[142] B. Ruch, M. Jech, G. Pobegen, and T. Grasser. "Applicability of Shockley-Read-Hall Theory for Interface States". In: *2020 IEEE International Electron Devices Meeting (IEDM)*. 2020, pp. 22.1.1–22.1.4. DOI: 10.1109/IEDM13553.2020.9372032.

[143] K. L. Yip and W. B. Fowler. "Electronic Structure of E_1' Centers in SiO_2". In: *Phys. Rev. B* 11 (6 Mar. 1975), pp. 2327–2338. DOI: 10.1103/PhysRevB.11.2327.

[144] E. P. O'Reilly and J. Robertson. "Theory of Defects in Vitreous Silicon Dioxide". In: *Phys. Rev. B* 27 (6 Mar. 1983), pp. 3780–3795. DOI: 10.1103/PhysRevB.27.3780.

[145] J. K. Rudra, W. B. Fowler, and F. J. Feigl. "Model for the E_2' Center in Alpha Quartz". In: *Phys. Rev. Lett.* 55 (23 Dec. 1985), pp. 2614–2617. DOI: 10.1103/PhysRevLett.55.2614.

[146] S. Mukhopadhyay, P. V. Sushko, A. M. Stoneham, and A. L. Shluger. "Modeling of the Structure and Properties of Oxygen Vacancies in Amorphous Silica". In: *Phys. Rev. B* 70 (19 Nov. 2004), p. 195203. DOI: 10.1103/PhysRevB.70.195203.

[147] F. Schanovsky, O. Baumgartner, W. Goes, and T. Grasser. "A Detailed Evaluation of Model Defects as Candidates for the Bias Temperature Instability". In: *2013 International Conference on Simulation of Semiconductor Processes and Devices (SISPAD)*. 2013, pp. 1–4. DOI: 10.1109/SISPAD.2013.6650559.

[148] T. Grasser, W. Goes, Y. Wimmer, F. Schanovsky, G. Rzepa, M. Waltl, K. Rott, H. Reisinger, V. Afanas'ev, A. Stesmans, A.-M. El-Sayed, and A. Shluger. "On the Microscopic Structure of Hole Traps in pMOSFETs". In: *2014 IEEE International Electron Devices Meeting*. 2014, pp. 21.1.1–21.1.4. DOI: 10.1109/IEDM.2014.7047093.

[149] Y. Wimmer, A.-M. El-Sayed, W. Gös, T. Grasser, and A. L. Shluger. "Role of Hydrogen in Volatile Behaviour of Defects in SiO_2-Based Electronic Devices". In: *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences* 472.2190 (2016), p. 20160009. DOI: 10.1098/rspa.2016.0009.

[150] A. T. Krishnan, S. Chakravarthi, P. Nicollian, V. Reddy, and S. Krishnan. "Negative Bias Temperature Instability Mechanism: The Role of Molecular Hydrogen". In: *Applied Physics Letters* 88.15 (2006), p. 153518. DOI: 10.1063/1.2191828.

[151] T. Aichinger, S. Puchner, M. Nelhiebel, T. Grasser, and H. Hutter. "Impact of Hydrogen on Recoverable and Permanent Damage Following Negative Bias Temperature Stress". In: *2010 IEEE International Reliability Physics Symposium*. 2010, pp. 1063–1068. DOI: 10.1109/IRPS.2010.5488672.

[152] K. Kajihara, L. Skuja, M. Hirano, and H. Hosono. "Diffusion and Reactions of Hydrogen in F_2-Laser-Irradiated SiO_2 Glass". In: *Phys. Rev. Lett.* 89 (13 Sept. 2002), p. 135507. DOI: 10.1103/PhysRevLett.89.135507.

[153] J. Godet and A. Pasquarello. "Proton Diffusion Mechanism in Amorphous SiO_2". In: *Phys. Rev. Lett.* 97 (15 Oct. 2006), p. 155901. DOI: 10.1103/PhysRevLett.97.155901.

[154] N. S. Saks and A. K. Agarwal. "Hall Mobility and Free Electron Density at the SiC/SiO_2 Interface in 4H-SiC". In: *Applied Physics Letters* 77.20 (2000), pp. 3281–3283. DOI: 10.1063/1.1326046.

[155] V. V. Afanas′ev and A. Stesmans. "Interfacial Defects in SiO_2 Revealed by Photon Stimulated Tunneling of Electrons". In: *Phys. Rev. Lett.* 78 (1997), pp. 2437–2440. DOI: 10.1103/PhysRevLett.78.2437.

[156] A.-M. El-Sayed, M. B. Watkins, V. V. Afanas′ev, and A. L. Shluger. "Nature of Intrinsic and Extrinsic Electron Trapping in SiO_2". In: *Phys. Rev. B* 89 (2014), p. 125201. DOI: 10.1103/PhysRevB.89.125201.

[157] J. P. Campbell and P. M. Lenahan. "Atomic-Scale Defects Associated with the Negative Bias Temperature Instability". In: *Bias Temperature Instability for Devices and Circuits*. Ed. by T. Grasser. Springer New York, 2014, pp. 177–228. ISBN: 978-1-4614-7909-3. DOI: 10.1007/978-1-4614-7909-3_8.

