DISSERTATION

Advanced Electrical Characterization of Charge Trapping in MOS Transistors

ZUR ERLANGUNG DES AKADEMISCHEN GRADES
Doktor der technischen Wissenschaften

EINGEREICHT AN DER
Technische Universität Wien
Fakultät für Elektrotechnik und Informationstechnik

VON
Dipl.-Ing. **Bernhard Stampfer**
Matr. Nr. 00927191
geboren am 27. Februar 1989 in Hall i.T., Österreich

UNTER BETREUUNG VON
Univ.Prof. Dipl.-Ing. Dr.techn. **Tibor Grasser**
UND
Dipl.-Ing. Dr.techn. **Michael Waltl**

Wien, 28. Oktober 2020

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.
TU Bibliothek WIEN Your knowledge hub

ACKNOWLEDGEMENTS

First, I would like to express my gratitude to PROFESSOR TIBOR GRASSER, for supervising me throughout my PhD study, supporting me in academic questions whenever needed, and also for securing the funding necessary for my employment.

I would like to thank MICHAEL WALTL, who has been my colleague when I started my PhD studies and later acquired funding for a project where I have been contributing to. He developed the custom measurement equipment I used for most of my characterization experiments and helped me whenever I encountered issues with my measurements.

I also want to thank reasearchers working at other institutions whom I collaborated with. Among them JOERG APPENZELLER from PURDUE UNIVERSITY, KEN SAWADA from SONY, FRANZ SCHANOVSKY from GTS, and the members of the device reliability group at IMEC LEUVEN, in particular BEN KACZER, MARKO SIMICIC and PIETER WECKX, where I spent an interesting time during an internship.

I want to thank my colleagues, especially ALEXANDER GRILL, MARKUS JECH, MARKUS KAMPL, THERESIA KNOBLOCH, ROBERT KOSIK, GERHARD RZEPA, BIANKA ULLMANN and YANNICK WIMMER, for valuable discussions and input, and JAKOB MICHL, CHRISTIAN SCHLEICH, and DOMINIC WALDHÖR who further contributed to this work by proofreading it.

Further, I want to thank our non-scientific staff at the institute for microelectronics, including EWALD HASLINGER, MANFRED KATTERBAUER, DIANA POP, and MARKUS SCHLOFFER, who supported me in various other matters, and for ensuring a good working environment on our institute.

Finally, I would like to thank my family and friends, especially my parents IRENE and VINZENZ, my brother ANDREAS, and my girlfriend CARMEN.

ABSTRACT

The metal-oxide-semiconductor field effect transistor (MOSFET) is the basic building block of integrated circuits and stands at the heart of most modern electronic devices. It is considered a key technology and had a considerable influence on the cultural, societal, and economic development of the past decades. In light of this, it is not surprising that enormous efforts are put in the improvement of the transistors. One important aspect affecting the performance and also the reliability of these devices are atomic scale defects, inevitably introduced during manufacturing or newly created during operation. These defects possess the ability to capture electric charges and thereby affect the operation of the transistors. With the ongoing miniaturization of transistors, the impact of such defects on the device behavior grows. Effects for which these defects are responsible are among others bias-temperature instabilities (BTI), random telegraph noise (RTN) and hot carrier degradation (HCD).

With the ongoing development of MOSFETs, various characterization methods have been developed to study the wealth of effects plaguing these devices. In the last years, the miniaturization of the transistors paved the way for the development of new methods which aim at the detailed characterization of single defects. These measurements allow to verify and improve detailed models of single defects describing their atomistic nature and their charge trapping behavior. At the same time, this poses new challenges for characterization of a technology, as many such defects need to be studied to obtain the averages and distributions of the defects' influences on the devices, which requires improvements of the experimental approaches and data processing.

This work covers the experimental characterization and theoretical description of these defects and elaborates the measurement data analysis in combination with TCAD simulation. Different characterization and simulation methods are used to study defects in three separate technologies to draw conclusions about their distributions and physical nature.

KURZFASSUNG

Der Metall-Oxid-Halbleiter Feldeffekttransistor (MOSFET) ist der Grundbaustein integrierter Schaltungen und damit der meisten modernen elektronischen Geräte. Er gilt als eine Schlüsseltechnologie und hatte maßgeblichen Anteil an der kulturellen, gesellschaftlichen und wirtschaftlichen Entwicklungen der letzen Jahrzehnte. In Anbetatracht dessen ist es nicht verwunderlich, dass zu dessen Verbesserung ein enormer Aufwand betrieben wird.

Ein wesentlicher Aspekt in der Leistung und Zuverlässigkeit dieser Bauteile sind unvermeidbare Defekte auf atomarer Ebene, die während der Herstellung oder des Betriebs entstehen. Diese Defekte können Ladungsträger einfangen, und damit das Verhalten der Transistoren beeinflussen. Mit fortschreitender Miniaturisierung der Transistoren gewinnen solche Defekte zunehmenden Einfluss auf deren Verhalten. Effekte die durch diese Defekte entstehen sind unter Anderem *Bias-Temperature-Instabilities* (BTI), *Random-Telegraph-Noise* (RTN) und *Hot-Carrier-Degradation* (HCD).

Mit der Weiterentwicklung des MOSFET wurden verschiedene Messverfahren entwickelt, um die Vielzahl an Effekten zu studieren welche diese Bauteile beeinflussen. In den letzen Jahren ermöglichte die Miniaturisierung der Transistoren die Entwicklung neue Messverfahren, welche die detailierte Charakterisierung einzelner Defekte erlauben. Diese Messungen erlauben die Verifizierung und Verbesserung detailierter Einzeldefektmodelle, welche die atomistische Natur oder den Ladungsträgeraustausch dieser Defekte beschreiben. Gleichzeitig stellt es neue Herrausforderungen für die Charaterisierung einer Technologie dar, da viele dieser Einzeldefekte erforscht werden müssen um Aussagen über die Mittelwerte und Verteilungen der Auswirkungen der Defekte auf die Bauteile treffen zu können. Dies erfordert Verbesserungen der experimentellen Ansätze und der Datenauswertungen.

Diese Arbeit befasst sich mit der messtechnischen Charakterisierung und der theoretischer Beschreibung dieser Defekte und beschreibt die Messdatenanalyse in Verbindung mit TCAD Simulationen. Unterschiedliche Charakterisierungs- und Simulationsmethoden werden verwendet um Defekte in drei verschiedenen Technologien zu untersuchen, und Rückschlüsse auf deren Verteilungen und physikalischer Natur zu ziehen.

CONTENTS

Acknowledgements **iii**

Abstract **v**

Kurzfassung **vii**

Contents **ix**

List of Figures **xiii**

List of Abbreviations **xxvii**

1 Introduction **1**
- 1.1 History of the Field Effect Transistor 1
- 1.2 Reliability Issues in Field Effect Transistors 3
- 1.3 This Work 5

2 Defects in Field Effect Transistors **7**
- 2.1 Interface Defects 7
- 2.2 Oxide Defects 9
  - 2.2.1 Oxide Defects in SiO_2 10
- 2.3 Semiconductor Bulk Defects 11

3 Modeling and Simulation of Defects **13**
- 3.1 Defect Models for Charge Trapping 13
  - 3.1.1 The Shockley-Read-Hall Model 15
  - 3.1.2 The Kirton and Uren Model 17
  - 3.1.3 The NMP Model 18
  - 3.1.4 The 4-State NMP Model 20
  - 3.1.5 The Hydrogen Release Model 21
- 3.2 Defect Simulations 23
  - 3.2.1 Ab-Initio Simulations 24
  - 3.2.2 Defects in Device Simulation 24
  - 3.2.3 Defects in Circuit Simulation 25
- 3.3 The Defect Centric Model 26

4 Methods of Defect Characterization 29
4.1 Electrical Methods based on Channel Conductivity 29
4.1.1 Random Telegraph Noise (RTN) Measurements 33
4.1.2 Measure-Stress-Measure (MSM) Methods 38
4.1.3 Extended Measure-Stress-Measure (eMSM) 38
4.1.4 On-the-Fly (OTF) Measurement 42
4.1.5 Hysteresis Measurements 43
4.2 Electrical Methods based on Defect Charge 43
4.2.1 Capacitance-Voltage (CV) 46
4.2.2 Charge Pumping (CP) 49
4.2.3 Deep-Level Transient Spectroscopy (DLTS) 51
4.2.4 Direct-Current IV (DCIV) 52
4.2.5 Thermal Dielectric Relaxation Current (TDRC) 53
4.3 Physical Characterization Methods 53
4.3.1 Electron Paramagnetic Resonance (EPR) 53
4.3.2 X-Ray Photoelectron Spectroscopy (XPS) 54
4.3.3 Secondary Ion Mass Spectroscopy (SIMS) 55
4.3.4 Neutron Activation Analysis (NAA) 55

5 Defect Parameter Extraction from RTN, TDDS, and CV Measurements 57
5.1 Random Telegraph and TDDS Signals 57
5.1.1 Histogram and Lag Methods 58
5.1.2 Edge Detection 59
5.1.3 Hidden Markov Models 65
5.1.4 Estimations for Trap Level and Position 69
5.1.5 Frequency Domain Methods 70
5.1.6 Time-Dependent Defect Spectroscopy 71
5.2 Capacitance-Voltage Measurements 73
5.2.1 Capacitance Methods 73
5.2.2 Conductance Method 77

6 Measurements and Results 79
6.1 Statistical Characterization of Defects Causing RTN in SiO_2 Transistors . 79
6.1.1 Devices and Measurements 80
6.1.2 Parameter Extraction 81
6.1.3 Step Heights 81
6.1.4 Energy and Position 82
6.1.5 Conclusions 85
6.2 Defect Centric Evaluation of RTN and BTI using pMOS Arrays 85
6.2.1 Array Chip 86
6.2.2 Measurements 87
6.2.3 Parameter Extraction 90

6.2.4 Simulation and Extrapolation 93
6.2.5 Bulk and Drain Bias Dependence 94
6.2.6 Conclusions 95
6.3 Single Defects in Few-Layer MoS_2 Devices 96
6.3.1 Device and Measurements 96
6.3.2 Charge Transition Times 98
6.3.3 Simulation and Results 100
6.3.4 Conclusions 102

Summary and Outlook **105**

References **109**

List of Publications **125**

LIST OF FIGURES

2.1 Illustration of P_b centers at the (100) Si/SiO_2 interface. Shown are P_{b0} ($\cdot$Si≡Si_3) and P_{b1} centers ($\cdot$Si≡Si_2O). Yellow: Silicon, Red: Oxygen, Blue: Silicon dangling bond. The dangling bonds carry a net spin moment and are thus visible in ESR measurements. Recreated from [32] 8
2.2 Illustration of the channel carrier densities in a scaled device and the resulting current percolation path. **(a)** Random dopands lead to locally decreased carrier densities. The neutral defect has no influence on the channel. **(b)** The defect is now charged and increases the restriction of the percolation path, leading to a decreased drain current and a higher threshold voltage. 10
2.3 Ball-and-stick models of two candidates for hole trap in amorphous SiO_2. **(a)** The hydrogen bridge and **(b)** the hydroxyl E′ center. Here, yellow are the silicon, red are the oxygen, and gray are the hydrogen atoms. Additionally, the blue clouds represent the spin density for neutral defects, i.e. the location of the unpaired electron, while for the positive defects it shows the distribution of the captured hole. Both defects may exist in four configurations, with two of them neutral (states 1 and 1′) and two of them charged (states 2 and 2′). To change between two states with the same charge, the defects transition between their regular and puckered configuration [43, 44]. For this, a silicon atom moves through the plane spanned by its three neighboring oxygen atoms. Transitions between the neutral and charged states additionally require a charge transfer with the substrate or gate. Originally published in [BSJ1] and adapted from [45] 11

3.1 Example of a Markov process with three states, together with the corresponding states $\vec{I}$, the state vector $\vec{P}$ in equilibrium, and the transition matrix **K**. 14
3.2 Schematic band diagram showing the capture and emission processes for acceptor-like traps (Blue) and donor-like traps (Red). Electron traps change between neutral and negative charge states while hole traps change between neutral and positive charge. States of the defects are shown as they are prior to the indicated transition. 15

3.3 Kirton-Uren model illustrated in a configuration coordinate diagram. A defect exchanges an electron with the reservoir. The position of the parabola depends on the carrier energy. The capture barrier ΔE_B accounts for thermal activation of the capture process. Blue: Empty defect, electron in the conduction band. Red: Electron captured at the defect site. **Dashed:** Energy zero of the system: empty defect, electron at the Fermi level. Recreated after [51] 18

3.4 Parabolic potential energy surfaces (PESs) for the charged (red) and a neutral state (blue) as used in the NMP defect model, together with the parameters of the model. 19

3.5 Markov chain **(a)** and potential energy surfaces **(b)** of the four-state defect. In addition to the neutral and charged states 1 and 2, there is an additional metastable state for each charge state marked with a prime ($'$). Transitions between neutral and charged states are modeled using NMP theory, while transitions between states of the same charge are purely thermal transitions described using classical transition state theory. Changing the gate voltage shifts the trap level which leads to a shift in the PES relative to each other. 21

3.6 Illustration of the hydrogen release model. A number of precursor sites (orange) exist in the oxide, which become defects once they capture a hydrogen atom. Vice versa, in one of the neutral configurations of the defect (blue) there is a realistic probability for the hydrogen atom to detach from the defect **(a)**. Once this happens, the hydrogen can diffuse in the oxide **(b)** and bond to another precursor **(c)**. The high diffusion rate for hydrogen in SiO_2, together with a low barrier of hydrogen capture by the precursors leads to a low probability of free hydrogen at any time. Additional hydrogen is supplied by the gate at harsh stress conditions **(d)**, leading to additional defects. States of the defects are shown as they are before the indicated transitions. **Inset:** Markov chain of the hydrogen release defect. Compared to the four-state defect there is an additional precursor state to or from which the defect can transition by emission or capture of hydrogen, respectively. 22

3.7 Experiment on long-term degradation of a 10 um × 10 um planar pMOS transitor. Measured permanent degradation component (circles), together with simulation results extracted using the hydrogen-release model (red). From [76] 24

3.8 Distributions of the defects per device and the step height per defect in the defect centric model. **Left:** The average number of defects in the devices is assumed to be Poisson distributed around a mean value N. **Right:** The impact of the single defects on the threshold voltage shift is given by an exponential distribution with the expectation value η. While the majority of defects ($1 - \frac{1}{e} \approx 63\%$) show step heights below η, some defects produce much larger steps. The variance in both the number of defects and step height per defect underlines the fact that for small devices with few defects on average, some devices can show degradation much worse than the average device. 26

3.9 **Left:** Measured distribution of ΔV_{th} after stress in comparison with the defect centric model. **Center:** The one-sided BTI component gets smaller after stress due to decreasing N. **Right:** The symmetric RTN component stays constant throughout recovery. 28

4.1 Schematic of a constant gate voltage measurement setup for channel conductivity based methods. The drain and gate voltages of the device are controlled by DACs. The gate voltage is applied directly to the gate of the DUT, while the drain current is supplied to the positive terminal of an OPAMP configured as an transimpedance amplifier. Its negative input terminal is connected to the drain connection of the DUT. The voltage output of the transimpedance amplifier, which corresponds to I_{D}, is read by an ADC. Originally published in [BSJ1] 30

4.2 Schematic of a constant source current measurement setup for channel conductivity based methods. The drain bias of the DUT is controlled directly, while the gate bias is controlled by the OPAMP. The OPAMP circuit controls the gate bias of the MOSFET in a way that the source current equals the reference current defined by the DAC bias $-I_{\mathrm{s}}R_{\mathrm{s}}$. Originally published in [BSJ1] 30

4.3 **Top:** Schematic of the measurement system [39] employed for this work. Analog and digital subsystems colored in blue and orange, respectively; control signals omitted for clarity. The system consists of a 19” rack case with 12 slots which can be equipped with a number of different inserts. The inserts connect to a back-panel which provides analog and digital power, a number of signaling lines, an I^2C bus and individual USB lines. For the methods based on channel conductivity as outlined in this section, the constant gate voltage scheme (Figure 4.1) has been used. **Bottom:** Using the system, the channel current can be measured either on the source terminal or on the drain terminal of the device as shown in the lower part of the figure. Additional inserts can be equipped to measure e.g. the gate current. 32

4.4 RTN measurements in the time domain **(left)** and frequency domain **(right)**. In large devices **(top)**, the noise signal in the time domain corresponds to a 1/f spectrum in the frequency domain. In small devices **(bottom)**, due to the small number of defects and the large influence of each defect on the channel current, individual steps can be observed. In the frequency domain, the Lorentzian PSDs of the individual defects which make up the 1/f shape in large devices might still be distinguished. Originally published in [BSJ1] 34
4.5 Energetical window for RTN characterization shown in a Si-SiO$_2$ band diagram. Defects located close to the Fermi level randomly exchange charge with the channel and produce RTN. Defects located far above (below) the Fermi level will be neutral (charged) for most of the time. The brief phases where they are charged (neutral) will most likely be too short to sample. Originally published in [BSJ1] 36
4.6 Flow chart illustrating parameter extraction from single defect RTN measurements. To obtain parameters suitable for defect simulation, the RTN signal parameters for each defect first have to be extracted from the measurements. Using these parameters, physical defect parameters such as their trap levels can be obtained using analytical estimates or TCAD simulation. 37
4.7 Gate and drain biases during a MSM measurement. After an initial measurement, phases of stress and subsequent measurements are performed. This allows to characterize the degradation due to the applied stress conditions. In this example, degradation due to gate bias stress is characterized using $I_\mathrm{D}(V_\mathrm{G})$ measurements. 39
4.8 Illustration of $I_\mathrm{D}(V_\mathrm{G})$ characteristics of a device before and after BTI stress. From this, the threshold voltage shift ΔV_th, changes in the subthreshold slope, the off-current and the transconductance can be obtained. Originally published in [BSJ1] 39
4.9 Gate bias and drain current during an eMSM measurement. Phases of stress alternate with phases of relaxation. During relaxation the drain current or threshold voltage shift is recorded. This allows to observe the impact of stress on the device behavior and also to understand the following relaxation behavior. 40
4.10 Measurement data (**circles**) and simulation data (**lines**) for a BTI eMSM experiment. Note that the simulations suggest that a significant part of the threshold voltage shift due to stress already recovered before the first recovery measurement point is taken, underlining the need for fast measurements. 40
4.11 Hot carrier stress. Unlike BTI stress, a non-negligible source-drain bias is applied, causing carriers to accelerate towards the drain. This causes damage at the interface or secondary carrier generation. 41

4.12 Measurement windows for RTN and TDDS single defect characterization, shown for an exemplary defect. While RTN allows to characterize a defect at gate voltages close to $E_\mathrm{F} = E_\mathrm{t}$, TDDS allows to measure its charge capture time at biases above this point and its charge emission time at biases below that point. By characterizing the defect using both methods, a comprehensive picture of the defects' behavior can be obtained. Originally published in [BSJ1] 42

4.13 Gate and drain biases during an on-the-fly measurement. A small drain bias is applied for the duration of the measurement. At the same time, the gate bias is pulsed to measure selected points on the $I_\mathrm{D}(V_\mathrm{G})$ curve. This keeps the interruption to stress at a minimum and thus allows to observe the degradation of the MOSFET. 43

4.14 Hysteresis measurements on a 4H-SiC nMOSFET with varied starting voltage $V_\mathrm{G,low}$ and constant high voltage $V_\mathrm{G,high}$. Decreasing $V_\mathrm{G,low}$ increases the hysteresis width. Data originally published in [99] 44

4.15 Schematics of a circuit as might be used for the charge based methods. With this setup, all signal processing has to be done in software. Particularly for CV measurements, part of the impedance extraction might be performed in hardware. In CP measurements, a hardware integrator might be used. 44

4.16 **Top:** Schematic of the measurement system [39] employed in this work, equipped for the measurement methods presented in this section. Analog and digital subsystems colored in blue and orange, respectively; control signals omitted for clarity. As compared to Figure 4.3 for channel current measurements, an additional sampling unit included to measure the phase of the input signal for CV and DLTS measurements. **Bottom:** Possible configurations for gate and bulk current measurements. For CV measurements both the gate or the bulk current can be used to measure the MOS capacitance. Especially when testing devices on a wafer this decision may however influence the noise level of the measurement. Source and drain terminals—if available on the test structure—may either be shorted to bulk or supplied separately (as shown), in case the gate current is measured optionally with or without the AC signal. Depending on the chosen configuration the gate-source and gate-drain capacitances may be included in the measurement. 45

4.17 Capacitance-voltage measurement. **(a)** DUT connected for measurement. **(b)** DUT in the various operation schemes with illustrated equivalent capacitance. In the depletion regime there are no free charges close to the interface. This can be interpreted as an increase in spacing between the plates of the equivalent capacitor. This in turn lowers the capacitance and gives the CV curve its characteristic shape. **(c)** Voltage and current signals during measurement. The phase and amplitudes of the voltage and current signals give the sought-after impedance. The offset voltage is swept in a staircase-like manner during the measurement. **(d)** Exemplary results of a CV measurement at multiple frequencies. The accumulation and inversion branches almost reach the oxide capacitance, while the capacitance in depletion is much lower. The inversion branch is generally steeper than the accumulation branch. 47

4.18 Extraction of the phase difference from a CV measurement. The current signal is mixed with the original voltage signal **(left)** and a phase-shifted signal **(right)**. This gives signals with their sum and difference frequencies, i.e. $2\omega_0$ and $0\omega_0$. After low-pass filtering or integrating, the DC I and Q signals which are proportional to the imaginary and real parts of the current remain. 48

4.19 The CP technique. **(a)** DUT during measurement. During the high bias phase defects charge with minorities supplied from source and drain. In the following low bias phase, the minorities are emitted and recombine with majorities sourced from the bulk. Their amount is measured as the bulk current. **(b)** Gate voltage and bulk current during measurement. **(c)** Gate voltage parameters for the two measurement schemes. **(d)** Illustration of a measurement result. 50

4.20 Illustration of deep-level transient spectroscopy using rate windows. As the temperature of the sample is increased, the time constants of defects in the device decrease. Once the response time of a defect band coincides with the chosen rate window given by $t_{1,2}$, the difference between the capacitance at these points—the DLTS signal—increases. At even higher temperatures the defects responses get too fast and the signal decreases again. 51

4.21 DCIV measurement in the top emitter configuration. The drain and source diodes are slightly forward biased, causing minority injection into the space charge region. As the voltage is swept from accumulation to inversion, defect charges recombine with the injected minorities. This causes majorities to move towards the space charge region, in turn creating a measurable bulk current. The scale of this current depends on the number of defects discharging, as illustrated on the right. 52

TU Bibliothek Your knowledge hub WIEN
Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

4.22 Principle of EPR spectroscopy. **(a)** A sample is placed in a homogeneous magnetic field and irradiated by microwave radiation. **(b)** The magnetic field causes the energy levels of unpaired electrons to split due to the Zeeman effect. When the split caused by the magnetic field coincides with the microwave energy, the absorption peaks. 54

4.23 Principle of XPS. **(a)** X-ray photons are directed at the target. Upon collision they remove an electron from an atom. The electron moves to the sample surface and is emitted with an kinetic energy E_k. The binding energy of the electron can then be calculated from the detected energy and the work functions of the sample and the detector. The probed depth can be influenced by the angle of the detector. **(b)** The resulting peaks in the observed binding energies can be linked to the targeted species and their concentrations. Graph from [112] 55

4.24 Principle of SIMS. **(a)** The sample is sputtered with ions. Upon collision they remove ions from the sample. These secondary ions are then analyzed in a mass spectrometer. **(b)** Exemplary results: the composition of the removed material over time is seen in the output of the mass spectrometer. Measurement results from [113] 56

5.1 Exemplary RTN signals **(a)** extracted using histograms **(b,c)** and time lag plots **(d,e)**. The original RTN signal is shown in blue. The corresponding histogram **(b)** shows four partially overlapping peaks, indicating two defects. The time lag plot **(d)** shows the same peaks, but with better separation. In addition, the transitions between the states are shown in the off-diagonals. For the orange signal, a linear drift of 1.5 µV/s was added to show the influence of drift and low-frequency noise. Due to the drift, both the corresponding histogram **(c)** and time lag plot **(e)** show deformed peaks and it becomes difficult to distinguish the small defect. Originally published in [BSJ1] 58

5.2 The BCSUM algorithm applied to an exemplary signal r. The columns show the initial step **(left)**, and the first recursion for the first part **(center)** and second part **(right)** of the signal. **Top:** Signal used during recursion. **Middle:** CSUM signal of the original data. **Bottom:** Histogram of γ^* values obtained from the bootstrapped data, together with the $\epsilon = 0.9$ quantile value (Black), and γ from the original data (Red). Recursion ends when $\gamma < \gamma^*_\epsilon$, as is the case in the bottom center plot. 61

5.3 The Canny algorithm for edge detection demonstrated for an exemplary measurement trace. The input signal $f(t)$ is convoluted with the first derivative of a Gaussian pulse with a chosen variance G'. This gives a signal $H(t)$, which exhibits peaks corresponding to the steps in the original signal. All peaks above a chosen threshold are recognized and used to mark the positions of the steps in $f(t)$. The height of the peaks can be obtained from the original signal or from the height of the peaks. Modified from [BSJ1] 63

5.4 Comparison of the Canny edge detector and the BCSUM method for signals with different noise floors. The signal contains two defects with step heights of $\eta_1 \approx 0.25\,\mathrm{mV}$ and $\eta_2 \approx 0.55\,\mathrm{mV}$. **(top)** Original signal, both methods agree within 2% of absolute deviation of the extracted charge transition times. **(middle)** $\sigma = 0.1\,\mathrm{mV}$ of added noise. The extracted capture time increases as the algorithms miss shorter transitions. **(bottom)** $\sigma = 0.2\,\mathrm{mV}$ of added noise. With a noise level similar to the step height of the extracted defect, spurious transitions affect the detections while proper steps are missed, leading to false results. 64

5.5 The measured drain current can be described by the evolution of the underlying Markov chain(s) of the defects. The reverse, however, may not be true. If reconstruction of the individual defect states is not possible from the recorded drain current data, a hidden Markov model in conjunction with the Baum-Welch algorithm may be used to estimate the model parameters, i.e. transition probabilities $\mathbf{P}$ and emissions $\vec{\mu}$. . . . 66

5.6 Defects found from TDDS measurements. Recovery traces **(top)** are analyzed using a step detection algorithm and the steps are plotted in the step height–emission time plane **(bottom)**. If enough traces are plotted, clusters will form for each defect which emitted during recovery. The clusters are distributed exponentially in time and normally in step height. To obtain the capture time of the defects, measurements at varying stress times have to be performed. From [135] 71

5.7 Illustration of charge capture time extraction from a TDDS measurement. Assuming a low occupancy of the defect at recovery conditions and a sufficient relaxation time, the average occupation of the defect at the end of the stress phase equals the probability of observing the defect discharging during recovery. By plotting this occupancy, i.e. the ratio of the number of emission events over the number of measurement repetitions, over the stress time, the exponential CDF for charge capture can be obtained. The charge capture time of the defect at the stress condition can then be extracted by fitting the theoretical distribution. . . 72

5.8 Schematic CV curves of a pMOS transistor, affected by defects located close to the conduction band. **(a)** Defect free CV curve. **(b)** Fixed charges or defects with $\tau f_{\mathrm{sweep}} > 1$ shift the CV curve along the voltage axis. **(c)** Defects fast enough to react to the DC sweep but too slow to react to the AC frequency cause a stretch-out of the CV curve. **(d)** Defects fast enough to react to the AC frequency cause additional capacitance in the CV curve. In general, a superposition of these effects will be obtained, depending on the chosen measurement parameters. 74

5.9 Equivalent circuits for CV measurements in the different regimes. **(a)** General equivalent circuit. **(b)** Schematic CV-measurments for varying measurement conditions. **(1–4)** Accumulation, depletion, low-frequency inversion and high-frequency inversion equivalent circuits. To obtain the low- and high-frequency curves (blue and orange), the DC sweep rate has to be slow enough for the minority carriers to follow, otherwise the deep depletion curve (red) is obtained. If in addition the AC frequency is low enough for the minority carriers to respond, the low-frequency curve is obtained. 75

5.10 MOSCAP in depletion between low and high frequency regimes for a single interface defect. **(a)** Equivalent circuit with lossy interface defect represented by a series RC circuit. **(b)** Simplified circuit with equivalent parallel capacitance and conductance. **(c)** Behavior of $\mathrm{C_P}$ and $\mathrm{G_P}$ with varying measurement frequency. At low frequencies, $C_\mathrm{P} = C_\mathrm{b}||C_\mathrm{it}$ and at high frequencies $C_\mathrm{P} = C_\mathrm{b}$, in both cases $G_\mathrm{P}/\omega \approx 0$. At the corner frequency $\omega\tau = 1$, however, G_P/ω peaks at C_it. This behavior is exploited in the conductance method [4]. 78

6.1 Illustration of large area **(a)** and small area **(b)** devices with oxide defects randomly distributed in the $\mathrm{SiO_2}$ layer. Devices with a smaller gate area have on average a lower number of defects, but the average impact of the defects is larger. This leads to a similar mean degradation for both types of devices, but a much larger variance among the small devices. Originally published in [BSJ2] 79

6.2 Transfer characteristics of the ≈300 devices measured. Shown are the individual transfer characteristics in gray and the average in blue. The bias and current range in which RTN has been measured is highlighted. Originally published in [BSJ2] 81

6.3 An example of a short RTN trace. Even though this is the faster kind of the measurements performed, the noise level at around 0.15 mV is low enough to clearly see the trapping of two defects in the signal. Originally published in [BSJ2] 81

6.4 Example of the Canny edge detector used on one of the long (1 ks) RTN traces. The original ΔV_{th} signal has been convoluted with the first derivative of a Gaussian pulse to yield a signal h which has peaks at the positions of the steps. Local maxima above a selected threshold give the positions of the steps, their magnitude η has been taken from the original signal. Originally published in [BSJ2] 82

6.5 Complimentary cumulative density function of the step heights observed in the measurements. The distribution seems to be composed of two separate exponential distributions with mean values of 0.39 mV and 1.09 mV. Originally published in [BSJ2] 83

6.6 Examples of the capture and emission time dependence on gate bias and temperature, for two defects which have been characterized in more detail. Originally published in [BSJ2] 83

6.7 Distribution of the extracted vertical positions, measured from the interface, and trap levels for around 100 defects. The extracted energy peaks at around 0.4 eV above the Fermi level, which is close to the conduction band edge during the measurements. The distribution of the depths of the defects in the oxide shows a maximum at 0.6 t_{ox}, which is where the effective trap levels of the defects coincide with the Fermi level. Originally published in [BSJ2] 84

6.8 Simulated band diagram showing the defects extracted using the estimations for depth and trap level. In addition, defect bands for SiO_2 from [49] are shown in gray. Originally published in [BSJ2] 84

6.9 Distribution of the vertical positions, measured from the interface, and trap levels, referenced to the Si midgap, extracted using both the estimation approach and TCAD simulations. The simulation results show a narrower distribution in position and indicate that the defects are slightly closer to the interface compared to the estimations. The extracted trap levels are slightly lower as well, peaking ≈0.3 eV above the conduction band edge. The distribution of trap levels which have been measured covers mainly the lower half of the distribution obtained in [49]. Originally published in [BSJ2] 85

6.10 Layout of the signal lines of the array which have been used for defect characterization. The gate terminals of the transistors in each row can be switched between externally supplied on- or off-biases using on-chip logic. Likewise, the drain terminals can be switched for each transistor column. This allows to address and thus to characterize each individual device in the array. The bulk and source terminals are common for all devices. More details about the array structures can be found in [141]. Originally published in [BSJ3] 87

6.11 A set of $I_D(V_G)$ curves recorded on the long devices (shown in blue). In addition, the mean and variance are given in white and red. From the curves for each device, the gate bias during relaxation has been determined based on a chosen relaxation current $I_{D,r} = 100\,\text{nA}$. Originally published in [BSJ3] 88

6.12 A set of $\Delta V_{th}(t_r)$ mapped from $I_D(t_r)$ using the corresponding set of $I_D(V_G)$ curves (shown in blue). The mean and variance are given in white and red, respectively. The vertical lines indicate moments in time when the distributions of ΔV_{th} have been drawn for further analysis. Originally published in [BSJ3] 89

6.13 Degradation in ΔV_{th} recorded 1 ms after 10 s of stress at $V_{G,str} = -1.45\,\text{V}$. **Top:** for each device as positioned in the matrix of short devices. **Bottom:** averaged over a number of devices. The plots show neither defective rows nor columns, nor clusters or overall inhomogeneities of the extracted degradation of the threshold voltage. Originally published in [BSJ3] 89

6.14 CDFs of ΔV_{th} during recovery. Blue: measured, Red: calculated. Most devices exhibit positive degradation ($-\Delta V_{th} > 0$ for pMOS) after stress due to BTI, while for some devices RTN causes the reverse $-\Delta V_{th} < 0$. Originally published in [BSJ3] 90

6.15 Extracted parameters obtained from the short devices after $t_r = 100\,\text{ms}$ of recovery. The average step height η (×), as well as the average number of RTN charges N_{RTN} (◇) remains constant over all sets. The average number of charges captured due to BTI N (●) depends on stress bias and stress time as expected. Originally published in [BSJ3] 91

6.16 Average number of captured defects N (shade) extracted 2 ms after stress release, over stress time and gate bias. The crosses indicate the measurement points, the dashed lines are contour lines separating the iso-surfaces. For weak stress conditions, only a fraction of the devices exhibits a charged defect after stress. Originally published in [BSJ3] 92

6.17 Average number of captured defects N (shade) extracted after stress at $V_G = -1.45\,\text{V}$, over stress and relaxation time. The lines are contour lines separating iso-surfaces for the number of charged defects. Only for short stress times, full recovery can be observed in a relatively short relaxation time window. Originally published in [BSJ3] 92

6.18 Average degradation of the short devices during sets of stress at various gate biases. The simulation data shown as lines and the measurement data given by the points agree well, except for the first moments after short stress which show unusual negative ΔV_{th} shifts in the measurement. The origin of this behavior remains open at this point and requires further experimental and simulation efforts. However, the simulations qualitatively explain the trend of the measurements. Originally published in [BSJ3] 93

6.19 Average number of captured defects N (shade) simulated after stress at $V_{\mathrm{G}} = -1.45\,\mathrm{V}$, over stress and relaxation time. The lines are contour lines separating iso-surfaces for the number of charged defects. The dashed line indicates a ten-year time frame. Originally published in [BSJ3] . . . 94

6.20 Dependence of the average number of charged defects $t_{\mathrm{r}} = 100\,\mathrm{ms}$ on the stress gate-bulk voltage V_{GB} for measurements with pure gate stress and measurements with part of the stress voltage applied to the bulk ($V_{\mathrm{G}} = -1.3\,\mathrm{V}$). Both stress cases have a similar effect on the degradation. Originally published in [BSJ3] 94

6.21 Evolution of the average number of charged defects during recovery after $t_{\mathrm{s}} = 100\,\mathrm{ms}$ stress at a number of gate-bulk voltages. Points show the measurements after pure gate stress while crosses show measurements with $V_{\mathrm{G}} = -1.3\,\mathrm{V}$ and varied V_{B}. Again, both stress cases seem to have a similar effect. Originally published in [BSJ3] 95

6.22 Dependence of the average number of charged defects on the drain bias. **Left:** After stress and $t_{\mathrm{r}} = 100\,\mathrm{ms}$ of recovery. **Right:** During recovery after $t_{\mathrm{s}} = 100\,\mathrm{ms}$ of stress. The mild drain bias stress applied does not seem to significantly affect the device threshold voltage degradation. Originally published in [BSJ3] 96

6.23 Colorized scanning electron microscope picture of one of the studied devices. The source and drain contacts are shown in yellow, below them is the structured $\mathrm{MoS_2}$ layer, shown in green. Originally published in [BSJ7](supporting material) 97

6.24 **(a)** Schematics of the devices used during measurement. Few layers of $\mathrm{MoS_2}$ are located on top of a $\mathrm{SiO_2}$ wafer. Source and drain have been contacted on top, while the channel has been controlled using the back gate. **(b)** Transfer characteristics $I_{\mathrm{D}}(V_{\mathrm{G}})$ of an exemplary device with the subthreshold slope of $S \approx 1.1\,\mathrm{V/dec}$. **(c)** Output characteristics $I_{\mathrm{D}}(V_{\mathrm{G}})$ for the same device. These devices show normally-on characteristics. Originally published in [BSJ7] 97

6.25 Impact of single defects on different device characteristics: **(a)** Steps in the drain current of a transistor during $I_D(V_G)$ sweeps, equivalent to around 300 mV in ΔV_{th}. In large area devices many such defects would be visible as hysteresis, leading to a similar width of the hysteresis, but the individual contributions would not be observable. **(b)** A defect causing RTN. The stochastic behavior of the defect causes a large variety of charge capture and emission times. **(c)** Similar to **(b)**, but this defect shows a significant gate bias dependence. **(d)** A defect showing aRTN, periods of noise are separated by periods of inactivity. The Markov chains necessary to model the respective defect's behavior are shown in the insets **(b)** and **(d)**. aRTN can not be described using a two state model and, in this case, needs an additional neutral state. Originally published in [BSJ7] 99

6.26 Capture and emission times extracted for four defects **(symbols)** and fits **(lines)**. The charge transition times of defects A and D show a significant dependence on the gate bias, while the charge transition times of defects B and C seem to be unaffected by the gate bias. Defect B, which shows aRTN, is described using two sets of transition times due to its additional meta-stable state. The fits for defects B to D are from numerical simulation, while the fit for defect A is a linear fit. This is due to the current inability to simulate these devices at the low temperature at which defect A was measured. Originally published in [BSJ7] 100

6.27 The SRH and NMP defect models both widely used for the description of interface and oxide defects. **(a)** The mechanisms for charge transfer as outlined in Section 3.1. **(b)** Simulations of the capture and emission times of defects placed at a distance of 2 nm above and below the channel. It can be seen that due to the lack of a backward energy barrier the SRH model, only one of the charge transition times features a meaningful temperature and bias dependence at any voltage. The SRH model can not describe the strongly voltage dependent capture and emission times as observed for defects A and D. For defect C the temperature dependence of both capture and emission time can not be described by the simple SRH model. Finally, multi state defects such as defect B can not be described at all using the simple SRH model, as it is limited to a defect with only a charged and a neutral state. Originally published in [BSJ7] 101

6.28 Extracted spatial and energetic positions for defects A-D shown in a band diagram simulated for a device considering a 6 nm thin MoS_2 layer. Defects A and D are located below the channel, while defects B and C are on top of it. The gate bias used for simulation was 1.2 V (opaque) and −4.2 V (semi-opaque). The charge transition times for defects A and D change with the gate voltage as their position relative to the Fermi level shifts. The gray lines depict defect bands extracted for SiO_2 [BSC6]. Originally published in [BSJ7] 102

List of Abbreviations

ADC	analog-to-digital coverter
aRTN	anomalous RTN
BCSUM	bootstrapping and cumulative sum
BJT	bipolar junction transistor
BTI	bias temperature instability
CCD	charge coupled device
CCDF	complementary cumulative density function
CDF	cumulative density function
CMOS	complementary MOS
CP	charge pumping
CV	capacitance-voltage
DAC	digital-to-analog coverter
DCIV	direct-current current-voltage
DFT	density functional theory
DIBL	drain-induced barrier lowering
DLTS	deep-level transient spectroscopy
DUT	device under test
eMSM	extended MSM
EPR	electron paramagnetic resonance
ESR	electron spin resonance
FET	field-effect transistor
FHMM	factorial HMM
GIDL	gate-induced drain leakage
HC	hot carrier
HMM	hidden Markov model
I^2C	inter-integrated circuit
IV	current-voltage
MIS	metal-insulator-semiconductor
MOS	metal-oxide-semiconductor
MoS_2	molybdenum disulfide
MOSCAP	MOS capacitor
MOSFET	MOS field-effect-transistor
MSM	measure stress measure
NMP	non-radiative multi-phonon
NMR	nuclear magnetic resonance
OPAMP	operational amplifier
OTF	on-the-fly
PDF	probability density function
PES	potential energy surface
PSD	power spectral density
RAM	random access memory
RTN	random telegraph noise
RTS	random telegraph signal
SILC	stress induced leakage current
SIMS	secondary ion mass spectroscopy
SMU	source measure unit
SNR	signal-to-noise ratio
SPICE	simulation program with integrated circuit emphasis
SRH	Shockley–Read–Hall
TCAD	technology computer-aided design
TDDS	time-dependent defect spectroscopy
TDRC	thermal dielectric relaxation current
USB	universal serial bus
WKB	Wentzel–Kramers–Brillouin
XPS	X-ray photoelectron spectroscopy

CHAPTER 1

INTRODUCTION

In this chapter a short summary of the history of the development of the metal-oxide-semiconductor (MOS) field-effect-transistor (MOSFET) will be given, followed by an overview of reliability issues in state-of-the-art MOSFETs. Finally, the focus of this work will be outlined and the structure of the thesis will be given.

1.1 History of the Field Effect Transistor

Early concepts of amplifying devices employing the field effect were first proposed in the 1920s and 1930s by Lilienfeld [1] and Heil [2]. The effect was not well understood at the time and also manufacturing of such devices was not possible. It required further development of the understanding of semiconductor physics until the field effect could be demonstrated by Schockley and Pearson [3] in 1948. In the basic field effect experiments, as Schockley and Pearson performed, a few micrometer thin strip of semiconducting material was used, separated from a gate electrode by thin sheets of quartz, mylar, mica, barium-, or strontium-crystals. The semiconducting strip was contacted at its ends using metal electrodes. These Ohmic contacts allow a majority current to flow, with its amplitude determined by the density of the majority carriers in the semiconducting strip, which in turn can be controlled by a bias applied at the gate electrode. This enabled not only studies on the modulation of the majority carriers, but also on defects located at the surface of the semiconducting strip. However, employing the conduction of majority carriers significantly limits the geometry of such devices. The semiconductor strip has to be very thin, as otherwise the current flux is not confined to the region affected by the electric field but spread over a wider region. In this case the field effect is not observed anymore and the majority current can not be controlled by the gate bias. In case of silicon, the maximum strip thickness is around 3 µm [4].