[158] J. P. Campbell, P. M. Lenahan, A. T. Krishnan, and S. Krishnan. "Identification of Atomic-Scale Defect Structure Involved in the Negative Bias Temperature Instability in Plasma-Nitrided Devices". In: *Applied Physics Letters* 91.13 (2007), p. 133507. DOI: 10.1063/1.2790776.

[159] J. P. Campbell, P. M. Lenahan, C. J. Cochrane, A. T. Krishnan, and S. Krishnan. "Atomic-Scale Defects Involved in the Negative-Bias Temperature Instability". In: *IEEE Transactions on Device and Materials Reliability* 7.4 (2007), pp. 540–557. DOI: 10.1109/TDMR.2007.911379.

[160] J. P. Campbell, P. M. Lenahan, A. T. Krishnan, and S. Krishnan. "Identification of the Atomic-Scale Defects Involved in the Negative Bias Temperature Instability in Plasma-Nitrided p-Channel Metal-Oxide-Silicon Field-Effect Transistors". In: *Journal of Applied Physics* 103.4 (2008), p. 044505. DOI: 10.1063/1.2844348.

[161] P. M. Lenahan and S. E. Curry. "First Observation of the ^{29}Si Hyperfine Spectra of Silicon Dangling Bond Centers in Silicon Nitride". In: *Applied Physics Letters* 56.2 (1990), pp. 157–159. DOI: 10.1063/1.103278.

[162] W. L. Warren and P. M. Lenahan. "Electron-Nuclear Double-Resonance and Electron-Spin-Resonance Study of Silicon Dangling-Bond Centers in Silicon Nitride". In: *Phys. Rev. B* 42 (3 July 1990), pp. 1773–1780. DOI: 10.1103/PhysRevB.42.1773.

[163] D. Griscom, E. Friebele, and G. Sigel. "Observation and Analysis of the Primary ^{29}Si Hyperfine Structure of the E′ Center in Non-Crystalline SiO_2". In: *Solid State Communications* 15.3 (1974), pp. 479–483. ISSN: 0038-1098. DOI: https://doi.org/10.1016/0038-1098(74)91124-7.

[164] D. Waldhör, A.-M. El-Sayed, Y. Wimmer, M. Waltl, and T. Grasser. "Atomistic Modeling of Oxide Defects". In: *Noise in Nanoscale Semiconductor Devices*. Ed. by T. Grasser. Springer International Publishing, 2020, pp. 609–648. ISBN: 978-3-030-37499-0. DOI: 10.1007/978-3-030-37500-3{_}18.

[165] J. Strand, M. Kaviani, D. Gao, A.-M. El-Sayed, V. V. Afanasév, and A. L. Shluger. "Intrinsic Charge Trapping in Amorphous Oxide Films: Status and Challenges". In: *Journal of Physics: Condensed Matter* 30.23 (May 2018), p. 233001. DOI: 10.1088/1361-648x/aac005.

[166] F. Bohra, B. Jiang, and J.-M. Zuo. "Textured Crystallization of Ultrathin Hafnium Oxide Films on Silicon Substrate". In: *Applied physics letters* 90.16 (2007), p. 161917.

[167] M. Kaviani, J. Strand, V. V. Afanas'ev, and A. L. Shluger. "Deep Electron and Hole Polarons and Bipolarons in Amorphous Oxide". In: *Phys. Rev. B* 94 (2 July 2016), p. 020103. DOI: 10.1103/PhysRevB.94.020103.

[168] F. Cerbu, O. Madia, D. V. Andreev, S. Fadida, M. Eizenberg, L. Breuil, J. G. Lisoni, J. A. Kittl, J. Strand, A. L. Shluger, V. V. Afanas'ev, M. Houssa, and A. Stesmans. "Intrinsic Electron Traps in Atomic-Layer Deposited HfO_2 Insulators". In: *Applied Physics Letters* 108.22 (2016), p. 222901. DOI: 10.1063/1.4952718.

[169] A. Kerber and E. A. Cartier. "Reliability Challenges for CMOS Technology Qualifications With Hafnium Oxide/Titanium Nitride Gate Stacks". In: *IEEE Transactions on Device and Materials Reliability* 9.2 (2009), pp. 147–162. DOI: 10.1109/TDMR.2009.2016954.

[170] B. Govoreanu, R. Degraeve, M. Zahid, L. Nyns, M. Cho, B. Kaczer, M. Jurczak, J. Kittl, and J. Van Houdt. "Understanding the Potential and Limitations of HfAlO as Interpoly Dielectric in Floating-Gate Flash Memory". In: *Microelectronic Engineering* 86.7 (2009). INFOS 2009, pp. 1807–1811. ISSN: 0167-9317. DOI: https://doi.org/10.1016/j.mee.2009.03.099.

[171] R. N. Hall. "Electron-Hole Recombination in Germanium". In: *Phys. Rev.* 87 (2 July 1952), pp. 387–387. DOI: 10.1103/PhysRev.87.387.