An alternative application of the field effect aims at the creation of a so-called inversion layer, which consists of minority carriers. Such inversion layer devices do not suffer from this limitation, as the thickness of the inversion layer is limited by the penetration depth of the electric field. However, the Ohmic contacts used in the early devices did not enable the formation of an inversion layer, as minorities could only be provided by thermal generation in the semiconductor bulk. But the forming of an inversion layer has been observed as a parasitic effect in bipolar junction transistors (BJTs) that time, which enabled to explore this effect [5]. In 1955, Ross proposed electrostatic creation of an inversion layer in the base region of a BJT [6]. Another important observation made in the 1950s by Atalla *et al.*, is the impact of the thermal oxidation on the interface trap density. By evaluating the high temperature oxidation of silicon surfaces it has been observed that the interface trap density can be significantly reduced [7]. The combination ob both, the exploration of the field effect and the initial attempts to create MOS structures paved the way for the successful development of MOSFETs in the 1960s.

The first MOSFET structures were shown by Kahng and Atalla in the 1960s [8, 9]. These early MOSFETs suffered from significant variations in their characteristics due to a high density of interface defects prevalent at the Si/SiO_2 interface and also from of oxide charges originating from ionic contamination. The first commercial process enabling the fabrication of MOSFETs with stable characteristics was shown in 1963 by Deal [10]. One year later, in 1964, Miller and Barson found that the ion contamination can be considerably reduced using getters [11] and Kerr and Young observed that this could be improved further by applying an electrical field to the gate during annealing [12]. Further notable improvements in the 1960s were the introduction of complementary MOS (CMOS) technology by Wanlass [13], which enabled lower power dissipation in logic circuits compared to NMOS-logic, and eventually higher switching speed and density. Another significant improvement of the fabrication processes has been achieved by the introduction of the polycrystalline gate [14], as now MOSFETs could be produced self-aligned by using the gate as a diffusion mask for the source and drain doping processes. The introduction of the floating gate [15] paved the way for non-volatile data storage, and the DRAM cell invented by Dennard [16] enabled a high storage density for volatile memory. A further milestone for controlled device fabrication is the application of ion implantion [17] for creating doped areas, as this offers high flexibility in transistor design and improves the precision in controlling the amount and depth of the dopands incorporated into the structure. A notable invention is also the charge coupled devices (CCDs) by Boyle and Smith [18], enabling a number of applications, including shift registers and image sensing.

Since the 1970s, the continuous improvements of the fabrication processes and MOS technology has allowed decreasing the feature size of integrated MOSFETs from around 10 µm down to below 10 nm today. This development required several updates of the techniques and equipment used for lithography as typical gate lengths became shorter than the wavelength of UV light. The enhancements include shorter wavelength light sources, phase-shift masks, optical proximity correction and immersion lithography. Not only the geometry, but also the layer thicknesses had to be significantly reduced. For instance commercial insulating layers exhibit an oxide thickness of only few nanometers. However, pure SiO_2 or SiON insulating layers with a physical thickness below 2 nm exhibit a drastic increase in tunneling current. To mitigate this, high-k gate dielectrics have been introduced, which allow for smaller equivalent oxide thicknesses without suffering the increased leakage currents. Another detrimental effect in scaled planar MOSFETS are short channel effects like drain-induced barrier lowering (DIBL) and gate-induced drain leakage (GIDL). To suppress such short-channel effects, the principal shape of the MOSFETs has been altered, e.g. in FinFETs, to improve electrical control over the channel.

As device miniaturization seems to approach its physical limits, the focus towards further improvement of the performance in logic devices is shifting to novel materials and new concepts. Next to silicon, other materials have recently emerged for applications in electronic power devices. For such applications, silicon carbide and gallium nitride show very promising properties. As both materials exhibit a larger bandgap than silicon, such devices can sustain higher electric fields. An additional decisive advantage of SiC over Si is also the higher thermal conductivity, making SiC more suitable for high power applications. Finally, it is worth mentioning that 2D materials are also considered for novel MOS devices. One advantage is that the interface between the 2D material and the insulator does not exhibit interface defects as these two materials are connected by van der Waals bonds, instead of having covalent bonds as present at Si/SiO_2 interfaces. However, the development of such devices is still in its infancy and requires significant improvement of the fabrication processes in order to establish this technology.

At the time of writing, power MOSFETs based on SiC and GaN are commercially available, GaAs field-effect transistors (FETs) are available for high frequency applications and SiGe-on-insulator substrates are used in logic transistors to improve mobility and decrease leakage. Although the device performance is continuously improved, the presence of a number of defects cannot be avoided. As such defects can affect the long-term stable operation of devices, an understanding of their behavior and impact on the device performance provides an essential input for circuit and application engineers. Over the recent years, four main reliably issues have been distinguished and will be discussed next.

1.2 Reliability Issues in Field Effect Transistors

The stable and performant operation of devices and circuits can be affected by a wide range of reliability issues. Potential sources of malfunction of devices are mechanical stress, elevated temperature, harsh electrical stress or radiation. The operation of devices under such conditions can negatively impact the reliability of interconnects, can lead to an increase of the oxide leakage currents and in the worst case even to oxide breakdown, and can significantly alter the transfer characteristics of the device. The latter can manifest as a shift in threshold voltage, reduced sub-threshold slope and decreased on-current. The focus of this work is placed on the last mentioned issues, which are considered electrical issues and are caused by defects in the atomic structure of the device. These defects can capture and emit charge and thus lead to changes in the device performance. In this context the following four main device performance degradation mechanisms are typically investigated:

- Bias temperature instability (BTI): BTI degradation manifests itself when a bias is applied at the gate of a device. The resulting electrical field in the oxide causes defects located within the oxide to capture a charge either from the channel or the gate. These trapped charges reduce the threshold voltage and the mobility in the channel, thus affect the transfer characteristics of the device. The effect of charge trapping is accelerated by both gate bias and temperature and is at least partially recoverable when the gate bias returns to a low value. BTI can be separated into positive BTI (PBTI) and negative BTI (NBTI) depending on the sign of the gate bias applied and is relevant in both nMOS and pMOS devices. For silicon, the pMOS/NBTI combinations shows the strongest degradation due to the defect density and energetic alignment of the hole defect band [19, 20, 21].

- Random telegraph noise (RTN): The same type of defects responsible for BTI cause an effect called RTN, when their energy level is located close to the Fermi level. Even at static bias conditions the defects randomly capture and emit charge, causing noise in the channel current. In large devices this is seen as 1/f noise, while in small gate area devices RTN can be observed as discrete steps in the current. The noise produced by the devices causes various issues in integrated circuits, e.g increased failure probabilities in SRAM and jitter in ring oscialltors. The effect has attracted growing attention lately as the amplitude of the effect increases with device scaling [22, 23, 24].

- Stress induced leakage current (SILC) or trap assisted tunneling: Defects in the oxide not only affect the channel of the device, but can also aid tunneling of charge between the channel and the gate, either by allowing the carriers to hop between the gate and the channel or by inhibiting direct tunneling between these carrier reservoirs. This leads to increased gate leakage currents which in turn increase power consumption of the devices and can cause thermal failure of the device. For EEPROM and FLASH memory cells in particular, the increased leakage currents due to SILC decrease the retention time of the devices after repeated write and erase cycles [25, 26, 27, 28].

- Hot carrier (HC) degradation: When a drain-source voltage is applied to a MOSFET, carriers in the channel are accelerated towards the source or the drain region, depending on the carrier type. For high source-drain voltages, carriers with high kinetic energy—so-called hot carriers—cause damage close to the interface, where they can either get trapped or break Si-H bonds, and thereby create interface states. The carriers gain the highest energy close to the end of the channel, where most of the damage is typically observed. The interface states can then charge and thus affect the channel similar to BTI defects [29, 30].

1.3 This Work

A wide variety of literature has amassed, devoted to reliability issues in MOSFETs. Many of these studies focus on characterizing the impact of defects on certain technologies, often using large area devices. However, as the gate area of MOSFETs continues to decrease, studying the underlying behavior of defects becomes increasingly important to explain the reliability of a technology. This is also important for investigating the feasibility of novel technologies. Ideally, a statistically significant amount of single defects should be characterized in sufficient detail to allow further studies, both using atomistic simulations to assess their underlying physical origin and to model their charge trapping dynamics, as well as device simulations to predict their detailed impact on the operating behavior of the devices. For this, more recently developed methods of defect characterization can be employed, which are currently still limited in usability due to the effort required. In this work various aspects of defect characterization including charge trapping models, experimental procedures and methods of data processing are discussed before presenting results obtained using improved methodologies on a number of technologies.

This work is divided in a number of chapters, discussing experimental and mathematical treatment of these defects and their effects in MOSFETs. The structure is as follows. Chapter 2 discusses the main types of electrically active defects found in MOSFETs, how they interact with the device and the issues they cause. Discussion is aided by examples for silicon technology. In Chapter 3, approaches to the mathematical modeling of these defects and their impact on the device are discussed. This includes concepts necessary later, such as the representation of a defect using a Markov chain with capture and emission rates. The following Chapter 4 discusses various experimental approaches available for the characterization of defects. Chapter 5 then discusses the extraction of defect parameters from measurements. This closes the gap between experimental defect characterization and defect modeling. Finally, in Chapter 6, defect characterization studies performed on three separate technologies using combinations of the methods as outlined in the previous sections will be presented.

CHAPTER 2

DEFECTS IN FIELD EFFECT TRANSISTORS

Electrically active defects are inevitably introduced or created during manufacturing of semiconductor devices. Even worse is the fact that such defects can also become newly created during operation at nominal bias conditions. Thus the understanding of the behavior of the defects and their impact on the device performance is essential for further optimization of the MOS transistors. The physical nature of these defects ranges from contaminating foreign atoms in the material or on surfaces to unsatisfied chemical bonds at material interfaces or even stretched bonds in an amorphous oxide. The defects either carry a fixed electrical charge or posses the ability to capture electrical charge during operation. Their presence negatively affects the operation of the device, and may lead to gradual or instant failure of the device if there is too many of them. While advantages in the manufacturing of devices have considerably reduced the number of defects found in mature technologies, a significant number of defects are still prevalent and require accurate physical description to estimate the device lifetime and yield for various operating conditions. For this, suitable defect models are typically calibrated to experimental data.

In the following chapter, an overview of the defect types and candidates which are considered in this work is given. The chapter is split into three sections, devoted to interface-, oxide-, and bulk defects. Agents of each of the defect categories exist in a different electrical and structural environment, hence their physical nature, as well as their impact on the device and the reliability issues they ultimately cause differ. Discussion of the defect classes is aided by the available knowledge on defects in the Si/SiO_2 material system, which is—due to its technological relevance—the best studied material system in semiconductor physics.

2.1 Interface Defects

During device manufacturing, in particular when growing or depositing the insulating layer on top of the semiconductor substrate, the structural mismatch between those materials usually results in defects in the interfacial layer between them. The density of these defects is strongly dependent on the semiconductor and insulator materials

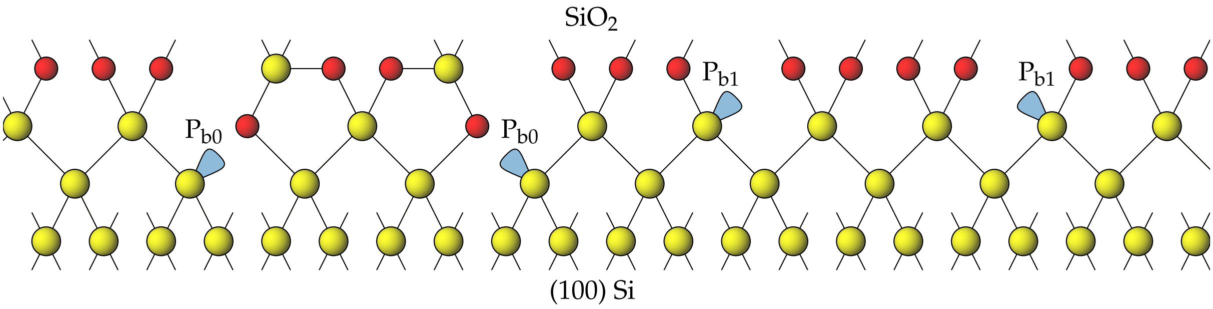


Figure 2.1. Illustration of P_b centers at the (100) Si/SiO$_2$ interface. Shown are P_{b0} ($\cdot Si \equiv Si_3$) and P_{b1} centers ($\cdot Si \equiv Si_2O$). Yellow: Silicon, Red: Oxygen, Blue: Silicon dangling bond. The dangling bonds carry a net spin moment and are thus visible in ESR measurements. Recreated from [32]

used, the processing parameters and other species available during processing, which may passivate these bonds. Control of interface defects is one of the main technological challenges faced when introducing a novel material system. For example, in-situ oxidation and hydrogen passivation has led to a very low achievable interface defect density in silicon devices on the order of $10^{10}\,\mathrm{cm}^{-2}\,\mathrm{eV}^{-1}$ [31]. This is together with the high quality of the oxide one of the main reasons why the Si/SiO$_2$ material system still remains the most successful technology today.

To study interface defects in silicon MOSFETs, ESR measurements have been performed in the 1960s. One of the earliest studies was carried out by Nishi et al. [33], where three distinct defect signatures were found on (111) interfaces. The P_b centers they identified can be found on (110) and (100) oriented surfaces as well [34, 35]. In addition to the original P_b center ($\cdot Si \equiv Si_3$), which is now widely known as P_{b0} center in the context of (110) surfaces, chemically different P_{b1} centers were found on (100) oriented Si interfaces, which are suspected to be oxidized P_b centers, i.e. $\cdot Si \equiv Si_2O$ [32]. An illustration of these defects is shown in Figure 2.1. These so-called “dangling bond” defects involve unpaired bonds located at silicon atoms. P_b centers have two trap levels in the silicon band gap [36], one above mid-gap where they can trap an extra electron and one below mid-gap where they can emit their electron—this property of the defects is called amphoteric. Due to their location in the device, interface defects exchange charge primarily with the semiconductor substrate. Compared to other defects, interface defects act as fast traps, as they are close to the substrate valence and conduction bands, which act as charge reservoirs, and their relaxation energies are low.

The main issue with large densities of interface defects is the decreased controllability of the carrier density in the channel. This is due to trapped charges at interface defects counteracting the charge accumulating at the gate. In addition to this, interface defects increase noise in the channel current due to RTN and decrease the channel mobility. Their number may increase with hot carrier stress due to depassivation, which can lead to long-term degradation of device parameters.

Aside from ESR measurements mentioned above, further methods of interface

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.
TU Bibliothek WIEN Your knowledge hub

defect characterization include capacitance-voltage (CV) measurements, charge pumping (CP) measurements, deep-level transient spectroscopy (DLTS) and direct-current current-voltage (DCIV) measurements, which are discussed in Chapter 4. All these methods have in common that they cannot study the impact of individual interface states on the device behavior, but record average properties of the defects.

2.2 Oxide Defects

Defects may also be located within the insulating oxide layer of a device. These oxide defects or border states[1] may be created intrinsically during oxide growth or can be introduced by impurities or locally distorted chemical bonds within the oxide. Oxide defects can exchange charges both with the gate and the channel. In general, charge trapping at oxide defects proceeds slower than at interface defects. This is partially due to the fact that electron tunneling has to occur during charge exchange due to the spatial separation of the defect and the respective charge reservoirs in the substrate or in the gate, and partially due to the much larger structural relaxation occurring at the defect site. Due to their position in the insulating oxide, the effective trap level of the oxide defects shifts with the applied gate bias. This leads to a strong gate bias dependence of their charge capture and emission rates. Once an oxide defect is charged, Coulomb interaction between its charge and the carriers in the inversion layer below leads to a reduction of the local carrier density, which affects the mobility in the subthreshold regime and the threshold voltage of the device.

The magnitude of its impact on the channel current depends on its distance to the channel and its position relative to the positions of random dopands which also affect the current flux through the channel [38]. Particularly noteworthy is that in devices which have a small gate area the consequences of charge trapping at oxide defects can lead to a particularly large reduction in the source-drain current, even for single charged defects, due to their effect on the percolation path in the channel. On average, the smaller capacitance of these devices results in each defect having a larger impact on the surface potential, and thus the threshold voltage shift. While exact calculation of their effect on the channel in a particular device is not possible due to the random distribution of the dopands, the expected impact of trapped oxide charges can be estimated using the charge sheet approximation (see Section 3.2.2). It has to be noted that this approximation cannot replace a detailed statistical analysis of single defects and their impact on the device performance. Typically, the charge sheet approximation underestimated the impact of the defects on the device [39].

The observed transition times of oxide defects are widely distributed, even more so than the step heights. This is mainly because each defect experiences a slightly different local environment in the amorphous oxide, which strongly affects their properties.

---

[1] *Border states* refers to near-interface oxide defects which are able to communicate with the semiconductor or the gate [37]

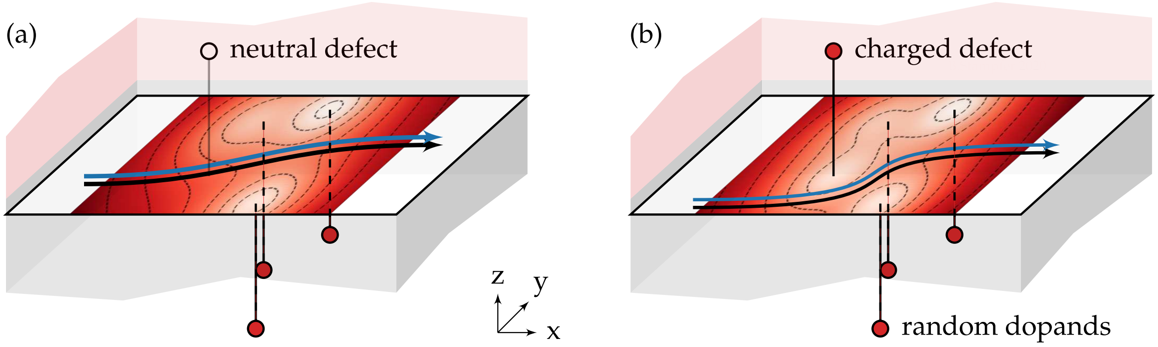


Figure 2.2. Illustration of the channel carrier densities in a scaled device and the resulting current percolation path. **(a)** Random dopands lead to locally decreased carrier densities. The neutral defect has no influence on the channel. **(b)** The defect is now charged and increases the restriction of the percolation path, leading to a decreased drain current and a higher threshold voltage.

In addition, the different positions of the defects in the oxide affect the tunneling probabilities between them and the carrier reservoirs. Thus, oxide defects can on the one hand exhibit very short charge transition times which might affect the momentary behavior the device, but on the other hand very large transition times which can cause long term shifts in the characteristics of the device. Momentarily, they cause increased noise in the drain-source current (RTN) and may lead to hysteresis-like switching behavior. During long term operation, slower defects can become charged, leading to degraded transfer behavior (BTI).

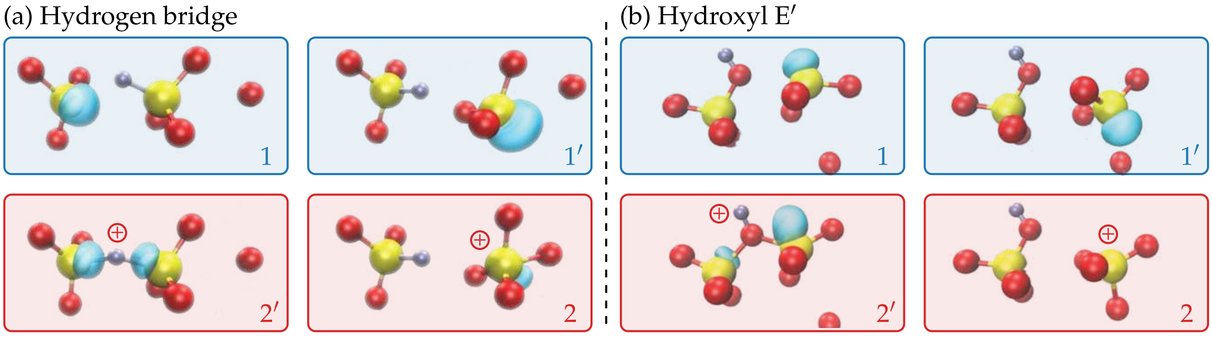
It should be mentioned that the current discussion of charge trapping mainly considers charge exchange between defects and the gate or the conducting channel. However, charge transitions where the defect captures a charge from the gate and emits it into the channel, and vice versa, can also happen. This is possible with intermediate defect to defect transitions. This case is referred to trap-assisted tunneling, and can be observed as leakage current through the insulator. TAT is typically observed as an increase of the gate leakage current with gate bias at low electric fields up to 4MV/cm. At higher electric fields Fowler-Nordheim tunneling starts to dominate and obscures contributions from TAT. Charged defects may further act as an additional barrier for direct tunneling, see Section 4.1.1.