[172] W. Shockley and W. T. Read. "Statistics of the Recombinations of Holes and Electrons". In: *Phys. Rev.* 87 (5 Sept. 1952), pp. 835–842. DOI: 10.1103/PhysRev.87.835.

[173] A. L. McWhorter et al. "1/f Noise and Related Surface Effects in Germanium." In: (1955).

[174] T. Grasser. "Stochastic Charge Trapping in Oxides: From Random Telegraph Noise to Bias Temperature Instabilities". In: *Microelectronics Reliability* 52.1 (2012), pp. 39–70. DOI: 10.1016/j.microrel.2011.09.002.

[175] J. Franck and E. G. Dymond. "Elementary Processes of Photochemical Reactions". In: *Trans. Faraday Soc.* 21 (February 1926), pp. 536–542. DOI: 10.1039/TF9262100536.

[176] E. Condon. "A Theory of Intensity Distribution in Band Systems". In: *Phys. Rev.* 28 (6 Dec. 1926), pp. 1182–1201. DOI: 10.1103/PhysRev.28.1182.

[177] M. Lax. "The Franck-Condon Principle and Its Application to Crystals". In: *The Journal of Chemical Physics* 20.11 (1952), pp. 1752–1760. DOI: 10.1063/1.1700283.

[178] W. Nolting. *Grundkurs: Theoretische Physik: Quantenmechanik*. Springer, 1997.

[179] R. H. Fowler and L. Nordheim. "Electron Emission in Intense Electric Fields". In: *Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character* 119.781 (1928), pp. 173–181. DOI: 10.1098/rspa.1928.0091.

[180] G. Rzepa. "Efficient Physical Modeling of Bias Temperature Instability". PhD thesis. Institut für Mikroelektronik, 2018. DOI: 10.34726/hss.2018.57326.

[181] A. Alkauskas, Q. Yan, and C. G. Van de Walle. "First-Principles Theory of Nonradiative Carrier Capture via Multiphonon Emission". In: *Phys. Rev. B* 90 (7 Aug. 2014), p. 075202. DOI: 10.1103/PhysRevB.90.075202.

[182] M. E. Turiansky, A. Alkauskas, M. Engel, G. Kresse, D. Wickramaratne, J.-X. Shen, C. E. Dreyer, and C. G. Van de Walle. "Nonrad: Computing Nonradiative Capture Coefficients from First Principles". In: *Computer Physics Communications* 267 (2021), p. 108056. ISSN: 0010-4655. DOI: https://doi.org/10.1016/j.cpc.2021.108056.

[183] R. Gross, A. Marx, D. Einzel, and S. Geprägs. *Festkörperphysik: Aufgaben und Lösungen*. De Gruyter, 2018. ISBN: 9783110566130. DOI: doi:10.1515/9783110566130.

[184] P. P. Schmidt. "Computationally Efficient Recurrence Relations for One- Dimensional Franck–Condon Overlap Integrals". In: *Molecular Physics* 108.11 (2010), pp. 1513–1529. DOI: 10.1080/00268971003762142.

[185] W. Gös, Y. Wimmer, A.-M. El-Sayed, G. Rzepa, M. Jech, A. L. Shluger, and T. Grasser. "Identification of Oxide Defects in Semiconductor Devices: A Systematic Approach Linking DFT to Rate Equations and Experimental Evidence". In: *Microelectronics Reliability* 87 (2018), pp. 286–320. DOI: 10.1016/j.microrel.2017.12.021.

[186] K. Huang and A. Rhys. "Theory of Light Absorption and Non-Radiative Transitions in *F*-Centres". In: *Selected Papers of Kun Huang*, pp. 74–92. DOI: 10.1142/9789812793720_0007.

[187] T. Holstein. "Quantal Occurrence-Probability Treatment of Small-Polaron Hopping". In: *Philosophical Magazine B* 37.1 (1978), pp. 49–62. DOI: 10.1080/13642817808245306.

[188] T. Markvart. "Determination of Potential Surfaces From Multiphonon Transition Rates". In: *Journal of Physics C: Solid State Physics* 14.15 (May 1981), pp. L435–L440. DOI: 10.1088/0022-3719/14/15/002.

[189] T. Markvart. "Semiclassical Theory of Non-Radiative Transitions". In: *Journal of Physics C: Solid State Physics* 14.29 (Oct. 1981), pp. L895–L899. DOI: 10.1088/0022-3719/14/29/006.

[190] T. Markvart. "Multiphonon Transitions Between Adiabatic Potential Curves". In: *Journal of Physics C: Solid State Physics* 17.35 (1984), pp. 6303–6316. ISSN: 00223719. DOI: 10.1088/0022-3719/17/35/006.

[191] J. Mathews, R. Walker, and W. L. *Mathematical Methods of Physics*. Addison-Wesley world student series. W. A. Benjamin, 1970. ISBN: 9780805370027.

[192] A. Nitzan. *Chemical Dynamics in Condensed Phases: Relaxation, Transfer and Reactions in Condensed Molecular Systems*. Oxford university press, 2006. ISBN: 9780198529798. DOI: 10.1093/oso/9780198529798.001.0001.

[193] K. F. Freed and J. Jortner. "Multiphonon Processes in the Nonradiative Decay of Large Molecules". In: *The Journal of Chemical Physics* 52.12 (1970), pp. 6272–6291. DOI: 10.1063/1.1672938.

[194] R. Wang, S. Guo, Z. Zhang, J. Zou, D. Mao, and R. Huang. "Complex Random Telegraph Noise (RTN): What Do We Understand?" In: *2018 IEEE International Symposium on the Physical and Failure Analysis of Integrated Circuits (IPFA)*. 2018, pp. 1–7. DOI: 10.1109/IPFA.2018.8452514.

[195] H. Ihantola and J. Moll. "Design Theory of a Surface Field-Effect Transistor". In: *Solid-State Electronics* 7.6 (1964), pp. 423–430. ISSN: 0038-1101. DOI: https://doi.org/10.1016/0038-1101(64)90039-5.

[196] C.-T. Sah. "Characteristics of the Metal-Oxide-Semiconductor Transistors". In: *IEEE Transactions on Electron Devices* 11.7 (1964), pp. 324–345. DOI: 10.1109/T-ED.1964.15336.

[197] H. Pao and C. Sah. "Effects of Diffusion Current on Characteristics of Metal-Oxide (Insulator)-Semiconductor Transistors". In: *Solid-State Electronics* 9.10 (1966), pp. 927–937. ISSN: 0038-1101. DOI: https://doi.org/10.1016/0038-1101(66)90068-2.

[198] C. C. McAndrew. "Compact Models for MOS Transistors: Successes and Challenges". In: *IEEE Transactions on Electron Devices* 66.1 (2019), pp. 12–18. DOI: 10.1109/TED.2018.2849943.

[199] J. Brews. "A Charge-Sheet Model of the MOSFET". In: *Solid-State Electronics* 21.2 (1978), pp. 345–355. ISSN: 0038-1101. DOI: https://doi.org/10.1016/0038-1101(78)90264-2.

[200] A. Beckers. *Cryogenic MOSFET Modeling for Large-Scale Quantum Computing*. Tech. rep. EPFL, 2021.

[201] G. Gildenblat, X. Li, W. Wu, H. Wang, A. Jha, R. Van Langevelde, G. Smit, A. Scholten, and D. Klaassen. "PSP: An Advanced Surface-Potential-Based MOSFET Model for Circuit Simulation". In: *IEEE Transactions on Electron Devices* 53.9 (2006), pp. 1979–1993. DOI: 10.1109/TED.2005.881006.

[202] C. C. Enz, F. Krummenacher, and E. A. Vittoz. "An Analytical MOS Transistor Model Valid in All Regions of Operation and Dedicated to Low-Voltage and Low-Current Applications". In: *Analog Integrated Circuits and Signal Processing* 8.1 (1995), pp. 83–114. ISSN: 1573-1979. DOI: 10.1007/BF01239381.

[203] C. Galup-Montoro, M. C. Schneider, A. I. A. Cunha, F. R. de Sousa, H. Klimach, and O. F. Siebel. "The Advanced Compact MOSFET (ACM) Model for Circuit Analysis and Design". In: *2007 IEEE Custom Integrated Circuits Conference*. 2007, pp. 519–526. DOI: 10.1109/CICC.2007.4405785.

[204] Y. S. Chauhan, S. Venugopalan, M.-A. Chalkiadaki, M. A. U. Karim, H. Agarwal, S. Khandelwal, N. Paydavosi, J. P. Duarte, C. C. Enz, A. M. Niknejad, and C. Hu. "BSIM6: Analog and RF Compact Model for Bulk MOSFET". In: *IEEE Transactions on Electron Devices* 61.2 (2014), pp. 234–244. DOI: 10.1109/TED.2013.2283084.

[205] S. M. Sze, Y. Li, and K. K. Ng. *Physics of Semiconductor Devices*. John Wiley & sons, 2021.

[206] W. B. Joyce and R. W. Dixon. "Analytic Approximations for the Fermi Energy of an Ideal Fermi Gas". In: *Applied Physics Letters* 31.5 (1977), pp. 354–356. DOI: 10.1063/1.89697.

[207] W. Bludau, A. Onton, and W. Heinke. "Temperature Dependence of the Band Gap of Silicon". In: *Journal of Applied Physics* 45.4 (1974), pp. 1846–1848. DOI: 10.1063/1.1663501.

[208] J. E. Lang, F. L. Madarasz, and P. M. Hemenger. "Temperature Dependent Density of States Effective Mass in Nonparabolic p-Type Silicon". In: *Journal of Applied Physics* 54.6 (1983), pp. 3612–3612. DOI: 10.1063/1.332397.

[209] M. A. Green. "Intrinsic Concentration, Effective Densities of States, and Effective Mass in Silicon". In: *Journal of Applied Physics* 67.6 (1990), pp. 2944–2954. DOI: 10.1063/1.345414.