For the electrical characterization of the effects of oxide defects, a number of measurement methods are available, as outlined in Chapter 4. These include also single defect characterization methods, such as RTN and time-dependent defect spectroscopy (TDDS), which can be used to study the trapping behavior of individual defects.

2.2.1 Oxide Defects in SiO_2

Density functional theory (DFT) studies on oxide defects in Si/SiO_2 MOSFETs have identified a number of possible defect candidates for SiO_2 [40, 41, 42]. The two

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.
TU Bibliothek WIEN Your knowledge hub


(a) Hydrogen bridge

(b) Hydroxyl E′

Figure 2.3. Ball-and-stick models of two candidates for hole trap in amorphous SiO$_2$. **(a)** The hydrogen bridge and **(b)** the hydroxyl E′ center. Here, yellow are the silicon, red are the oxygen, and gray are the hydrogen atoms. Additionally, the blue clouds represent the spin density for neutral defects, i.e. the location of the unpaired electron, while for the positive defects it shows the distribution of the captured hole. Both defects may exist in four configurations, with two of them neutral (states 1 and 1′) and two of them charged (states 2 and 2′). To change between two states with the same charge, the defects transition between their regular and puckered configuration [43, 44]. For this, a silicon atom moves through the plane spanned by its three neighboring oxygen atoms. Transitions between the neutral and charged states additionally require a charge transfer with the substrate or gate. Originally published in [BSJ1] and adapted from [45]

most likely candidates—the hydrogen bride and the hydroxyl E′ center—are shown in Figure 2.3. It can be seen that, unlike interface defects, their more complex atomistic structure allows them to exist in more than one configuration with the same net charge. Usually, this is due to one of the silicon atoms transitioning through the triangle spanned by three oxygen atoms to enter a so-called puckered configuration. Due to the amorphous nature of SiO$_2$, these defects show a wide variation in their properties, i.e. their trap levels and transition barriers, which leads to the wide variety in observed time constants mentioned above.

2.3 Semiconductor Bulk Defects

Defects in the semiconductor bulk material are of minor concern for the operation of MOS devices, but are mentioned here for the sake of completeness. As the name suggests, bulk defects are located within the crystalline semiconductor material. They are in part intentionally introduced during device manufacturing in the form of dopands, to control the type, the conductivity and other parameters of a layer or volume. Apart from the introduction of dopands, deep level impurities are unintentionally introduced due to contamination. As bulk defects mainly affect the lifetime of minorities, they play a secondary role in inversion layer devices. However, in devices where a certain minority lifetime is required for proper operation, e.g. CMOS optical sensors, random access memory (RAM) or photodiodes, bulk defects are of significant importance. For the characterization of bulk defects electrical methods, e.g. DLTS, and various physical or chemical methods can be used to quantify their presence and concentration. This kind of defects will not be considered more closely throughout this thesis, as the main focus is on the reliability of MOS devices.

CHAPTER 3

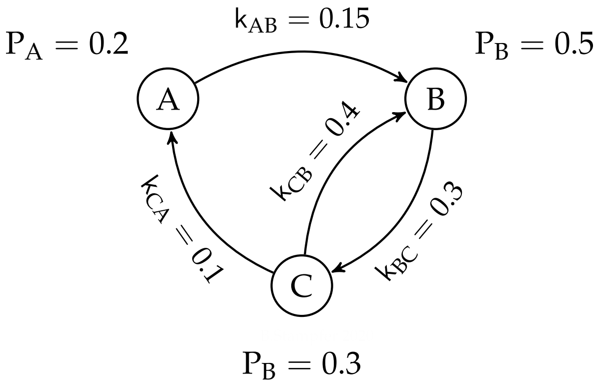
MODELING AND SIMULATION OF DEFECTS

To describe the impact of defects present within a device on its function and behavior, the charge trapping kinetics of the defects and their interaction with the device have to be modeled. For this, empirical and physics based models have been proposed. While empirical models typically rely on simple analytic formulas which are used to describe the measurement data, physics-based approaches aim at the explanation of the behavior of single defects and describe the observed behavior by the superposition of the impact of many such defects. While the former is a more straight-forward approach and much faster to calculate, it lacks in accuracy and secondary effects like, e.g. the saturation of the drift of the threshold voltage with increasing stress time, are usually not considered in such models. Thus, physics-based approaches are preferred for the scientific investigation of the defects, and will be used this work.

In the first section of this chapter, models describing the trapping behavior of individual defects are discussed. Section two then outlines how defects can be studied and used in computer simulations. Finally, in section three, the defect centric model, which phenomenologically describes the statistical distribution of the threshold voltage shift observed in an ensemble of devices, is discussed.

3.1 Defect Models for Charge Trapping

Charge trapping models aim to describe the charging and discharging dynamics of individual defects. Common among all defect models which are discussed within the scope of this work is the usage of a continuous-time Markov chain [46] to mathematically describe the defect. The fundamental difference between the models is in the interpretation of the transition rates between the individual states, as well as in the number of possible states in the Markov chain. Modeling the defects using Markov chains implies that the Markov property is assumed for the defects. This is that the defects retain no memory of their past and their behavior depends solely on the state they are currently in—they are memoryless. Makrov chains consist of a finite number



$$\vec{I} = \begin{bmatrix} A \\ B \\ C \end{bmatrix} \quad \vec{P} = \begin{bmatrix} 0.20 \\ 0.50 \\ 0.30 \end{bmatrix}$$

$$\mathbf{K} = \begin{bmatrix} 0.85 & 0.15 & 0.00 \\ 0.00 & 0.70 & 0.30 \\ 0.10 & 0.40 & 0.50 \end{bmatrix}$$

Figure 3.1. Example of a Markov process with three states, together with the corresponding states $\vec{I}$, the state vector $\vec{P}$ in equilibrium, and the transition matrix $\mathbf{K}$.

of states, in any of which the defect has to be at any time. Mathematically, the Markov chain can be described in terms of probabilities. The probabilities of being in any of the n states $\vec{I}$ is given by the state vector $\vec{P}(t)$. It follows that the expectation values P_i of being in any state i have to sum up to unity:

$$\sum_i P_i(t) = 1. \tag{3.1}$$

The transition rates from state i to j are k_{ij} and can be written as the $n \times n$ transition matrix $\mathbf{K}$. An example of a Markov chain in equilibrium is shown in Figure 3.1. The temporal evolution of $\vec{P}(t)$ can be described by the so-called master-equation [47]:

$$\frac{\mathrm{d}P_i(t)}{\mathrm{d}t} = \sum_{i \neq j} \left(P_j(t) k_{ji} - P_i(t) k_{ij} \right). \tag{3.2}$$

This equation contains the individual transition rates k between the states, which are described by the physical defect model. By assigning the diagonal elements in $\mathbf{K}$, related to the dwelling times, i.e. the time that no transition occurs,

$$k_{ii} = 1 - \sum_{i \neq j} k_{ij}, \tag{3.3}$$

the master equation can be written in vector form:

$$\dot{\vec{P}} = \mathbf{K}^T \vec{P}, \tag{3.4}$$

with $\mathbf{K}^T$ being the transpose of $\mathbf{K}$. With the general mathematical formulation at hand, a link between its parameters and the physical behavior of the defects has to be established. Thus, in the following the physical models describing the states in the Markov chain and the transition rates between them are discussed.

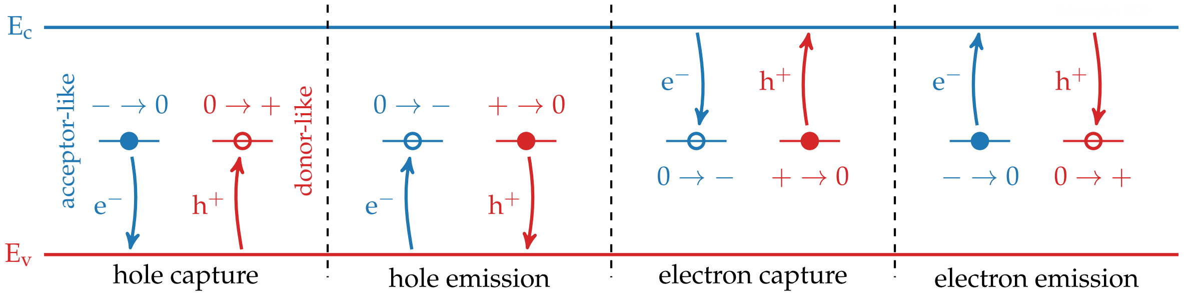


Figure 3.2. Schematic band diagram showing the capture and emission processes for acceptor-like traps (Blue) and donor-like traps (Red). Electron traps change between neutral and negative charge states while hole traps change between neutral and positive charge. States of the defects are shown as they are prior to the indicated transition.

3.1.1 The Shockley-Read-Hall Model

First proposed in 1952, the Shockley–Read–Hall (SRH) model [48] provides a statistical description for the recombination of electrons and holes in a semiconductor through a recombination center in the band gap. In this model, defects are assumed to have a charged and a neutral state. Defects which are electrically negative in their charged state are termed acceptor-like traps and defects which can be positively charged are termed donor-like traps. However, the defects are also frequently called electron- or hole-traps, depending either on their charge in the charged state, or with which band they primarily interact with[1]. The model itself distinguishes between four separate processes, as shown in Figure 3.2:

- Hole capture: A hole moves from the valence band to the trap. This is equivalent to an electron moving from the trap to the valence band and recombines with a hole there.
- Hole emission: A hole moves from the trap to the valence band. This is equivalent to an electron moving from the valence band to the trap, thereby generating a hole there.
- Electron capture: An electron moves from the conduction band to the trap.
- Electron emission: An electron moves from the trap to the conduction band.

Using the hole capture process as an example, the capture probability for a defect located within the band gap is modeled simply as the product of the thermal velocity $v_{\mathrm{th}} = \sqrt{8k_{\mathrm{B}}T/(\pi m)}$ of the carriers in the reservoir and a capture cross section σ:

$$c_p = v_{\mathrm{th}}\sigma. \tag{3.5}$$

---

[1] Note that this leads to some ambiguity in terminology, as an electron trap *emitting* to the valence band and a hole trap *capturing* from the valence band describe the same process, for example.

For the thermal velocity, k_B is the Boltzmann constant, T the lattice temperature, and m the effective mass of the carrier. From this, the capture rate can be obtained by integrating over the interacting band:

$$k_{12} = \int_{-\infty}^{E_\mathrm{v}} c_p(E) f_p(E) g_p(E)\,\mathrm{d}E. \tag{3.6}$$

Here, f_p is the hole occupancy probability, given by the Fermi-Dirac distribution in thermal equilibrium, and g_p the density of states. Similarly, the emission rate can be expressed as

$$k_{21} = \int_{-\infty}^{E_\mathrm{v}} e_p(E) f_n(E) g_p(E)\,\mathrm{d}E. \tag{3.7}$$

Using $f_n(E) = 1 - f_p(E)$ and the property of Fermi-Dirac statistics

$$\frac{f(E)}{1 - f(E)} = \mathrm{e}^{-\beta(E - E_\mathrm{F})}, \tag{3.8}$$

with $\beta = (k_\mathrm{B}T)^{-1}$ leads to

$$k_{21} = \int_{-\infty}^{E_\mathrm{v}} e_p(E) \mathrm{e}^{-\beta(E - E_\mathrm{F})} f_p(E) g_p(E)\,\mathrm{d}E. \tag{3.9}$$

Assuming thermal equilibrium, the principle of detailed balance must hold in equilibrium, which means that the probability of the defect being neutral $p_1 = f(E_1)$ and capturing a carrier at a specific energy must be balanced with the probability of it being charged p_2 and emitting a carrier at that energy:

$$p_1\, c_p(E) f_p(E) = p_2\, e_p(E) f_n(E). \tag{3.10}$$

With this, the emission probability e_p can be related to the capture probability c_p and the energy of the captured carrier E_1:

$$\begin{aligned} e_p(E) &= c_p(E)\, \frac{p_1}{1 - p_1} \frac{1 - f_n(E)}{f_n(E)} && \text{(3.11a)} \\ &= c_p(E)\, \mathrm{e}^{-\beta(E_1 - E_\mathrm{F})} \mathrm{e}^{\beta(E - E_\mathrm{F})} && \text{(3.11b)} \\ &= c_p(E)\, \mathrm{e}^{\beta(E - E_1)}. && \text{(3.11c)} \end{aligned}$$

If Boltzmann statistics hold, p can be expressed as

$$p = \mathrm{N_v} \mathrm{e}^{\beta(E_\mathrm{v} - E_\mathrm{F})}. \tag{3.12}$$

with N_v being the valence band weight. Combining Equation (3.9) and Equation (3.11c), yields

$$k_{21} = \int_{-\infty}^{E_\mathrm{v}} c_p(E) \mathrm{e}^{\beta(E-E_1)} \mathrm{e}^{-\beta(E-E_\mathrm{F})} f_p(E) g_p(E)\,\mathrm{d}E. \tag{3.13}$$

Finally, by approximating all carriers to be located at the valence band edge, which is known as band edge approximation [49], and substituting Equation (3.12) in k_{21} yields simple analytical expressions for the charge capture and emission rates:

$$k_{12} = p\,v_\mathrm{th}\sigma \tag{3.14}$$

$$k_{21} = N_\mathrm{V} v_\mathrm{th} \sigma \mathrm{e}^{\beta(E_\mathrm{V}-E_1)}. \tag{3.15}$$

The model is frequently used to model interface and bulk defects, but has also been extended for oxide defects. To model defects located in the oxide, a Wentzel–Kramers–Brillouin (WKB) tunneling coefficient λ is commonly included in σ to account for elastic tunneling between the defect and the carrier reservoir(s), i.e.

$$\sigma = \sigma_0 \lambda. \tag{3.16}$$

The transmission coefficient for a trapezoidal potential barrier can be written as [50]:

$$\lambda = \exp\left[-\frac{4\sqrt{2m^*}}{3\hbar q F}\left((q\phi_2 - E)^{\frac{3}{2}} - (q\phi_1 - E)^{\frac{3}{2}}\right)\right], \tag{3.17}$$

with the start and end potentials $q\phi_1$ and $q\phi_2$ of the barrier, the electric field F and the effective carrier mass in the oxide m^*. The issue with using this model to describe oxide defects is that its neither able to correctly describe the bias-dependence nor the temperature-dependence of the defects observed in measurements. As compared to experiments, calculated transition times are often magnitudes too small. To model defects located energetically above or below the band gap, modified equations have to be used [50].

3.1.2 The Kirton and Uren Model

It was found that the SRH model does not adequately describe the temperature dependence observed in measurements of oxide defects. An extended approach to explain charge trapping at oxide defects has been developed by Kirton and Uren. They realized that the model has to account for the structural relaxation of the defects and they supposed that the thermal barriers required for the description stem from multiphonon processes. To account for this, they incorporated a phenomenological Boltzmann factor

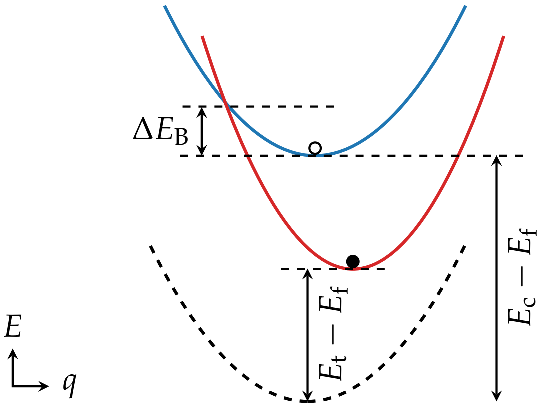


Figure 3.3. Kirton-Uren model illustrated in a configuration coordinate diagram. A defect exchanges an electron with the reservoir. The position of the parabola depends on the carrier energy. The capture barrier ΔE_B accounts for thermal activation of the capture process. Blue: Empty defect, electron in the conduction band. Red: Electron captured at the defect site. **Dashed:** Energy zero of the system: empty defect, electron at the Fermi level. Recreated after [51]

in the effective cross section [52, 51]:

$$\sigma = \sigma_0 \lambda \mathrm{e}^{-\beta \Delta E_B}. \tag{3.18}$$

Here, ΔE_B is the capture barrier of the defect responsible for the thermal activation, as illustrated in Figure 3.3. This additional term allows to model the temperature dependence of 1/f noise and the average charge transition times. However, it introduces a correlation between capture and emission time which can not be observed in measurements. Further, the barrier is independent of gate bias and thus still does not correctly describe bias dependence of charge trapping.

3.1.3 The NMP Model

To describe the bias dependence as observed in measurements the phenomenological energy barrier as introduced by Kirton and Uren is inadequate, and has to be replaced by one calculated in dependence on the effective trap level of the defect which is shifted with the gate bias. Thus, the model has been extended accordingly to the non-radiative multi-phonon (NMP) model. The NMP model [53, 54, 55] is based on the physical description of the system consisting of all electrons and nuclei involved in the charge transition.

Due to the difference in timescales at which the electrons and nuclei act, the transition can be split into an electronic and a vibrational part using the Born-Oppenheimer approximation [56]. In this approximation, the nuclei move in a potential given by their positions, the adiabatic potential energy. Local minima of this surface correspond to the previously mentioned states in the abstract Markov-chain description. The transition from one state to another is associated with a thermal barrier which is determined by a

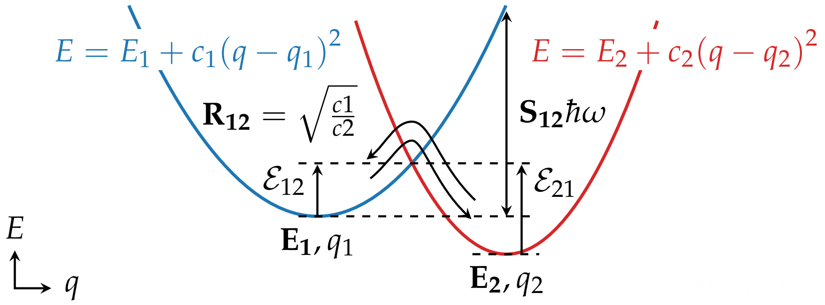

$E = E_1 + c_1(q - q_1)^2$
$E = E_2 + c_2(q - q_2)^2$
$\mathbf{R_{12}} = \sqrt{\frac{c1}{c2}}$
$\mathbf{S_{12}}\hbar\omega$
$\mathcal{E}_{12}$
$\mathcal{E}_{21}$
$\mathbf{E_1}, q_1$
$\mathbf{E_2}, q_2$
E
q

Figure 3.4. Parabolic potential energy surfaces (PESs) for the charged (red) and a neutral state (blue) as used in the NMP defect model, together with the parameters of the model.

transition state on the potential energy surface (PES), i.e. the highest point along the minimum energy transition path. While the potential energy surfaces can be calculated using DFT simulations, they are not directly accessible experimentally. To obtain a usable model, the situation is thus simplified. The multi-dimensional PES is reduced to one dimension along a reaction or configuration coordinate q, described by the lowest energy path between the states. This PES is then replaced by a Taylor expansion around the minimum, truncated after the quadratic term, which enables a description of the defect states as harmonic oscillators. This results in a configuration coordinate diagram with parabolic potential energy curves as shown in Figure 3.4.

Following from the Born-Oppenheimer approximation, the Frank-Condon principle [57, 58] and Fermi's golden rule [59], the transition rate from one state to the other can be written as [60]

$$k_{ij} = A_{ij} f, \tag{3.19}$$

with the electronic matrix element A between the initial and final electronic states and the line shape function f. The electronic matrix element cannot be calculated[2] for the systems concerned, and is thus commonly approximated using a WKB tunneling factor together with a prefactor. The line shape function can be calculated, but is commonly approximated by neglecting nuclear tunneling [63] below the highest point in the minimum energy path. In this classical limit [64], transitions happen at the intersection point of the PES and the line shape function is written as a Dirac function at the energy of intersection. Thus, the energy barrier between the two states can be found by the intersection of the barriers as shown in Figure 3.4. In the semi-classical regime—i.e. if tunneling below the barrier can be neglected—the transition can be modeled using a Boltzmann factor using this barrier energy.

With all these simplifications the model effectively reduces to two parabolas, which needs to be parametrized. A commonly used parametrization of the parabolas consists of

- the ratio of curvatures $R = \sqrt{c_1/c_2}$,

---

[2] Recent works [61, 62] propose a methodology to do so using DFT, for this, however, the detailed atomic structure of the defect has to be known.

- the relaxation energy $S\hbar\omega$ with the Huang-Rhys factor S,
- the ground state energies E_1 and E_2.

Note that this parametrization eliminates the reaction coordinate between the states, which would be necessary for calculating nuclear tunneling but is not needed in the semi-classical approximation. From these parameters, and using $E_{21} = E_2 - E_1$, the energy barrier from state 1 to 2 can be calculated as [50]

$$\mathcal{E}_{12} = \frac{S\hbar\omega}{(R^2-1)^2}\left(1 - R\sqrt{1 + \frac{S\hbar\omega + (R^2-1)E_{21}}{S\hbar\omega}}\right)^2. \tag{3.20}$$

Note that for $R = 1$, Equation (3.20) has a singularity, and in this case the energy can be calculated as

$$\mathcal{E}_{12} = \frac{(S\hbar\omega + E_{21})^2}{4S\hbar\omega}. \tag{3.21}$$

And finally the backward barrier $\mathcal{E}_{21}$ can be calculated from $E_1 + \mathcal{E}_{12} = E_2 + \mathcal{E}_{21}$ to

$$\mathcal{E}_{21} = \mathcal{E}_{12} - E_{21}. \tag{3.22}$$

Using again the valence band as an example, and including the band edge approximation, leads to the final capture and emission rates for the two state NMP defect model:

$$k_{12} = p\, v_{\mathrm{th}}\sigma_0\lambda \mathrm{e}^{-\beta\mathcal{E}_{12}} \tag{3.23}$$
$$k_{21} = N_{\mathrm{V}} v_{\mathrm{th}}\sigma_0\lambda \mathrm{e}^{\beta(E_{\mathrm{V}}-E_1)}\mathrm{e}^{-\beta\mathcal{E}_{21}}. \tag{3.24}$$

3.1.4 The 4-State NMP Model

Single defect measurements, such as RTN and TDDS measurements have shown that individual defects can exhibit more than one characteristic dwelling time in a single charge state. When RTN signals are studied, this behavior can be observed as anomalous RTN (aRTN), where periods of inactivity of the defect follow periods of RTN signals [22]. In TDDS measurements it has further been observed that individual defects seem to disappear for a number of measurements and later reappear [65]. These observations, together with knowledge gained from DFT calculations [45, 55] for suitable defect candidates have led to the introduction of the four-state defect model. In this model, the Markov chain describing the defect has a meta-stable and a stable state for both the charged and neutral charge state, as shown in Figure 3.5a. The transitions between the neutral and charged states, i.e. $1 \leftrightarrows 2'$ and $1' \leftrightarrows 2$, are modeled as NMP transitions, while transitions between states of the same charge are modeled by purely thermal barriers of constant height. The resulting energy profiles along the multiple reaction paths of such a defect are schematically illustrated in Figure 3.5b.

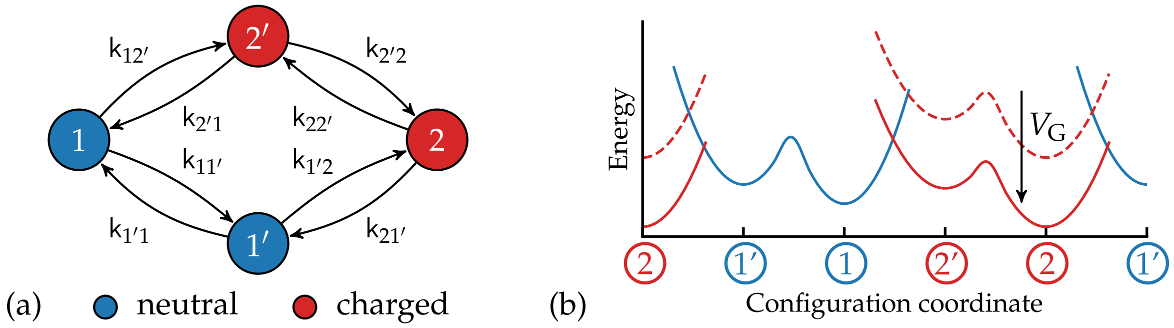


Figure 3.5. Markov chain **(a)** and potential energy surfaces **(b)** of the four-state defect. In addition to the neutral and charged states 1 and 2, there is an additional metastable state for each charge state marked with a prime (′). Transitions between neutral and charged states are modeled using NMP theory, while transitions between states of the same charge are purely thermal transitions described using classical transition state theory. Changing the gate voltage shifts the trap level which leads to a shift in the PES relative to each other.

In this picture, a defect can transit between its charged and neutral states using two separate pathways. This enables the description of complex capture and emission time behavior as observed in RTN and TDDS measurements. In particular, this enables to distinguish between *fixed* and *switching* traps [66]. Fixed traps show bias-dependent capture times, but at low gate bias the emission time can become bias-independent. This behavior can be explained by them using the 1–2′–2 path for both capture and emission. In this case, the emission time is determined mainly by the bias-independent barrier between the states 2 and 2′. Switching traps on the other hand use the same 1–2′–2 path for charge capture, but the 2–1′–1 path for emission, which yields bias-dependent rates for both processes. Using this model the charge trapping kinetics of a number of defects extracted from planar Si MOS devices [50, 67, 68], from devices employing high-k gate stacks [69], but also defects extracted from 2D devices [BSJ7] has been explained.

3.1.5 The Hydrogen Release Model

While the four state model seems to be able to describe the recoverable component of BTI, measurements also reveal a smaller “permanent” degradation of the threshold voltage of the devices [70, BSC3]. The defects which are considered responsible for this permanent component can not be explained using the four state model. However, to explain the permanent component of BTI a hydrogen release mechanism has been recently proposed [70]. In silicon devices, hydrogen is purposely introduced during manufacturing to passivate dangling bonds at the Si-SiO$_2$ interface. It is thus readily available in varying concentrations throughout the device. Various studies have suggested possible interaction between hydrogen and interface or near interface states even after passivation [71, 72, 73, 74]. With two of the suspected defect candidates—the Hydroxyl E′ center and the Hydrogen-bridge—incorporating hydrogen, it has recently been suggested [74] that such defects may be activated or disabled by available atomic

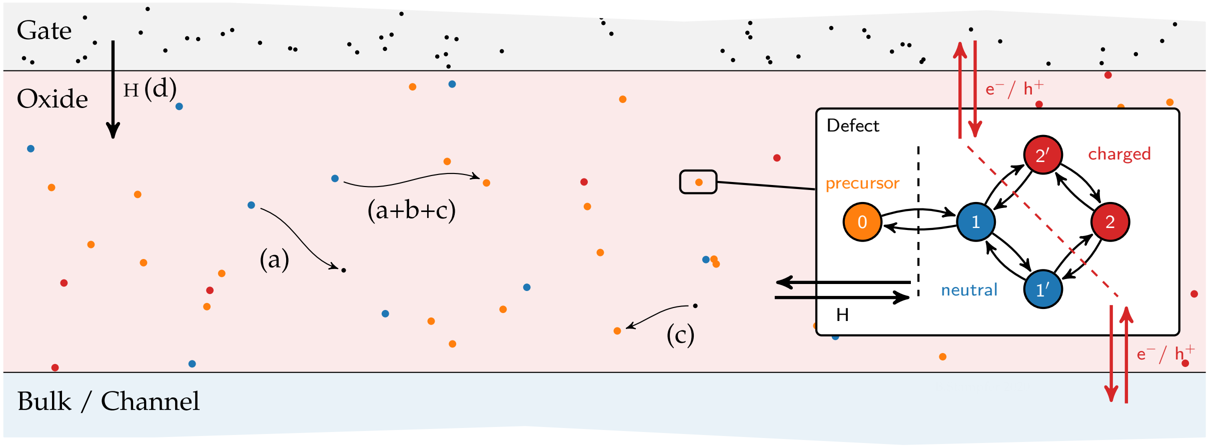


Figure 3.6. Illustration of the hydrogen release model. A number of precursor sites (orange) exist in the oxide, which become defects once they capture a hydrogen atom. Vice versa, in one of the neutral configurations of the defect (blue) there is a realistic probability for the hydrogen atom to detach from the defect **(a)**. Once this happens, the hydrogen can diffuse in the oxide **(b)** and bond to another precursor **(c)**. The high diffusion rate for hydrogen in SiO_2, together with a low barrier of hydrogen capture by the precursors leads to a low probability of free hydrogen at any time. Additional hydrogen is supplied by the gate at harsh stress conditions **(d)**, leading to additional defects. States of the defects are shown as they are before the indicated transitions. **Inset:** Markov chain of the hydrogen release defect. Compared to the four-state defect there is an additional precursor state to or from which the defect can transition by emission or capture of hydrogen, respectively.

hydrogen during operation. The basic idea of the hydrogen release model [70, BSC4] relies on this statement and assumes that hydrogen-related defects may be able to surrender their hydrogen atom in the neutral state. This hydrogen can then activate preexisting defects at other locations in the oxide or create additional defects at the interface by depassivating Si dangling bonds, i.e. $\mathrm{Si{-}H} + \mathrm{H} \longrightarrow \mathrm{Si} + \mathrm{H_2}$. In addition, more hydrogen might be released from the gate into the oxide at harsh stress conditions, thereby increasing the number of active defects in the oxide or at the interface. An illustration of the model is shown in Figure 3.6. Due to the high hopping rate of atomic hydrogen in SiO_2 [75] the hydrogen in the oxide is available to all precursors. Together with low barriers for hydrogen capture [41], this leads to a system governed by the release of hydrogen from neutral defects or the gate.

For a mathematical description of the model, the very low diffusion barrier of interstitial hydrogen can be approximated as zero. This allows to treat each material as a single reservoir for free hydrogen. The total amount of hydrogen, which is assumed to stay constant, is thus the sum of the hydrogen in the gate reservoir H_R, the interstitial hydrogen in the oxide H_0 and hydrogen trapped at each defect site $H_{\mathrm{T},i}$:

$$H_\mathrm{tot} = H_\mathrm{R} + H_0 + \sum_i H_{\mathrm{T},i}. \tag{3.25}$$

The exchange between the gate reservoir and the oxide can be formulated as a rate equation:

$$\frac{\partial H_\mathrm{R}}{\partial t} = -k_{\mathrm{R}0} H_\mathrm{R} + k_{0\mathrm{R}} H_0 \tag{3.26}$$

with temperature dependent diffusion rates $k_{\mathrm{R}0}$ and $k_{0\mathrm{R}}$ to and from the oxide, respectively. To simplify the implementation, the defects may be decoupled from the hydrogen system. For this, the defects are implemented by using a regular four-state defect model and a probability of possessing hydrogen is assigned to each defect (or defect site). The expectation value of hydrogen atoms at the defect site can then be calculated using $p_{1,i}$, the probability of the defect being in the stable neutral state, as

$$\frac{\partial H_{\mathrm{T},i}}{\partial t} = -k_{01} p_{1,i} H_{\mathrm{T},i} + k_{10} H_0 (H_{\mathrm{T,max},i} - H_{\mathrm{T},i}), \tag{3.27}$$

with hydrogen capture and emission rates k_{01} and k_{10}, and the maximum number of hydrogen at the defect site $H_{\mathrm{T,max},i}$ (= 1 for single defects).

The effective charge impacting the device is determined by the probability of the defect to be active (to possess hydrogen) and the probability of the active defect to be in a charged state. Thus, to calculate the charge of the defects or defect sites, their probabilities of being charged $p_{2,i} + p_{2',i}$ need to be weighted by their expectation value of attached hydrogen:

$$q_{\mathrm{T},i} = q H_{\mathrm{T},i} \left(p_{2,i} + p_{2',i} \right). \tag{3.28}$$

To give an example, in Figure 3.7 results of a long term experiment characterizing the permanent component of the degradation of a pMOS transistor are shown, together with the results simulated with the hydrogen release model. During the experiment the gate voltage has been cycled between −1.5 V and 0 V, and also the temperature has been varied. As can be seen, the degradation saturates at 300 °C, but increases again at the next stress phase of 350 °C due to release of hydrogen from the gate. This temporal behavior can be explained by the proposed hydrogen release model.

3.2 Defect Simulations

A commonly used tool to investigate the complex behavior of devices and how defects affect their operation is numerical simulation. The kind of simulation ranges from circuit level simulations, where very abstract degradation models are used to capture the essence of the effects on device and circuit properties caused by defects, to *ab-initio* simulations investigating the atomistic behavior of specific defect structures. Inbetween are physical device simulations, which are commonly used in conjunction with measurements to calibrate physical models and to find defect parameters.

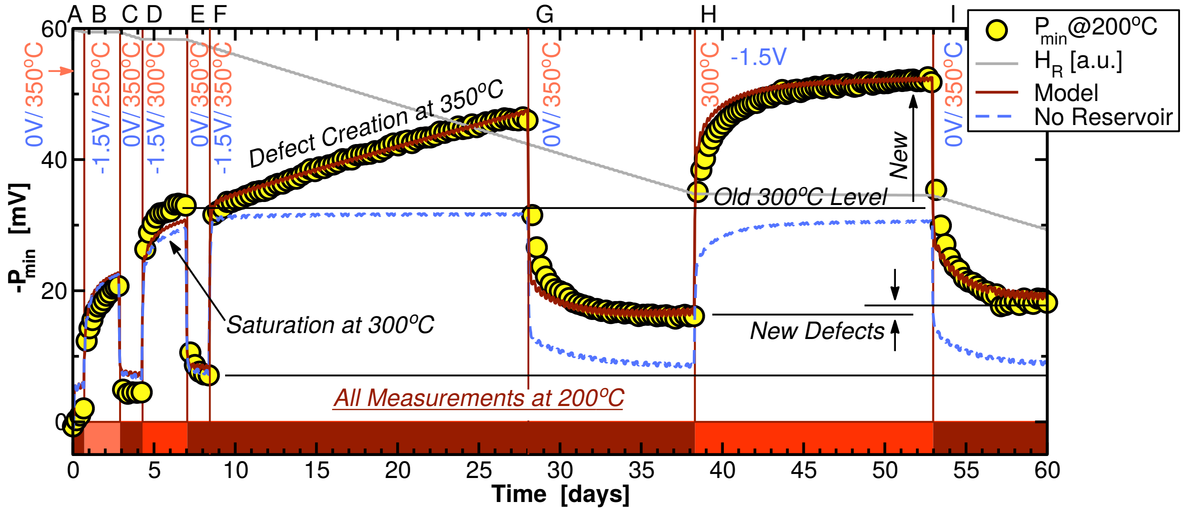


Figure 3.7. Experiment on long-term degradation of a 10 um × 10 um planar pMOS transitor. Measured permanent degradation component (circles), together with simulation results extracted using the hydrogen-release model (red). From [76]

3.2.1 Ab-Initio Simulations

Ab-initio simulations can be used to study the properties of materials or molecules. In microelectronics they have been traditionally used to identify the electronic structure of materials, material interfaces and specific defect configurations. This is done by approximately solving the Schrödinger equation for a number of atoms of the material, together with appropriate boundary conditions, e.g. to represent bulk material. This allows, among other things, to estimate energy levels of defects, barrier energies or formation mechanisms of defects. The most commonly used method for defect studies is the density functional theory (DFT) method, which is, compared to other ab initio methods, computationally inexpensive and allows to simulate thousands of atoms on current hardware [77].

3.2.2 Defects in Device Simulation

Defect models, e.g. the SRH, NMP and hydrogen release models, as discussed in the previous section, are typically used in TCAD (technology computer-aided design) simulations which aim at simulating the degradation of components on the device level. For this, a device model taking geometry, material properties, dopand densities, and mobility models into account is used in conjunction with a defect model as outlined in Section 3.1 to simulate the altering of the device. Such simulations are often performed in a Monte Carlo [78] fashion by placing a large number of defects with randomly drawn parameters in the device. This, however, requires prior knowledge about the distributions of defect parameters and their density. To obtain this information TCAD simulations can be calibrated to explain dedicated measurements, but also DFT studies on the suspected defects can be used to estimate defect parameters.

The simulations typically start with the calculation of an initial equilibrium state of the system. Afterwards, the evolution of defects in the device is calculated based on the applied bias conditions and temperatures for each following time step. Depending on the density of defects in the devices, the calculations may have to be performed self-consistently. For high densities of defects the charges captured by the defects can significantly alter the device electrostatics. If this is the case, the charge of the defects has to be considered in the Poisson equation as otherwise the results will not be accurate. If possible, however, calculations are performed in a non self-consistent manner as this decreases the convergence time of the solver.

To estimate the impact of the defects on the channel without knowing the exact distribution of the random dopands in the device, the charge sheet approximation can be used. It allows to approximate the threshold voltage shift at low defect concentrations and for little deviations of the results from the ones a defect free device. In this approximation, a charge trapped at a defect is considered to be distributed homogeneously over the entire gate area in a thin sheet parallel to the insulator/semiconductor interface. The resulting threshold voltage shift can be simply calculated from the capacitance between the charge sheet and the gate:

$$\Delta V_{\mathrm{th}} = \pm q_0 \frac{1 - d/t_{\mathrm{ox}}}{C}, \tag{3.29}$$

$$C = \varepsilon_0 \varepsilon_{\mathrm{ox}} \frac{A}{t_{\mathrm{ox}}}, \tag{3.30}$$

with the vacuum permittivity ε_0, the relative permittivity of the oxide $\varepsilon_{\mathrm{ox}}$, the gate area A, the gate oxide thickness t_{ox}, the defects distance from the interface d, and the elementary charge q_0. It should be noted that this approach generally leads to an underestimation of the average impact of the defects. For a more accurate estimation of the average impact of a single defect on the threshold voltage shift, the distribution function of step heights has to be determined [79, 39].

3.2.3 Defects in Circuit Simulation

Most simulators used for the development of integrated circuits are based on a simulation program with integrated circuit emphasis (SPICE) [80] written by L. Nagel in 1973. SPICE models describe components within the circuit and thus are physically more abstract than device simulation. They are based on relatively simple analytical models, often with a number of empirical parameters. Even though it is generally accepted that BTI is due to the interaction of single defects with the device, to date compact BTI models do not directly calculate the impact of such defects for this. The impact the defects have, e.g. on the threshold voltage and mobility, is calculated based on model parameters calibrated to device simulation or measurements on a specific technology. However, these models cannot describe many facets of BTI which might be important for the development of circuits considering scaled technologies. The most fundamental problem is that calculating the vast number of defects per device becomes inefficient

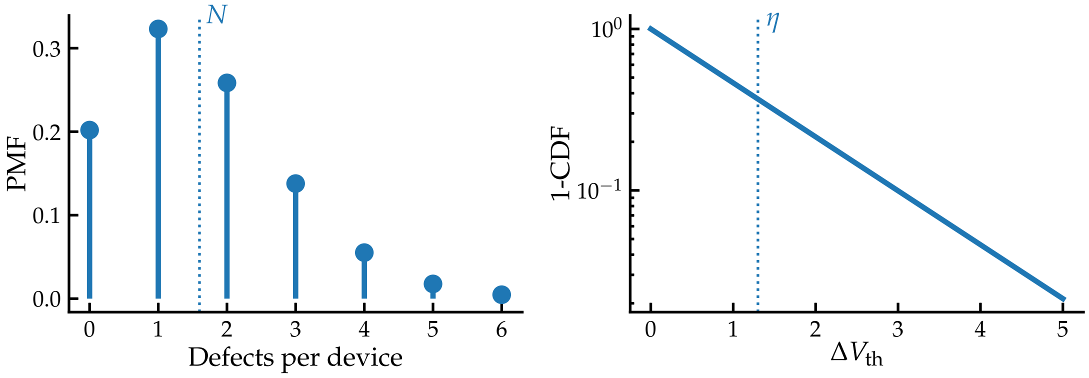


Figure 3.8. Distributions of the defects per device and the step height per defect in the defect centric model. **Left:** The average number of defects in the devices is assumed to be Poisson distributed around a mean value N. **Right:** The impact of the single defects on the threshold voltage shift is given by an exponential distribution with the expectation value η. While the majority of defects ($1 - \frac{1}{e} \approx 63\%$) show step heights below η, some defects produce much larger steps. The variance in both the number of defects and step height per defect underlines the fact that for small devices with few defects on average, some devices can show degradation much worse than the average device.

when it has to be done for a number of transistors. For this more compact descriptions of the phenomena are required. One approach which is based on the observations from single defects is the defect-centric perspective, which will be discussed next.

3.3 The Defect Centric Model

The defect centric model [81, 79] describes the effects of defects on the threshold voltage shift in a statistical manner. It is based on phenomenological descriptions of the statistical properties of defect ensembles, such as the number of defects per device and their individual contributions. This allows both to obtain defect parameters from measured distributions of changes of the threshold voltage, as well as to estimate these distributions from defect parameters, obtained for example from single defect measurements.