[210] B. Kaczer, T. Grasser, P. J. Roussel, J. Franco, R. Degraeve, L.-A. Ragnarsson, E. Simoen, G. Groeseneken, and H. Reisinger. "Origin of NBTI Variability in Deeply Scaled pFETs". In: *2010 IEEE International Reliability Physics Symposium*. 2010, pp. 26–32. DOI: 10.1109/IRPS.2010.5488856.

[211] A. N. Tikhonov and V. Y. Arsenin. *Solutions of Ill-Posed Problems*. V. H. Winston & Sons, 1977, pp. xiii+258.

[212] C. Schleich, D. Waldhör, K. A. Waschneck, M. Feil, H. Reisinger, T. Grasser, and M. Waltl. "Physical Modeling of Charge Trapping in 4H-SiC DMOSFET Technologies". In: *IEEE Transactions on Electron Devices* 68.8 (2021), pp. 4016–4021. DOI: 10.1109/TED.2021.3092295.

[213] M. Waltl. "Ultra-Low Noise Defect Probing Instrument for Defect Spectroscopy of MOS Transistors". In: *IEEE Transactions on Device and Materials Reliability* 20.2 (2020), pp. 242–250. DOI: 10.1109/TDMR.2020.2988650.

[214] M. Waltl. "Experimental Characterization of Bias Temperature Instabilities in Modern Transistor Technologies". PhD thesis. Institut für Mikroelektronik, 2016. DOI: 10.34726/hss.2016.38201.

[215] C. S. Chen, L. Lil, Q. Lim, H. H. Teh, N. F. Binti Omar, C. L. Ler, and J. T. Watt. "A Compact Array for Characterizing 32k Transistors in Wafer Scribe Lanes". In: *2014 International Conference on Microelectronic Test Structures (ICMTS)*. 2014, pp. 227–232. DOI: 10.1109/ICMTS.2014.6841497.

[216] M. Simicic, V. Putcha, B. Parvais, P. Weckx, B. Kaczer, G. Groeseneken, G. Gielen, D. Linten, and A. Thean. "Advanced MOSFET Variability and Reliability characterization array". In: *2015 IEEE International Integrated Reliability Workshop (IIRW)*. 2015, pp. 73–76. DOI: 10.1109/IIRW.2015.7437071.

[217] E. Bury, B. Kaczer, K. Chuang, J. Franco, P. Weckx, A. Chasin, M. Simicic, D. Linten, and G. Groeseneken. "Statistical Assessment of the Full V_G/V_D Degradation Space Using Dedicated Device Arrays". In: *2017 IEEE International Reliability Physics Symposium (IRPS)*. 2017, pp. 2D-5.1-2D–5.6. DOI: 10.1109/IRPS.2017.7936265.

[218] K.-H. Chuang, E. Bury, R. Degraeve, B. Kaczer, G. Groeseneken, I. Verbauwhede, and D. Linten. "Physically Unclonable Function Using CMOS Breakdown Position". In: *2017 IEEE International Reliability Physics Symposium (IRPS)*. 2017, pp. 4C-1.1-4C–1.7. DOI: 10.1109/IRPS.2017.7936312.

[219] E. Bury, A. Chasin, M. Vandemaele, S. Van Beek, J. Franco, B. Kaczer, and D. Linten. "Array-Based Statistical Characterization of CMOS Degradation Modes and Modeling of the Time-Dependent Variability Induced by Different Stress Patterns in the $\{V_G, V_D\}$ Bias Space". In: *2019 IEEE International Reliability Physics Symposium (IRPS)*. 2019, pp. 1–6. DOI: 10.1109/IRPS.2019.8720592.

[220] Y. Taur and T. H. N. *Fundamentals of Modern VLSI Devices*. 2nd ed. Cambridge University Press, 2009. DOI: 10.1017/CBO9781139195065.

[221] A. Beckers, D. Beckers, F. Jazaeri, B. Parvais, and C. Enz. "Generalized Boltzmann Relations in Semiconductors Including Band Tails". In: *Journal of Applied Physics* 129.4 (2021), p. 045701. DOI: 10.1063/5.0037432.

[222] T. Grasser, B. J. O'Sullivan, B. Kaczer, J. Franco, B. Stampfer, and M. Waltl. "CV Stretch-Out Correction after Bias Temperature Stress: Work-function Dependence of Donor-/Acceptor-like Traps, Fixed Charges, and Fast States". In: *Proceedings of the IEEE International Reliability Physics Symposium (IRPS)*. 2021, pp. 1–6. ISBN: 978-1-7281-6893-7. DOI: 10.1109/IRPS46558.2021.9405184.

[223] P. A. 'T Hart, M. Babaie, E. Charbon, A. Vladimirescu, and F. Sebastiano. "Characterization and Modeling of Mismatch in Cryo-CMOS". In: *IEEE Journal of the Electron Devices Society* 8 (2020), pp. 263–273. DOI: 10.1109/JEDS.2020.2976546.

[224] H. Yang, M. Robitaille, X. Chen, H. Elgabra, L. Wei, and N. Y. Kim. "Random Telegraph Noise of a 28-nm Cryogenic MOSFET in the Coulomb Blockade Regime". In: *IEEE Electron Device Letters* 43.1 (2022), pp. 5–8. DOI: 10.1109/LED.2021.3132964.