In the model, the number of defects per device is assumed to follow a Poisson distribution with mean N, while the effect of a single defect on the threshold voltage is assumed to be exponentially distributed with expectation value η [81], as shown in Figure 3.8. From this, the distribution of threshold voltage shifts in a device can be calculated [79, 82]. One core assumption of the model follows from experimental observations and is the exponential distribution of the threshold voltage shift caused by single defects:

$$f(\Delta V_\mathrm{th}, \eta) = \frac{1}{\eta} \mathrm{e}^{-\frac{\Delta V_\mathrm{th}}{\eta}} . \tag{3.31}$$

TU Bibliothek Your knowledge hub WIEN
Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

Assuming the defects to be independent, i.e. their total effect on ΔV_{th} is given by their superposition, the shift of the threshold voltage for n defects can be found as the (n) autoconvolution of the contributions of the defects:

$$f_n(\Delta V_{\mathrm{th}}, \eta) = f * f * \cdots * f = \frac{\mathrm{e}^{-\frac{\Delta V_{\mathrm{th}}}{\eta}}}{(n-1)!} \frac{{\Delta V_{\mathrm{th}}}^{n-1}}{\eta^n}. \tag{3.32}$$

Alternatively, with the gamma distribution for $p \in \mathbb{N}$

$$\gamma_{p,b}(x) = \frac{b^p}{(p-1)!} x^{p-1} \mathrm{e}^{-bx}, \tag{3.33}$$

with $p = n$ and $b = 1/\eta$ the function yields

$$f_n(\Delta V_{\mathrm{th}}, \eta) = \gamma_{n,1/\eta}(\Delta V_{\mathrm{th}}). \tag{3.34}$$

The probability of having n defects in a device is given by the Poisson distribution:

$$P_N(n) = \frac{\mathrm{e}^{-N} N^n}{n!}. \tag{3.35}$$

The total probability density function (PDF) of the threshold voltage shifts is thus the sum of Poisson weighted Gamma distributions for n defects:

$$f_N(\Delta V_{\mathrm{th}}, \eta) = \sum_{n=1}^{\infty} P_N(n) \gamma_{n,1/\eta}(\Delta V_{\mathrm{th}}). \tag{3.36}$$

Using this model, simple relations between the first moments of the distribution of threshold voltage shifts of a set of devices and the statistical properties of the defects can be found for the number of defects per device:

$$N = \frac{\langle \Delta V_{\mathrm{th}} \rangle}{\eta}, \tag{3.37}$$

and for the average impact of a defect on the threshold voltage:

$$\eta = \frac{\mathrm{Var}(\Delta V_{\mathrm{th}})}{2 \langle \Delta V_{\mathrm{th}} \rangle}. \tag{3.38}$$

The goal is to apply the compact formulas to calibrate the model to measurement data. However, distributions measured using an MSM scheme on actual devices will have additional components due to RTN and measurement noise. This can be included in the model by convoluting the BTI contribution as outlined above with contributions due to RTN and Gaussian noise, as shown in Figure 3.9, which yields

TU Bibliothek, WIEN Your knowledge hub
Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN

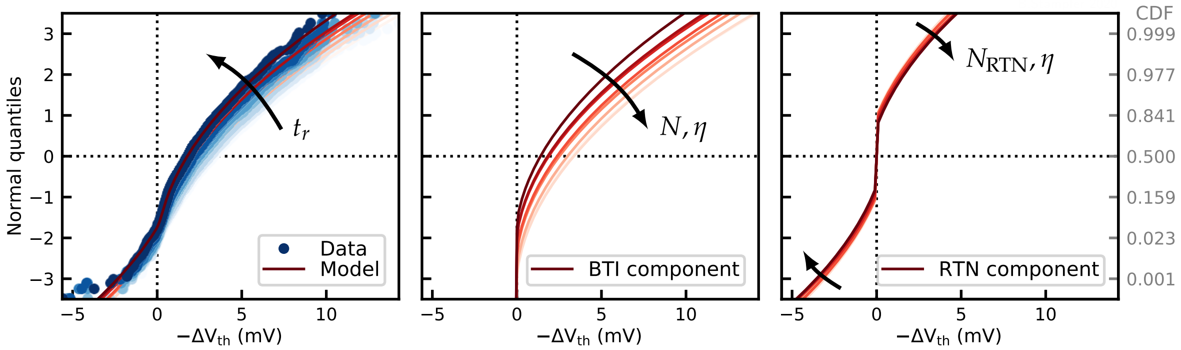


Figure 3.9. Left: Measured distribution of ΔV_{th} after stress in comparison with the defect centric model. **Center:** The one-sided BTI component gets smaller after stress due to decreasing N. **Right:** The symmetric RTN component stays constant throughout recovery.

$$\begin{aligned} p(\Delta V_{\mathrm{th}}) = f_N(&\Delta V_{\mathrm{th}}, \eta) \\ &* g_{N_{\mathrm{RTN}}}(\Delta V_{\mathrm{th}}, \eta) \\ &* h_{\mathrm{Noise}}(\Delta V_{\mathrm{th}}, m, \sigma). \end{aligned} \tag{3.39}$$

The RTN component is similar to the BTI component, but symmetric in ΔV_{th}:

$$\begin{aligned} g_{N_{\mathrm{RTN}}}(\Delta V_{\mathrm{th}}, \eta) = &\sum_{n=0}^{\infty} \mathcal{P}_{N_{\mathrm{RTN}}/2}(n) \gamma_{n,1/+\eta}(\Delta V_{\mathrm{th}}) \\ * &\sum_{n=0}^{\infty} \mathcal{P}_{N_{\mathrm{RTN}}/2}(n) \gamma_{n,1/-\eta}(\Delta V_{\mathrm{th}}). \end{aligned} \tag{3.40}$$

This is because an defect showing RTN can also be charged during the initial $I_{\mathrm{D}}(V_{\mathrm{G}})$ sweep of the measurement sequence, and become neutral after stress. The component accounting for measurement noise is a normal distribution $\mathcal{N}$, with mean m and variance σ:

$$h_{\mathrm{Noise}}(\Delta V_{\mathrm{th}}, m, \sigma) = \mathcal{N}(m, \sigma^2). \tag{3.41}$$

The method, as shown above, describes a single defect distribution. If the oxide stack consists of more than one material, e.g. in high-k devices, the model can be extended accordingly, see for example [83]. In Section 6.2, this method is applied to find the statistical behavior of defects for various stress cases in a high-k/metal gate technology.

CHAPTER 4

METHODS OF DEFECT CHARACTERIZATION

Defects play a substantial role in semiconductor devices. They impact the feasibility of devices made from novel materials as well as the reliable and failure-safe operation of devices manufactured in mature technologies. Hence, it is no surprise that numerous methods have been developed over the last decades devoted to characterization of their physical nature.

In this chapter, methods used for experimental defect characterization will be discussed. The main focus of this chapter will be placed on electrical methods which have been implemented or employed by the author during his PhD studies. The discussion of these methods will also include details on the capabilities required of the measurement equipment. The chapter is split into three sections. In Section 4.1, electrical characterization methods which derive information on the device performance from the channel current in MOSFETs. This includes RTN, BTI/TDDS, and $I_D(V_G)$ measurements. Section 4.2 will discuss electrical methods which measure defect charging currents in MOS capacitor (MOSCAP) or MOSFET structures, e.g. capacitance-voltage (CV), charge pumping (CP), deep-level transient spectroscopy (DLTS), direct-current current-voltage (DCIV), and thermal dielectric relaxation current (TDRC) measurements. Finally, in Section 4.3, physical methods which do not involve direct measurement of device currents are discussed.

4.1 Electrical Methods based on Channel Conductivity

As described in Chapter 2, charged defects located spatially in between the gate and the conducting channel influence the electrostatics of the channel in inversion. This is due to Coulomb interaction between the trapped charge and the channel carriers, which in the most common case effectively lowers the carrier density in the vicinity of the defect. The impact of a defect on the channel current depends on the resulting perturbation of the current percolation path [84]. This affects the drain-source conductivity which can be measured and employed for defect characterization. In this chapter characterization methods based on this effect are discussed. The fact that the influence of such defects increases with device scaling also enables the possibility to observe charge transitions of single defects in these measurements, given the gate area is small enough. Unlike for other methods discussed in this chapter, full MOSFET structures are required for all methods presented here.

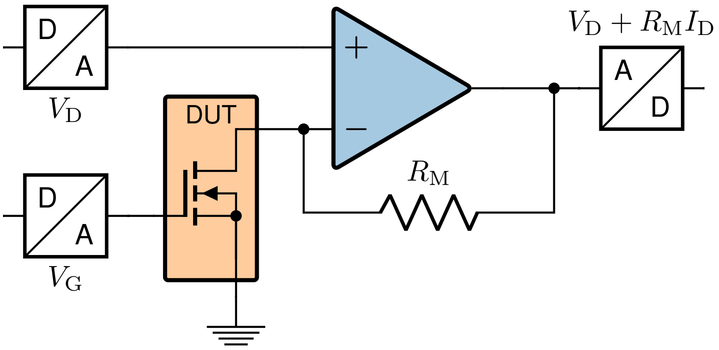


Figure 4.1. Schematic of a constant gate voltage measurement setup for channel conductivity based methods. The drain and gate voltages of the device are controlled by DACs. The gate voltage is applied directly to the gate of the DUT, while the drain current is supplied to the positive terminal of an OPAMP configured as an transimpedance amplifier. Its negative input terminal is connected to the drain connection of the DUT. The voltage output of the transimpedance amplifier, which corresponds to I_D, is read by an ADC. Originally published in [BSJ1]

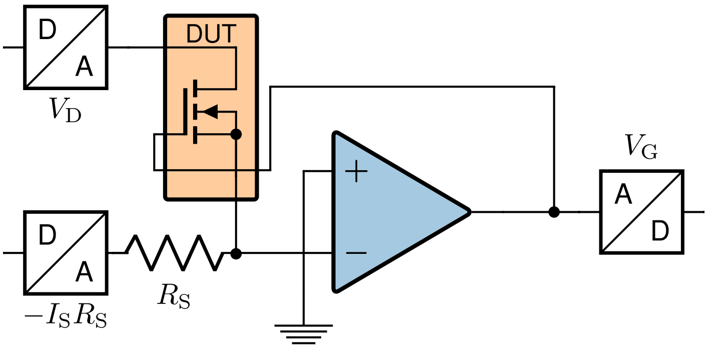


Figure 4.2. Schematic of a constant source current measurement setup for channel conductivity based methods. The drain bias of the DUT is controlled directly, while the gate bias is controlled by the OPAMP. The OPAMP circuit controls the gate bias of the MOSFET in a way that the source current equals the reference current defined by the DAC bias $-I_s R_s$. Originally published in [BSJ1]

The measurements used to observe changes in the channel conductivity rely on recording the drain-source current at constant drain-source bias and device temperature. Two distinct measurement schemes exist for this type of characterization [85]. In the more popular constant-voltage measurement scheme, both the drain-source bias and the gate bias of the DUT are controlled externally, while the drain or source current is converted to a voltage and subsequently measured by an additional digital-to-analog coverter (DAC), see Figure 4.1. In contrast, the constant-current scheme aims at controlling the drain-source bias directly, while the gate voltage is controlled via a feedback loop of an operational amplifier, see Figure 4.2. The control input for the feedback loop is a second external voltage which defines the drain or source current of the DUT. The gate voltage applied via the feedback loop is measured directly by an ADC or an external device like an oscilloscope. The decision which method should preferably be used is difficult: The constant-voltage scheme has the benefit of simplicity but the drawback is that to obtain the threshold voltage shift ΔV_{th} the obtained current has to be mapped with a pre-recorded $I_{\mathrm{D}}(V_{\mathrm{G}})$ curve. However, the characteristic of the $I_{\mathrm{D}}(V_{\mathrm{G}})$ can alter during the measurement, i.e. not only the threshold voltage but also the sub-threshold slope can change, which cannot be considered when mapping the current to ΔV_{th}. The constant-current scheme may seem more suited in this case, but due to the feedback loop formed by the DUT and the OPAMP, it is more challenging to implement as the stability of the feedback signal has to be ensured.

The circuits which have been used in practice, however, incorporate a number of additional components. To switch between multiple amplification ranges, relays may be added which enable utilization of different resistor values. Additional passives may further be included to either improve noise figures by limiting the frequency range, or for circuit stability in case of the constant-current scheme. In addition, one might also want to measure the gate current to characterize SILCs.

Possible choices for control and measurement devices may be any combination of custom built voltage sources and sampling circuits, commercial source units or digital storage oscilloscopes. However, the control of the timing of the experiments and also achieving the highest SNR remains a formidable challenge for developing of measurement tools in general.

Measurement system used for this work

The measurement system which has been used by the author for the characterization methods outlined in this chapter is a custom designed, modular system, specifically developed for defect characterization. The reasons for using a custom system is in the requirements that come with TDDS and RTN measurements, which require long recording times at relatively high sampling rates, accurate synchronization between the outputs and inputs, as well as a high resolution of the recorded current signals. Part of the work of the author during his PhD studies was in developing this system and advancing its capabilities.

The current version of the system is based on a 19″ rack case with slots for 12 inserts. Inserts can be equipped from the front and connect to the case via back-panel connectors. Slots one and two of the system are occupied with a power supply unit for the analog and digital circuits. The back-panel supplies the remaining slots with the analog and digital power and individual universal serial bus (USB) connections which terminate at a USB hub in the back of the case to interface the units to a measurement PC for control and data transmission. In addition, the back-panel offers an inter-integrated circuit (I^2C) bus for communication between the units and signaling lines such as a sampling clock and a trigger line to synchronize the measurement outputs and inputs. In the minimal configuration for measurements of the channel current of a MOSFET, one of each of the following units has to be equipped as shown in Figure 4.3:

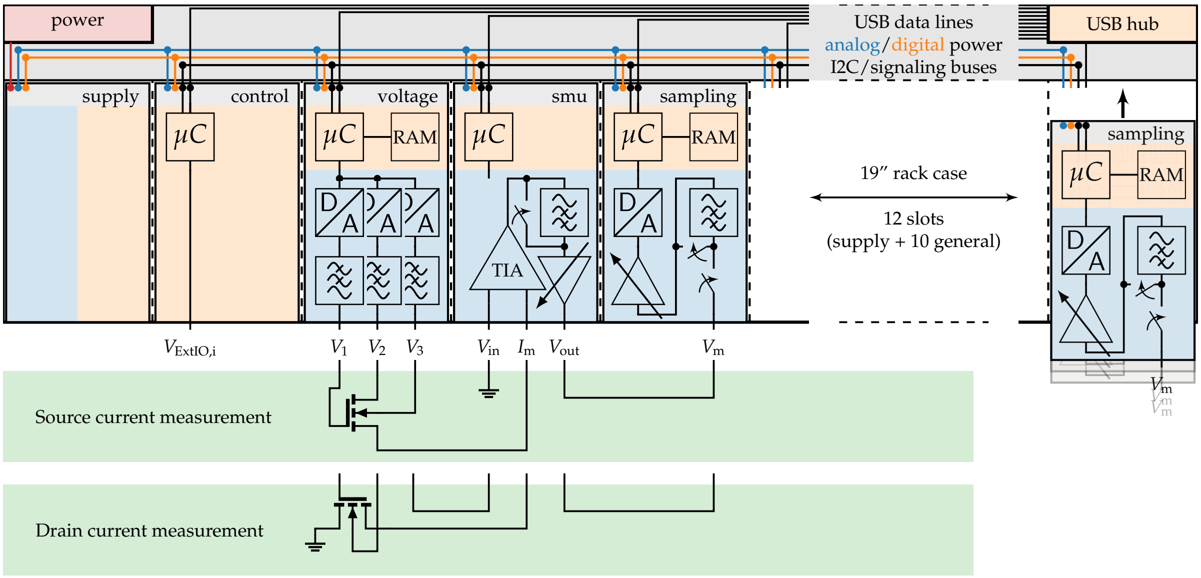

power
USB data lines
analog/digital power
I2C/signaling buses
USB hub
supply
control
voltage
smu
sampling
μC
RAM
D
A
TIA
19″ rack case
12 slots
(supply + 10 general)
sampling
$V_{\mathrm{ExtIO,i}}$
V_1
V_2
V_3
V_{in}
I_{m}
V_{out}
V_{m}
Source current measurement
Drain current measurement

Figure 4.3. Top: Schematic of the measurement system [39] employed for this work. Analog and digital subsystems colored in blue and orange, respectively; control signals omitted for clarity. The system consists of a 19″ rack case with 12 slots which can be equipped with a number of different inserts. The inserts connect to a back-panel which provides analog and digital power, a number of signaling lines, an I^2C bus and individual USB lines. For the methods based on channel conductivity as outlined in this section, the constant gate voltage scheme (Figure 4.1) has been used. **Bottom:** Using the system, the channel current can be measured either on the source terminal or on the drain terminal of the device as shown in the lower part of the figure. Additional inserts can be equipped to measure e.g. the gate current.

- **Control unit:** A purely digital insert which controls measurement clock, trigger and abort lines, and additionally supplies digital IO signals $V_{\mathrm{ExtIO},i}$ via a front-panel connector, e.g for the control of array chips (see Section 6.2) or other periphery.
- **Voltage unit:** Supplies three separate analog signals V_{1-3}. The signal output can be synced to the sampling unit(s) and switched at defined points in time between phases of the measurement, e.g. for BTI measurements. User defined ramps and arbitrary signal output is possible up to a frequency of 2 MHz, the signals are loaded via USB and stored in the connected RAM. Variants with different voltage ranges are available.
- **Source measure unit (SMU):** Contains a transimpedance amplifier stage as shown in Figure 4.1 which mirrors the voltage on terminal V_{in} on terminal I_{m} while at

the same time measuring the current on this terminal. The current is translated to a proportional voltage and optionally DC-filtered, post-amplified, or attenuated before being output on V_{out}. Multiple measurement ranges are available for selection.

- **Sampling unit:** An insert which samples a voltage at a frequency of up to 2 MHz with 18 bit resolution. It may be used in conjunction with the source-measure unit to measure a current or on its own to measure a voltage, e.g. for CV measurements. The analog input V_{m} can be DC-filtered and pre-amplified or attenuated before sampling. Sampled data can optionally be averaged using an averaging table to reduce the effective sampling rate and improve noise, or to achieve logarithmically increasing time steps. Finally they are stored in RAM before being transmitted to the measurement PC via USB.

Possible configurations for measuring the channel current of a MOSFET at its source or drain terminals are shown in the lower part of Figure 4.3. The measurements are controlled from the PC using a python framework which allows the user to configure and run measurements. After completion, the framework receives the acquired data, corrects them using calibration values stored in the respective units, calculates the current values from the sampled voltages using the values from the trans-impedance amplifier and writes the data in a file.

The flexibility this system offers allows to configure it for many characterization methods, including RTN, BTI, TDDS, on-the-fly (OTF), and current-voltage (IV) / hysteresis measurements as outlined in this section or CV, CP and DLTS measurements as shown in the following section. The gate current can be recorded alongside the drain or source current by adding a second set of SMU and sampling units.

4.1.1 Random Telegraph Noise (RTN) Measurements

RTN describes an effect which causes noise in the drain-source current of MOSFETs. It is caused by individual oxide defects stochastically charging and discharging during device operation. Noise in MOS transistors typically shows a 1/f spectral behavior, also known as pink noise or flicker noise. McWorther first proposed that this phenomenon originates from individual defects dynamically charging and discharging during device operation [86]. It has been further observed that the measured noise power scales roughly inversely proportional to the gate area of the devices [87]. As the feature size of MOSFETs decreased, Ralls et al. were able to measure discrete steps in the channel current, and thus show that the observed 1/f noise can be linked to individual defects which modulate the resistance of the inversion channel [88]. This effect paved the way for the observation of the behavior of individual defects when measured on small gate area devices.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.
TU Bibliothek WIEN Your knowledge hub

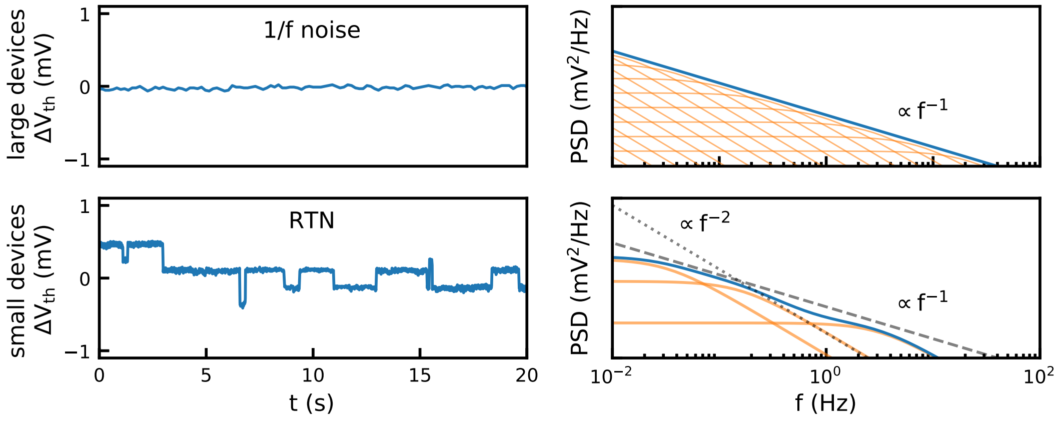


Figure 4.4. RTN measurements in the time domain **(left)** and frequency domain **(right)**. In large devices **(top)**, the noise signal in the time domain corresponds to a 1/f spectrum in the frequency domain. In small devices **(bottom)**, due to the small number of defects and the large influence of each defect on the channel current, individual steps can be observed. In the frequency domain, the Lorentzian PSDs of the individual defects which make up the 1/f shape in large devices might still be distinguished. Originally published in [BSJ1]

Measurement

The characterization of RTN appears—in terms of the measurement requirements—as one of the most straight-forward methods for defect characterization. The whole measurement procedure consists of applying defined bias voltages and recording the drain or source current. A common measurement scheme performed on small devices is to record the drain current over time at multiple gate voltages to obtain the bias dependence of the charge transition times of single defects. On large area devices, where only the superposition of the contribution of many individual defects can be studied, the main interest is the noise amplitude or power, which may also be recorded for a range of voltages. For the bias selection, typically a small drain-source voltage—usually on the order of ±100 mV for nMOS/pMOS devices—is chosen, sufficiently large to measure the channel current but low enough not to cause any hot carrier related effects. The gate bias is often varied in the sub-threshold to threshold region.