[225] Y. Ono, K. Nishiguchi, A. Fujiwara, H. Yamaguchi, H. Inokawa, and Y. Takahashi. "Conductance Modulation by Individual Acceptors in Si Nanoscale Field-Effect Transistors". In: *Applied Physics Letters* 90.10 (2007), p. 102106. DOI: 10.1063/1.2679254.

[226] S. Bonen et al. "Cryogenic Characterization of 22-nm FDSOI CMOS Technology for Quantum Computing ICs". In: *IEEE Electron Device Letters* 40.1 (2019), pp. 127–130. DOI: 10.1109/LED.2018.2880303.

[227] H.-C. Han, F. Jazaeri, A. D'Amico, A. Baschirotto, E. Charbon, and C. Enz. "Cryogenic Characterization of 16 nm FinFET Technology for Quantum Computing". In: *ESSCIRC 2021 - IEEE 47th European Solid State Circuits Conference (ESSCIRC)*. 2021, pp. 71–74. DOI: 10.1109/ESSCIRC53450.2021.9567747.

[228] P. A. 'T Hart, M. Babaie, E. Charbon, A. Vladimirescu, and F. Sebastiano. "Subthreshold Mismatch in Nanometer CMOS at Cryogenic Temperatures". In: *ESSDERC 2019 - 49th European Solid-State Device Research Conference (ESSDERC)*. 2019, pp. 98–101. DOI: 10.1109/ESSDERC.2019.8901745.

[229] G. Niu, J. Cressler, S. Mathew, and D. Ahlgren. "Enhanced Low-Temperature Corner Current-Carrying Inherent to Shallow Trench Isolation (STI)". In: *IEEE Electron Device Letters* 20.10 (1999), pp. 520–522. DOI: 10.1109/55.791929.

[230] P. Welch. "The Use of Fast Fourier Transform for the Estimation of Power Spectra: A Method Based on Time Averaging Over Short, Modified Periodograms". In: *IEEE Transactions on Audio and Electroacoustics* 15.2 (1967), pp. 70–73. DOI: 10.1109/TAU.1967.1161901.

[231] M. S. Bartlett. "Periodogram Analysis and Continuous Spectra". In: *Biometrika* 37.1/2 (1950), pp. 1–16. ISSN: 00063444.

[232] F. Hooge. "1/f Noise Is No Surface Effect". In: *Physics Letters A* 29.3 (1969), pp. 139–140. ISSN: 0375-9601. DOI: https://doi.org/10.1016/0375-9601(69)90076-0.

[233] K. . Hung, P. Ko, C. Hu, and Y. Cheng. "A Unified Model For the Flicker Noise in Metal-Oxide-Semiconductor Field-Effect Transistors". In: *IEEE Transactions on Electron Devices* 37.3 (1990), pp. 654–665. DOI: 10.1109/16.47770.

[234] J. Craninckx, A. Potočnik, B. Parvais, A. Grill, S. Narasimhamoorthy, S. Van Winckel, S. Brebels, M. Mongillo, R. Li, B. Govoreanu, and I. Radu. "CMOS Cryo-Electronics for Quantum Computing". In: *2020 IEEE International Electron Devices Meeting (IEDM)*. 2020, pp. 25.1.1–25.1.4. DOI: 10.1109/IEDM13553.2020.9371897.

[235] J. Canny. "A Computational Approach to Edge Detection". In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* PAMI-8.6 (1986), pp. 679–698. DOI: 10.1109/TPAMI.1986.4767851.

[236] A. Grill. "Charge Trapping and Single-Defect Extraction in Gallium-Nitride Based MIS-HEMTs". PhD thesis. Institut für Mikroelektronik, 2018. DOI: 10.34726/hss.2018.60228.

[237] N. Otsu. "A Threshold Selection Method from Gray-Level Histograms". In: *IEEE Transactions on Systems, Man, and Cybernetics* 9.1 (1979), pp. 62–66. DOI: 10.1109/TSMC.1979.4310076.

[238] A. Chambolle. "An Algorithm for Total Variation Minimization and Applications". In: *Journal of Mathematical Imaging and Vision* 20.1 (2004), pp. 89–97. ISSN: 1573-7683. DOI: 10.1023/B:JMIV.0000011325.36760.1e.

[239] F. M. Puglisi and P. Pavan. "Factorial Hidden Markov Model Analysis of Random Telegraph Noise in Resistive Random Access Memories". In: *ECTI Transactions on Electrical Engineering, Electronics, and Communications* 12.1 (Jan. 2014), pp. 24–29. DOI: 10.37936/ecti-eec.2014121.170814.

[240] A. Asenov, R. Balasubramaniam, A. Brown, and J. Davies. "RTS Amplitudes in Decananometer MOSFETs: 3-D Simulation Study". In: *IEEE Transactions on Electron Devices* 50.3 (2003), pp. 839–845. DOI: 10.1109/TED.2003.811418.

[241] T. Knobloch. "On the Electrical Stability of 2D Material-Based Field-Effect Transistors". PhD thesis. Institut für Mikroelektronik, 2021.

[242] M. G. Peters, J. I. Dijkhuis, and L. W. Molenkamp. "Random Telegraph Signals and 1/f Noise in a Silicon Quantum Dot". In: *Journal of Applied Physics* 86.3 (1999), pp. 1523–1526. DOI: 10.1063/1.370924.

[243] J. E. Thomas and D. R. Young. "Space-Charge Model for Surface Potential Shifts in Silicon Passivated with Thin Insulating Layers". In: *IBM Journal of Research and Development* 8.4 (1964), pp. 368–375. DOI: 10.1147/rd.84.0368.

[244] Y. Miura and Y. Matukura. "Investigation of Silicon-Silicon Dioxide Interface Using MOS Structure". In: *Japanese Journal of Applied Physics* 5.2 (Feb. 1966), pp. 180–180. DOI: 10.1143/jjap.5.180.

[245] B. E. Deal, M. Sklar, A. S. Grove, and E. H. Snow. "Characteristics of the Surface-State Charge (Q_{ss}) of Thermally Oxidized Silicon". In: *Journal of The Electrochemical Society* 114.3 (1967), p. 266. DOI: 10.1149/1.2426565.

[246] T. Grasser, M. Waltl, Y. Wimmer, W. Gös, R. Kosik, G. Rzepa, H. Reisinger, G. Pobegen, A.-M. El-Sayed, A. L. Shluger, and B. Kaczer. "Gate-Sided Hydrogen Release as the Origin of "Permanent" NBTI Degradation: From Single Defects to Lifetimes". In: *Proceedings of the IEEE International Electron Devices Meeting (IEDM)*. 2015, pp. 535–538. DOI: 10.1109/IEDM.2015.7409739.

[247] T. Grasser, B. Stampfer, M. Waltl, G. Rzepa, K. Rupp, F. Schanovsky, G. Pobegen, K. Puschkarsky, H. Reisinger, B. J. O'Sullivan, and B. Kaczer. "Characterization and Physical Modeling of the Temporal Evolution of Near-Interfacial States Resulting from NBTI/PBTI Stress in nMOS/pMOS Transistors". In: *Proceedings of the IEEE International Reliability Physics Symposium (IRPS)*. 2018, 2A.2-1–2A.2-10.

[248] B. Kaczer, T. Grasser, P. J. Roussel, J. Martin-Martinez, R. O'Connor, B. J. O'Sullivan, and G. Groeseneken. "Ubiquitous Relaxation in BTI Stressing-New Evaluation and Insights". In: *Proceedings of the International Reliability Physics Symposium (IRPS)*. 2008, pp. 20–27.

[249] T. Grasser, H. Reisinger, P. Wagner, B. Kaczer, F. Schanovsky, and W. Gös. "The Time Dependent Defect Spectroscopy (TDDS) for the Characterization of the Bias Temperature Instability". In: *Proceedings of the International Reliability Physics Symposium (IRPS)*. 2010, pp. 16–25. ISBN: 978-1-4244-5431-0.

[250] M. Waltl, B. Stampfer, G. Rzepa, B. Kaczer, and T. Grasser. "Separation of Electron and Hole Trapping Components of PBTI in SiON nMOS Transistors". In: *Microelectronics Reliability* 114 (2020), pp. 113746-1–113746-5. DOI: 10.1016/j.microrel.2020.113746.

[251] T. Grasser, W. Gös, Y. Wimmer, F. Schanovsky, G. Rzepa, M. Waltl, K. Rott, H. Reisinger, V. V. Afanas'Ev, A. Stesmans, A.-M. El-Sayed, and A. L. Shluger. "On the Microscopic Structure of Hole Traps in pMOSFETs". In: *Proceedings of the IEEE International Electron Devices Meeting (IEDM)*. 2014, pp. 530–533. ISBN: 978-1-4799-8001-7. DOI: 10.1109/IEDM.2014.7047093.

[252] R. Courant, P. Lax, and D. Hilbert. *Methoden der Mathematischen Physik*. Springer Berlin Heidelberg, 2013. ISBN: 9783642580390.

List of Publications

Scientific Journals, Book Contribution

[MJJ1] **J. Michl**, A. Grill, D. Waldhör, W. Goes, B. Kaczer, D. Linten, B. Parvais, B. Govoreanu, I. Radu, M. Waltl, and T. Grasser. "Efficient Modeling of Charge Trapping at Cryogenic Temperatures-Part I: Theory". In: *IEEE Transactions on Electron Devices* 68.12 (2021), pp. 6365–6371. DOI: 10.1109/TED.2021.3116931.

[MJJ2] D. Waldhör, C. Schleich, **J. Michl**, B. Stampfer, K. Tselios, E. Ioannidis, H. Enichlmair, M. Waltl, and T. Grasser. "Toward Automated Defect Extraction From Bias Temperature Instability Measurements". In: *IEEE Transactions on Electron Devices* 68.8 (2021), pp. 4057–4063. DOI: 10.1109/TED.2021.3091966.