Typical measurement results are depicted in Figure 4.4. For large area devices, typically a signal containing 1/f noise is obtained, while for small area device, individual steps are visible. The spectra of each contribution of an individual defect can be described by a Lorentzian PSD, which shows a plateau below a certain frequency and decreases proportional to f^{-2} above that. For many defects with border frequencies spread equidistant over the logarithmic frequency axis, this adds up to a 1/f PSD, as can be observed for large area devices.

1/f noise

To show the relation between 1/f noise in large area devices and RTN in small devices, we start with the Lorentzian PSD of a two-state RTN signal, as shown in Figure 4.4. It can be calculated using the Wiener-Khintchine theorem from the autocorrelation function of the RTN signal [89, 90]:

$$S_{\mathrm{single}}(f) = \frac{(2d\tau_0)^2}{(\tau_{\mathrm{c}} + \tau_{\mathrm{e}})(1 + (2\pi f \tau_0)^2)}. \tag{4.1}$$

Here, d is the step height, and $1/\tau_0 = 1/\tau_{\mathrm{c}} + 1/\tau_{\mathrm{e}}$ is the mean frequency of the noise signal. For a number of defects, the spectral density is given by a superposition of the individual Lorentzian spectra:

$$S(f) = \sum_i \frac{(2d_i\tau_{0,i})^2}{(\tau_{\mathrm{c},i} + \tau_{\mathrm{e},i})(1 + (2\pi f \tau_{0,i})^2)}. \tag{4.2}$$

In large gate area transistors there is a large number of defects present with an apparently uniform distribution of $\log(\tau)$. Equation (4.2) can be simplified by condensing all defects with similar corner frequencies. For this, all defects within a small range $[\log(\tau_0) - \delta, \log(\tau_0) + \delta]$ are substituted with a defect with $\tau_{\mathrm{c}} = \tau_{\mathrm{e}} = \tau$ and an equivalent step height $d = d_{\mathrm{e}}$, which simplifies the expression for the PSD to:

$$S(f) \approx \frac{d_{\mathrm{e}}^2}{2} \sum_i \frac{\tau_i}{1 + (\pi f \tau_i)^2}. \tag{4.3}$$

In Equation (4.3), one term of the sum will be dominating at any frequency due to the characteristic shape of the Lorentzian PSDs. The maximum of the inner function is at $\tau_i = \tau_m = 1/(\pi f)$, thus the dominating term is the one with its corner frequency at πf. By considering only the dominant defects, Equation (4.3) simplifies to

$$S(f) \approx \frac{d_{\mathrm{e}}^2 \tau_m}{4} = \frac{d_e^2}{4\pi f}. \tag{4.4}$$

While this approximation is rather crude, it shows the link between RTN and 1/f noise. This also underlines the fact that characterization of RTN data is possible both in time- and frequency-domain. In this work, I will focus mostly on characterization of time domain data, as this enables a more direct observation of the behavior of individual defects.

Measurement window

The defects which may be characterized by RTN are those which are located energetically close to the Fermi level of the channel or the gate at the applied measurement conditions, as shown in Figure 4.5. The occupancy of such defects will be significantly

TU Bibliothek Your knowledge hub WIEN
Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

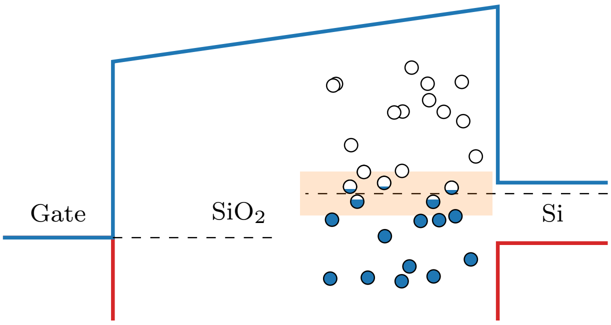


Figure 4.5. Energetical window for RTN characterization shown in a Si-SiO$_2$ band diagram. Defects located close to the Fermi level randomly exchange charge with the channel and produce RTN. Defects located far above (below) the Fermi level will be neutral (charged) for most of the time. The brief phases where they are charged (neutral) will most likely be too short to sample. Originally published in [BSJ1]

larger than zero and smaller than unity, which causes these defects to stochastically change their charge state within a reasonable measurement time window. By modifying the gate bias, the area in the band diagram which is scanned can be changed. The energy range which can be characterized depends on the maximum ratio between the transition times which can be extracted from the measured data and the temperature. From Fermi statistics it follows that

$$f(E_\mathrm{t}) = \frac{1}{\mathrm{e}^{\beta(E_\mathrm{t}-E_\mathrm{F})}+1} = \frac{\tau_\mathrm{e}}{\tau_\mathrm{e}+\tau_\mathrm{c}} \tag{4.5}$$

$$(E_\mathrm{t}-E_\mathrm{F})|_\mathrm{max} = k_\mathrm{B}T\ln\left(\left.\frac{\tau_\mathrm{c}}{\tau_\mathrm{e}}\right|_\mathrm{max}\right). \tag{4.6}$$

For example, at a given maximum ratio of measurable transition times $(\tau_\mathrm{c}/\tau_\mathrm{e})|_\mathrm{max} = (\tau_\mathrm{e}/\tau_\mathrm{c})|_\mathrm{max} = 100$ and $T = 300K$, $(E_\mathrm{t}-E_\mathrm{F}) \approx \pm 120\,\mathrm{mV}$, which means that in this case defects located in the range 120 mV above and below the Fermi level can be characterized.

Single Defect Measurements

Due to the stochastic nature of the capture and emission events, the actual characterization range for single defects is smaller than the previously estimated measurement window spanned by the sampling time and the measurement length. This is both due to the statistical variation of the dwelling times of a certain defect, as well as due to the requirement to measure multiple capture and emission cycles to accurately extract the average times. To predict the error due to a limited number of observations, the chi-square distribution can be used. It estimates the relative one-sigma confidence limits for the charge transition times [91]. For more than a handful of transition events observed, the width of the Gaussian distribution is close to the chi-square distribution and might be used instead to approximate the error as $1/\sqrt{N}$, with the number of observations N. From this, one can see that to achieve a reasonable error margin of approximately 10 %,

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.
TU Bibliothek WIEN Your knowledge hub

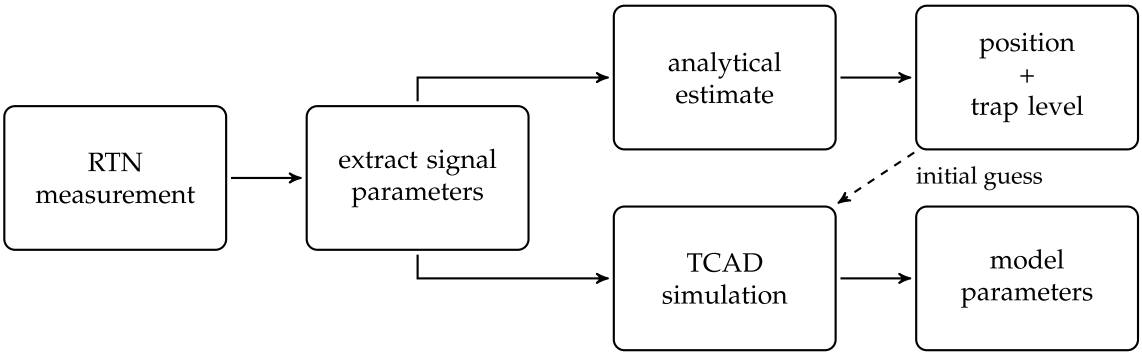


Figure 4.6. Flow chart illustrating parameter extraction from single defect RTN measurements. To obtain parameters suitable for defect simulation, the RTN signal parameters for each defect first have to be extracted from the measurements. Using these parameters, physical defect parameters such as their trap levels can be obtained using analytical estimates or TCAD simulation.

100 transition times have to be averaged. With only 10 observations, the error is still at approximately 30 %.

Finally, the measurement window is further reduced by the fact that the lowest dwelling times should be at least a few sampling times long. As the single defects are assumed memoryless, i.e. their statistical properties do not depend on their history, their transition rates are constant as long as the system is in equilibrium. This results in exponential distributions of their capture and emission times. The quantile function of the exponential distribution is $-\tau \ln(1-p)$, with the quantile p. This gives $\approx 0.01\tau$ and $\approx 4.6\tau$ for the 0.01 and 0.99 quantile, respectively. Thus, in practical terms the measurement window is nearly two orders of magnitude smaller than the scanning range at both the top and the bottom [92]. This requires a large sampling buffer length in excess of 100 kS to scan for defects. If only specific defects are to be analyzed, shorter sampling buffers might be possible.

Parameter Extraction The parameters of the RTN signals prevalent in the single defect measurements are the average charge capture and emission times, and the step height of each defect. A number of methods have been developed for extracting these parameters for the defects observed in such a measurement. They will be discussed in detail in Section 5.1. Once the characteristic parameters of an individual defect have been obtained at a number of measurement conditions, e.g. at different gate biases and device temperatures, they may be evaluated by comparison with TCAD simulation as outlined in Section 3.2. This finally yields the parameters of the employed defect model (see Section 3.1), such as the defect energy, position and curvatures of the parabolas. Alternatively to typically computational expensive TCAD simulations, an analytical method can be used to estimate the trap level and vertical position of a defect from the transition times, given in Section 5.1.4. However, the parameters of this method are often not very accurate. The results from this method may also be used as an initial guess for TCAD simulation. A flowchart illustrating the process is shown in Figure 4.6.

Gate Current Noise

So far the main focus has been put on noise in the drain-source current, but also the gate current can show noise similar to the one observed in the channel current. This may be aided by defects which assist tunneling of carriers between the gate and the channel. Recently, it was shown [93, 94] that there is indeed correlation between discrete gate and drain current noise. However, the reports available from literature are quite contradictory. While charge transitions are observed which lead to an increase of both the gate and the drain current, the opposite behavior is also observed. This opposite effect seems to stem from the inhibition of direct tunneling between the gate and the channel by charged oxide defects [95]. Additionally, SILC gives rise for an increase of the noise in the gate current as more defects get activated and created which can interact with the gate. Although the noise of the drain-source current is currently the primary subject of most RTN studies, the evaluation of the gate current is of equal importance and may provide missing information to the physical understanding of charge trapping in MOS devices.

4.1.2 Measure-Stress-Measure (MSM) Methods

A common way to learn about the reliability of a technology is to subject devices to stress conditions fairly exceeding the nominal operating range, in order to accelerate degradation mechanisms and allow for observations within a reasonable time window. This can include elevated gate and/or drain biases, device temperature or exposure to light or ionizing radiation. The basic measure stress measure (MSM) scheme consists—as the name suggests—of three phases. First, a measurement of a virgin device is taken. This is followed by the application of defined stress conditions, and finally another measurement. The original measurement can then be compared to the post stress measurement to evaluate the effect of stress. The stress and measurement part may be repeated multiple times, commonly with harshening stress conditions. An illustration of this scheme is given in Figure 4.7. In the example, the device is subjected to gate bias stress, and in the measurement phases $I_\mathrm{D}(V_\mathrm{G})$ curves are recorded. Apart form IV measurements, this basic scheme is also used with e.g. CV and CP measurements.

IV Measurements

The transfer characteristics $I_\mathrm{D}(V_\mathrm{G})$ is one of the most important characteristics of the MOSFET. By comparing $I_\mathrm{D}(V_\mathrm{G})$ characteristics of devices before and after stress, the threshold voltage shift, and changes of the sub-threshold slope, the mobility and the on and off current can be obtained, as illustrated in Figure 4.8.

4.1.3 Extended Measure-Stress-Measure (eMSM)

In addition to the information that can be obtained by comparing a device before and after stress, further information can be gained by observing the behavior of the

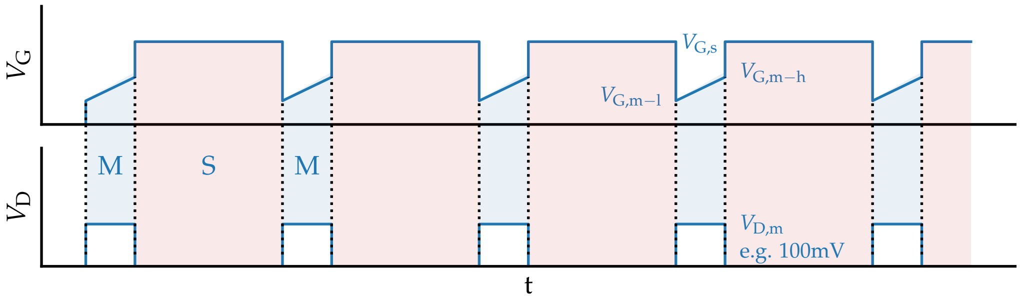


Figure 4.7. Gate and drain biases during a MSM measurement. After an initial measurement, phases of stress and subsequent measurements are performed. This allows to characterize the degradation due to the applied stress conditions. In this example, degradation due to gate bias stress is characterized using $I_D(V_G)$ measurements.

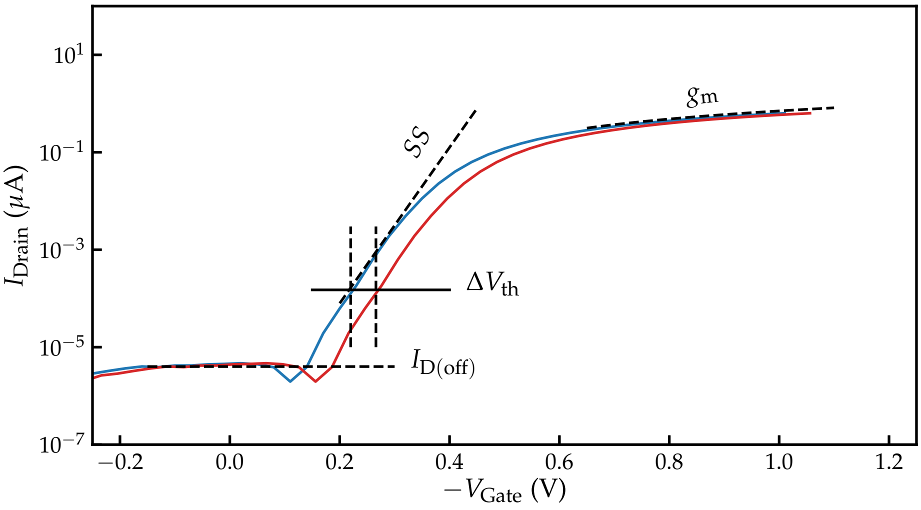


Figure 4.8. Illustration of $I_D(V_G)$ characteristics of a device before and after BTI stress. From this, the threshold voltage shift ΔV_{th}, changes in the sub-threshold slope, the off-current and the transconductance can be obtained. Originally published in [BSJ1]

device after stress release. In the extended MSM (eMSM) scheme [96], phases of stress are alternated with measurement phases at recovery conditions as shown in Figure 4.9. This is of particular interest if the effect of stress is—at least in part—reversible, as is the case for BTI stress.

BTI and HC Measurements on Large Devices

In large devices, eMSM measurments are often used to characterize the effect of BTI or HC stress. To measure BTI stress using the eMSM scheme, the drain and gate terminals of the device are kept at ground while a gate bias at or above the threshold voltage is applied during the stress phase. During the relaxation or recovery phase, the gate bias is set to a lower readout value while a small drain bias—commonly 50 mV or 100 mV for Si technology—is applied in order to ensure a small channel current. By using an initial $I_D(V_G)$ measurement recorded before stress, the temporal behavior of the channel current can be mapped to a threshold voltage shift during

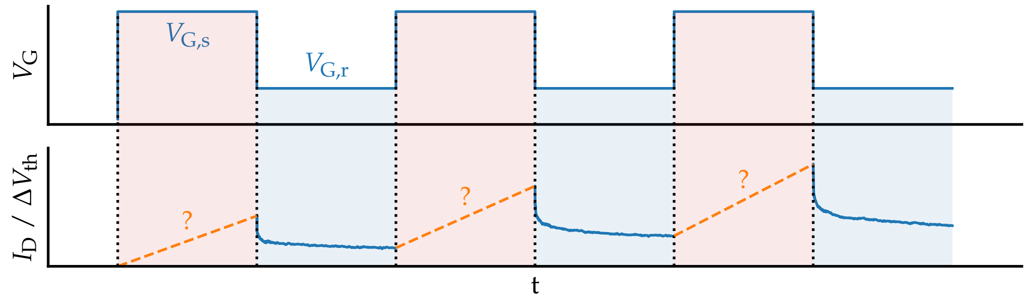


Figure 4.9. Gate bias and drain current during an eMSM measurement. Phases of stress alternate with phases of relaxation. During relaxation the drain current or threshold voltage shift is recorded. This allows to observe the impact of stress on the device behavior and also to understand the following relaxation behavior.

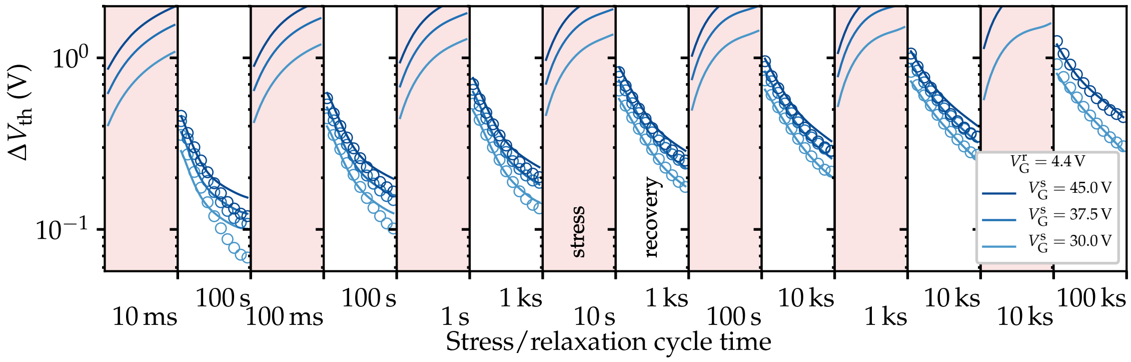


Figure 4.10. Measurement data (**circles**) and simulation data (**lines**) for a BTI eMSM experiment. Note that the simulations suggest that a significant part of the threshold voltage shift due to stress already recovered before the first recovery measurement point is taken, underlining the need for fast measurements.

relaxation. After the recovery phase, another stress phase may be performed and the recovery behavior monitored afterwards. Subsequent stress phases are often chosen with increasing stress times, as shown in Figure 4.10. The corresponding recovery phases are usually chosen long enough that most of the recoverable part of the degradation vanishes. As the recovery times used for MSM sequences are very long compared to noise measurements, sampling is usually performed at logarithmic time instances for this type of measurement in order to keep the amount of samples feasible.

Hot carrier stress can be characterized in a similar manner, but with different bias conditions applied during stress. For HC stress, the gate bias applied is commonly chosen lower than for BTI but a significant drain bias, often larger than the gate bias, is applied. The drain bias which causes the most severe degradation depends on the geometry of the device and ranges from $V_D = V_G$ to $2 \times V_G$. The drain-source voltage causes inversion layer charges to gain energy by accelerating towards the gate. The hot carriers may then either damage the interface directly or generate other hot carriers as shown in Figure 4.11.

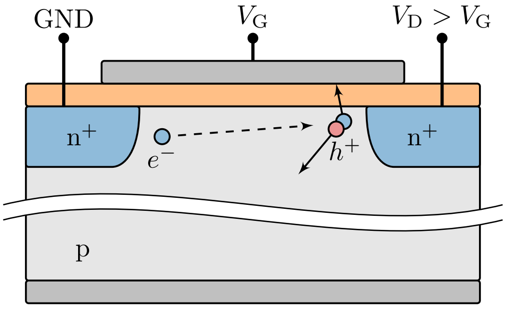


Figure 4.11. Hot carrier stress. Unlike BTI stress, a non-negligible source-drain bias is applied, causing carriers to accelerate towards the drain. This causes damage at the interface or secondary carrier generation.

Time Dependent Defect Spectroscopy

Another application which makes use of eMSM measurements is TDDS. More precisely, the TDDS aims at the characterization of single defects. On small gate area devices, as described in Section 4.1.1, charging and discharging of such defects can be observed within the relaxation traces. This can be used in conjunction with BTI or HC measurements to characterize their charge trapping kinetics at various biases and temperatures in order to develop a physical explanation for their behavior.