[MJJ3] K. Tselios, D. Waldhör, B. Stampfer, **J. Michl**, E. Ioannidis, H. Enichlmair, T. Grasser, and M. Waltl. "On the Distribution of Single Defect Threshold Voltage Shifts in SiON Transistors". In: *IEEE Transactions on Device and Materials Reliability* 91.2 (2021), pp. 199–206. DOI: 10.1109/TDMR.2021.3080983.

[MJJ4] A. Beckers, **J. Michl**, A. Grill, B. Kaczer, M. G. Bardon, B. Parvais, B. Govoreanu, K. De Greve, G. Hiblot, and G. Hellings. "Physics-Based and Closed-Form Model for Cryogenic MOSFET Subthreshold Swing". In: *IEEE Transactions on Electron Devices* (2022). Submitted.

[MJJ5] **J. Michl**, A. Grill, D. Waldhör, W. Goes, B. Kaczer, D. Linten, B. Parvais, B. Govoreanu, I. Radu, T. Grasser, and M. Waltl. "Efficient Modeling of Charge Trapping at Cryogenic Temperatures-Part II: Experimental". In: *IEEE Transactions on Electron Devices* 68.12 (2021), pp. 6372–6378. DOI: 10.1109/TED.2021.3117740.

[MJJ6] C. Schleich, D. Waldhör, T. Knobloch, W. Zhou, B. Stampfer, J. Michl, M. Waltl, and T. Grasser. "Single-Versus Multi-Step Trap Assisted Tunneling Currents – Part I: Theory". In: *IEEE Transactions on Electron Devices* (2022), pp. 1–7. DOI: 10.1109/TED.2022.3185966.

Conference Proceedings

[MJC1] K. Tselios, B. Stampfer, **J. Michl**, E. Ioannidis, H. Enichlmair, and M. Waltl. "Distribution of Step Heights of Electron and Hole Traps in SiON nMOS Transistors". In: *Proceedings of the International Integrated Reliability Workshop (IIRW)*. 2020, pp. 1–6. DOI: 10.1109/IIRW49815.2020.9312871.

[MJC2] K. Tselios, **J. Michl**, T. Knobloch, H. Enichlmair, E. Ioannidis, R. Minixhofer, T. Grasser, and M. Waltl. "Evaluation of the Impact of Defects on Threshold Voltage Drift Employing SiO_2 pMOS Transistors". In: *Abstracts of the 26th European Symposium on Reliability of Electron Devices, Failure Physics and Analysis*. Accepted. 2022, p. 99.

[MJC3] A. Grill, E. Bury, **J. Michl**, S. E. Tyaginov, D. Linten, T. Grasser, B. Parvais, B. Kaczer, M. Waltl, and I. Radu. "Reliability and Variability of Advanced CMOS Devices at Cryogenic Temperatures". In: *Proceedings of the IEEE International Reliability Physics Symposium (IRPS)*. 2020, pp. 1–6. ISBN: 978-1-7281-3199-3. DOI: 10.1109/IRPS45951.2020.9128316.

[MJC4] **J. Michl**, A. Grill, B. Stampfer, D. Waldhoer, C. Schleich, T. Knobloch, E. Ioannidis, H. Enichlmair, R. Minixhofer, B. Kaczer, B. Parvais, B. Govoreanu, I. Radu, T. Grasser, and M. Waltl. "Evidence of Tunneling Driven Random Telegraph Noise in Cryo-CMOS". In: *2021 IEEE International Electron Devices Meeting (IEDM)*. 2021, pp. 31.3.1–31.3.4. DOI: 10.1109/IEDM19574.2021.9720501.

[MJC5] A. Grill, V. John, **J. Michl**, A. Beckers, E. Bury, S. Tyaginov, B. Parvais, A. V. Chasin, T. Grasser, M. Waltl, B. Kaczer, and B. Govoreanu. "Temperature Dependent Mismatch and Variability in a Cryo-CMOS Array with 30k Transistors". In: *2022 IEEE International Reliability Physics Symposium (IRPS)*. 2022, 10A.1-1-10A.1–6. DOI: 10.1109/IRPS48227.2022.9764594.

[MJC6] T. Knobloch, **J. Michl**, D. Waldhör, Y. Illarionov, B. Stampfer, A. Grill, R. Zhou, P. Wu, M. Waltl, J. Appenzeller, and T. Grasser. "Analysis of Single Electron Traps in Nano-scaled MoS_2 FETs at Cryogenic Temperatures". In: *Proceedings of the Device Research Conference (DRC)*. 2020, pp. 52–53.

[MJC7] **J. Michl**, A. Grill, D. Claes, G. Rzepa, B. Kaczer, D. Linten, I. Radu, T. Grasser, and M. Waltl. "Quantum Mechanical Charge Trap Modeling to Explain BTI at Cryogenic Temperatures". In: *Proceedings of the IEEE International Reliability Physics Symposium (IRPS)*. 2020, pp. 1–6. ISBN: 978-1-7281-3199-3. DOI: 10.1109/IRPS45951.2020.9128349.