Measurement The measurement conditions for TDDS are the same as those used for BTI or HC measurements on large devices. The difference is that each stress condition is measured multiple times to gain sufficient statistics on the capture probabilities and emission times of the single defects, as outlined for RTN single defect measurements (Section 4.1.1). The characterization windows for TDDS measurements are determined by the charge capture time during stress and the charge emission time during relaxation of the respective defects. For the stress phase the measurement window is defined by:

- The minimum stress time which can be consistently applied, which is given by the measurement equipment
- The maximum stress time applied
- The highest gate bias at which the oxide field does not cause oxide breakdown
- The lowest gate bias at which the defect under investigation has an occupancy close to one.

For the recovery phase, the measurement window is given by:

- The delay after stress at which the first point can be measured, which is given by the measurement equipment
- The maximum relaxation time applied
- The highest gate bias at which the defect still has an occupancy close to zero.

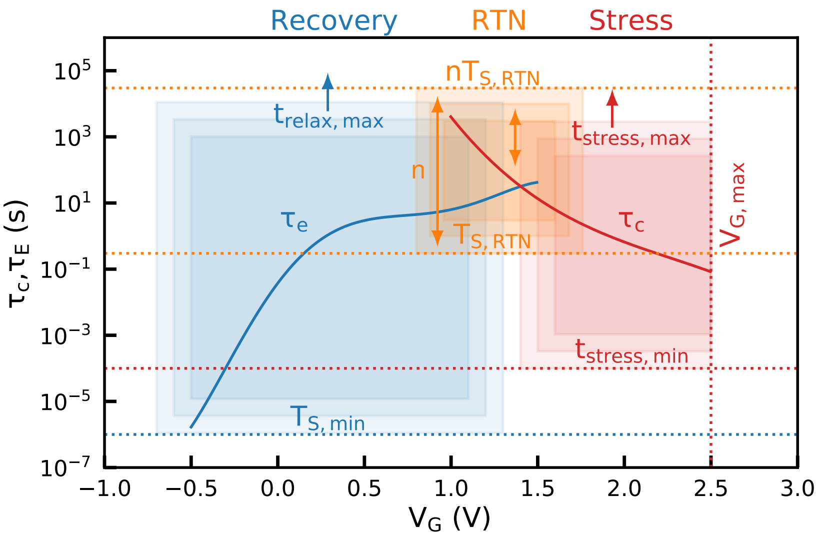


Figure 4.12. Measurement windows for RTN and TDDS single defect characterization, shown for an exemplary defect. While RTN allows to characterize a defect at gate voltages close to $E_F = E_t$, TDDS allows to measure its charge capture time at biases above this point and its charge emission time at biases below that point. By characterizing the defect using both methods, a comprehensive picture of the defects' behavior can be obtained. Originally published in [BSJ1]

- The minimum voltage at which single recovery steps can be measured.

In general, applying the TDDS to extract the charge transitions times over wide ranges for stress and recovery bias is very time consuming. However, TDDS measurements can be performed in conjunction with RTN measurements, which allows to extract the capture and emission times of a defect over wide ranges of gate biases while also capturing both close to the intersection bias, as shown in Figure 4.12.

Parameter Extraction TDDS data are commonly processed using step detection algorithms, similar to RTN data. Compared to RTN data, defect parameter extraction is comparatively simple, as the emissions probabilities for the defects are not constant over the time of the experiment, but increase drastically after applying the recovery conditions. This allows to directly obtain the emission times from the relaxation phases. Defect parameter extraction for TDDS data will be discussed in Section 5.1.

4.1.4 On-the-Fly (OTF) Measurement

The measurement methods discussed so far aimed at analyzing the behavior of the devices after a set amount of stress. In case the primary objective of characterization is the evolution of the device during stress, the OTF [97, 98] scheme may be used. For this, phases of stress are interrupted by short gate pulses to sample single points on the $I_D(V_G)$, with the goal to cause minimal disruption of the applied stress. This is illustrated in Figure 4.13. Its main drawback as compared to the MSM method is that

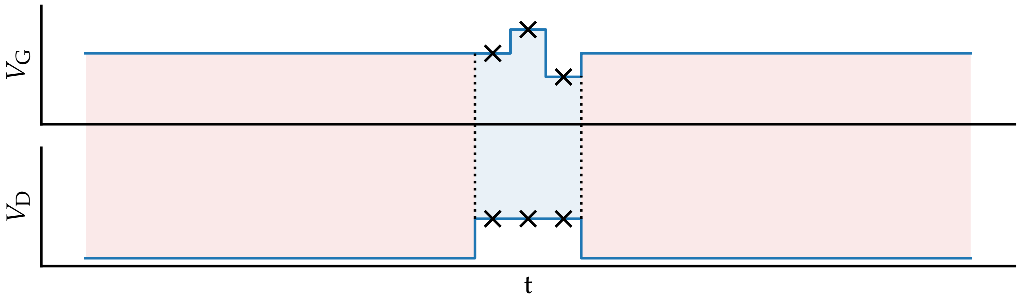


Figure 4.13. Gate and drain biases during an on-the-fly measurement. A small drain bias is applied for the duration of the measurement. At the same time, the gate bias is pulsed to measure selected points on the $I_D(V_G)$ curve. This keeps the interruption to stress at a minimum and thus allows to observe the degradation of the MOSFET.

only a very small part of the $I_D(V_G)$ curve is sampled, which complicates interpretation of the results.

4.1.5 Hysteresis Measurements

A measurement technique which can be used to learn about defects in devices of certain technologies is to record the hysteresis between voltage sweeps. For these devices, the gate voltage sweep during the recording of an $I_D(V_G)$ can subject them to a sufficient amount of stress to significantly alter its shape. In this case, MSM schemes may not be possible due to the failure to obtain a proper reference curve. To characterize these devices, repeated up- and down-sweeps can be recorded to obtain hysteresis curves as shown in Figure 4.14. These curves may be recorded at different sweep rates to obtain information about the numbers and time constants of the defects affecting the device. This method has been applied in the past to novel technologies for which a considerable amount of defects is still prevalent. For instance, devices based on graphene or SiC exhibit a significant hysteresis, while Si devices show almost no hysteresis behavior. Quite interestingly, while large area devices typically exhibit a continuous hysteresis behavior, single charge transition events become evident in voltage sweeps measured at nanoscale devices [70, BSJ7].

4.2 Electrical Methods based on Defect Charge

Another way of characterizing defects in MOS structures is to measure the charge they emit or capture during (dis-)charging. This can be done for example by pulsing the gate bias, applying an AC voltage to the gate terminal, or in the case of TDRC by cooling the device, changing the bias, and then heating up the device. To obtain a sufficient response, these methods usually require larger devices than the methods based on channel conductivity, as well as a sufficient defect density.

Schematics for a measurement circuit as might be used for these measurements is shown in Figure 4.15. The gate bias at the DUT is supplied by a DAC and recorded using

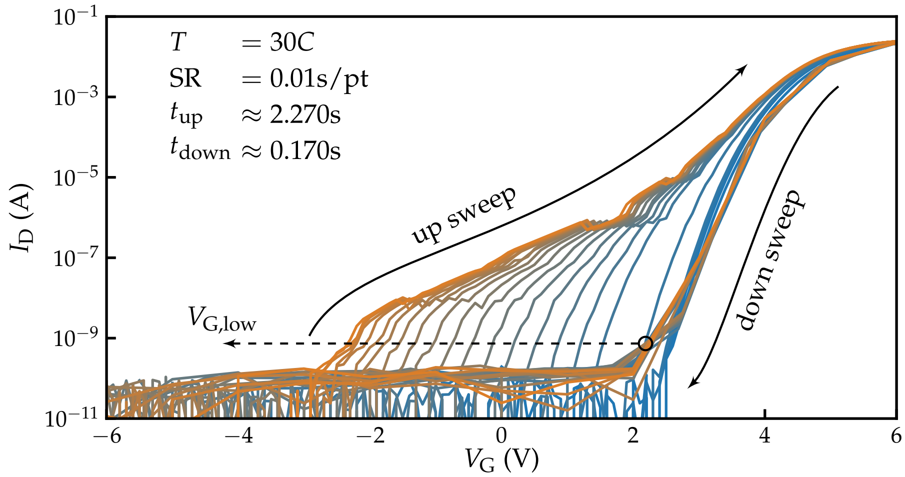


Figure 4.14. Hysteresis measurements on a 4H-SiC nMOSFET with varied starting voltage $V_{G,low}$ and constant high voltage $V_{G,high}$. Decreasing $V_{G,low}$ increases the hysteresis width. Data originally published in [99]

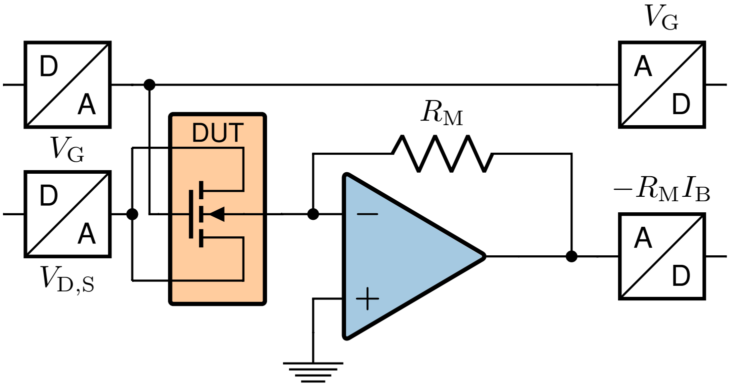


Figure 4.15. Schematics of a circuit as might be used for the charge based methods. With this setup, all signal processing has to be done in software. Particularly for CV measurements, part of the impedance extraction might be performed in hardware. In CP measurements, a hardware integrator might be used.

an ADC, often with separate wires from the DUT to mitigate parasitic effects introduced by the measurement lines. The bulk current is converted to an equivalent voltage using an OPAMP for trans-impedance conversion and is measured using a second ADC. The source and/or drain biases might be either supplied by a second DAC, shorted to ground or connected to bulk, depending on the measurement performed. Note that while a circuit as shown here allows a number of measurements to be performed, high-speed sampling is necessary and all processing has to be performed in software. If only a certain type of measurement is to be performed, parts of complexity might be shifted to the analog part of the circuit, e.g. signal mixing in the case of CV measurements and signal integration for CV, CP and DLTS measurements.

Measurement system used for this work

For the measurement methods as outlined in this section the same system as shown in the previous section (Section 4.1) has been used. Possible configuration for measurements of the gate and bulk current are shown in Figure 4.16. Note that with the configurations shown the current is never measured on a terminal where an AC signal is applied so as not to measure spurious currents resulting from stray capacitances.

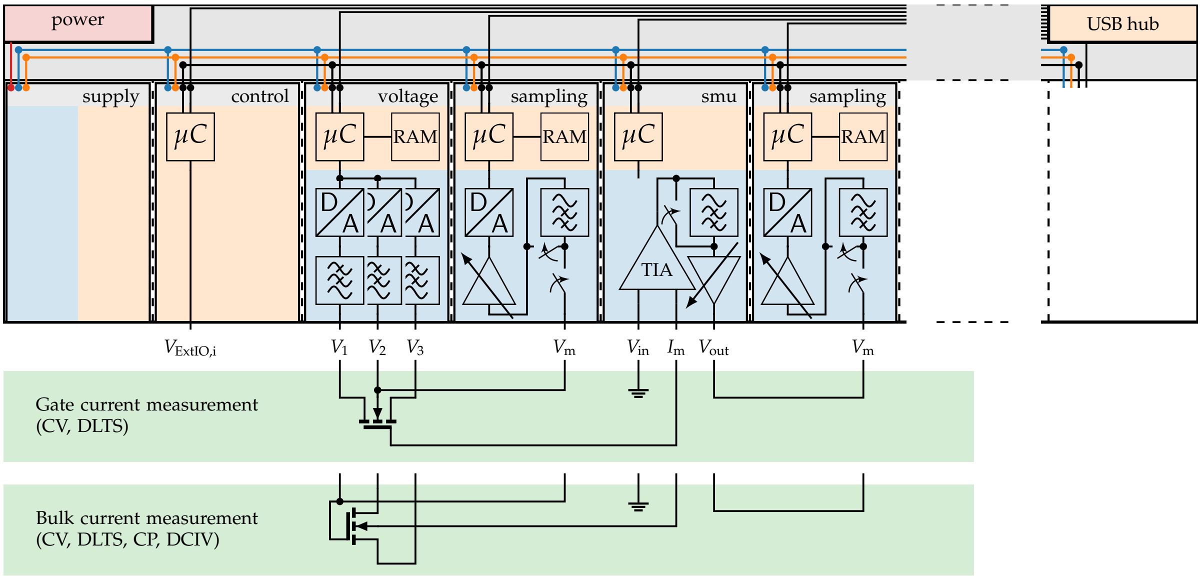


Figure 4.16. Top: Schematic of the measurement system [39] employed in this work, equipped for the measurement methods presented in this section. Analog and digital subsystems colored in blue and orange, respectively; control signals omitted for clarity. As compared to Figure 4.3 for channel current measurements, an additional sampling unit included to measure the phase of the input signal for CV and DLTS measurements. **Bottom:** Possible configurations for gate and bulk current measurements. For CV measurements both the gate or the bulk current can be used to measure the MOS capacitance. Especially when testing devices on a wafer this decision may however influence the noise level of the measurement. Source and drain terminals—if available on the test structure—may either be shorted to bulk or supplied separately (as shown), in case the gate current is measured optionally with or without the AC signal. Depending on the chosen configuration the gate-source and gate-drain capacitances may be included in the measurement.

As compared to the previous section, a second sampling unit is shown here, which is used for CV and DLTS measurements to sample the phase of the AC part of the applied signal. To implement these measurements, the hardware and firmware of the units, which had originally been developed for channel current measurements, had to be modified. Most importantly, the sampling units have been redesigned to include a faster ADC, DC filters and (secondary) amplifiers, while the firmware of the voltage unit has been modified to allow for arbitrary signal output, and a synchronization clock line on the back-panel has been added to eliminate phase drift between the units due to the slightly different processor clocks. Additional functions have been implemented in the python software framework which allow to process the captured AC waveforms to extract the parallel capacitances and conductances either per voltage (for CV) or over time (for DLTS).

4.2.1 Capacitance-Voltage (CV)

In CV measurements, the small signal capacitance $\mathrm{d}Q/\,\mathrm{d}V$ is recorded. It yields information on the device itself, as well as on defects contained. This can be achieved either by measuring the displacement current while sweeping the gate voltage (ramp method) or by superimposing a small AC signal on top of the DC gate bias and measuring the AC current (impedance method). The following discussion will focus on the impedance method, as it allows for greater flexibility in measurement parameters and further enables observation of the small signal conductance and extraction methods based on this.

Measurement

The principle of a CV measurement is illustrated in Figure 4.17. A small sinusoidal AC voltage—typically 50 mV at 100 kHz—is added to a DC bias which is slowly varied. This voltage is applied to the gate, while the bulk is forced to ground. Both, the gate voltage and the bulk current are measured simultaneously. Alternatively, the roles of gate and bulk may be switched, i.e. the voltage is applied to bulk and the gate current is measured. As stated before, this measurement can be performed either on MOSCAP structures or on MOSFET structures. The source and drain contacts are not necessarily required for the measurement. If they are available, they are usually shorted to ground or connected to bulk if the source and drain currents are to be recorded. The most fundamental difference of MOSCAP and MOSFET structures is that for MOSFETs the minorities required for inversion can be quickly supplied from the source and drain regions. This allows for faster gate bias sweeps compared to MOSCAP structures, where they have to be supplied by generation. The disadvantages of MOSFETs are the additional parasitics due to overlap of the source and drain areas with the gate. MOSCAP structures have the additional advantage of being effectively one-dimensional, allowing simple and accurate models to be used for description.

When CV measurements are performed, the gate bias is usually stepped from accumulation to inversion or vice versa to obtain the C-V and G-V curves, i.e. the small signal capacitance and conductance curves. The most important parameter for this measurement is the frequency of the applied AC signal, which may vary from kHz to GHz, but is usually limited to 1 MHz due to the significant increase in measurement effort beyond that point. The AC frequency determines how defects contribute to the measured CV curve as defects which can not follow the signal will not contribute to the measured capacitance. In MOSCAP structures, where minorities are not readily available from source or drain, the sweep rate in comparison to the minority response further determines the type of result obtained from the measurement.

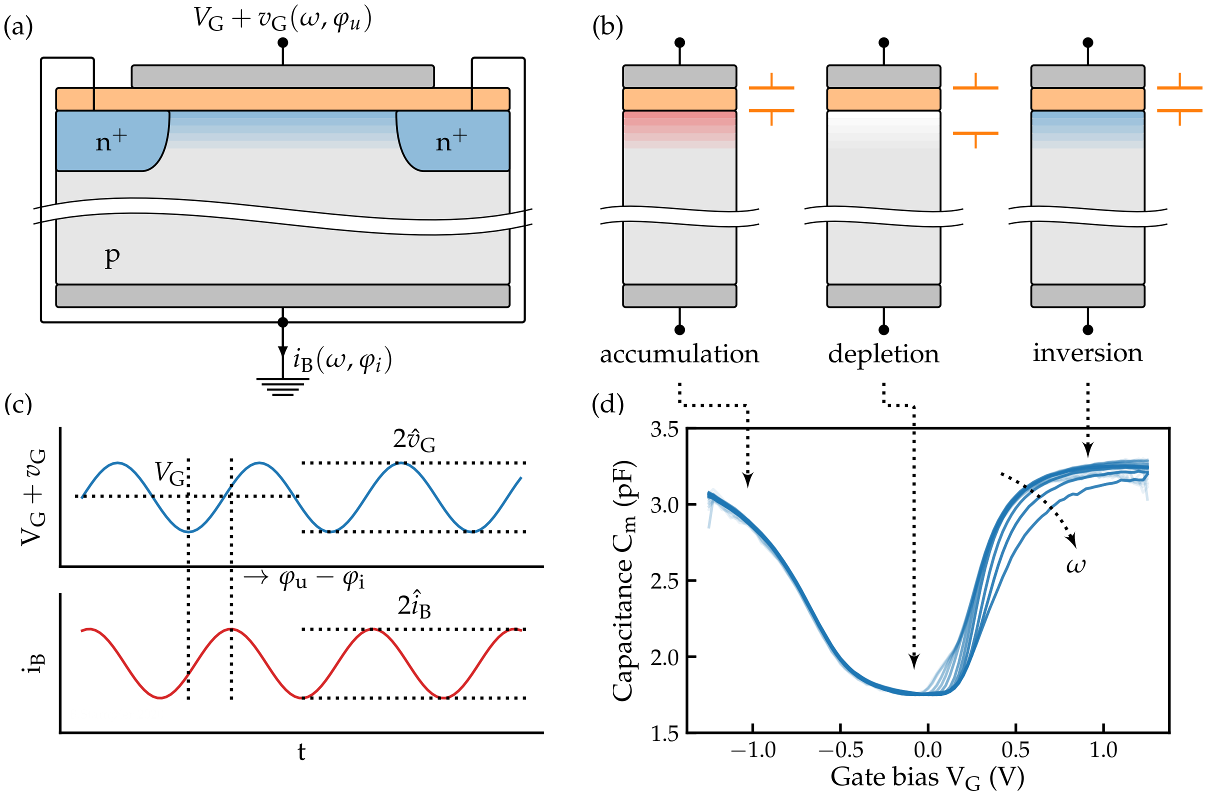

(a)
$V_G + v_G(\omega, \varphi_u)$
n+
n+
p
$i_B(\omega, \varphi_i)$
(b)
accumulation
depletion
inversion
(c)
$V_G + v_G$
V_G
$2\hat{v}_G$
$\rightarrow \varphi_u - \varphi_i$
i_B
$2\hat{i}_B$
t
(d)
Capacitance C_m (pF)
3.5
3.0
2.5
2.0
1.5
−1.0
−0.5
0.0
0.5
1.0
Gate bias V_G (V)
ω

Figure 4.17. Capacitance-voltage measurement. **(a)** DUT connected for measurement. **(b)** DUT in the various operation schemes with illustrated equivalent capacitance. In the depletion regime there are no free charges close to the interface. This can be interpreted as an increase in spacing between the plates of the equivalent capacitor. This in turn lowers the capacitance and gives the CV curve its characteristic shape. **(c)** Voltage and current signals during measurement. The phase and amplitudes of the voltage and current signals give the sought-after impedance. The offset voltage is swept in a staircase-like manner during the measurement. **(d)** Exemplary results of a CV measurement at multiple frequencies. The accumulation and inversion branches almost reach the oxide capacitance, while the capacitance in depletion is much lower. The inversion branch is generally steeper than the accumulation branch.

Impedance Extraction

The small signal impedance Z is given by the division of the complex values of the AC gate voltage v_G and bulk current i_B:

$$Z = \frac{v_G}{i_B} \tag{4.7}$$

$$|Z| = \frac{|v_G|}{|i_B|} \tag{4.8}$$

$$\arg(Z) = \arg(v_G) - \arg(i_B) \tag{4.9}$$

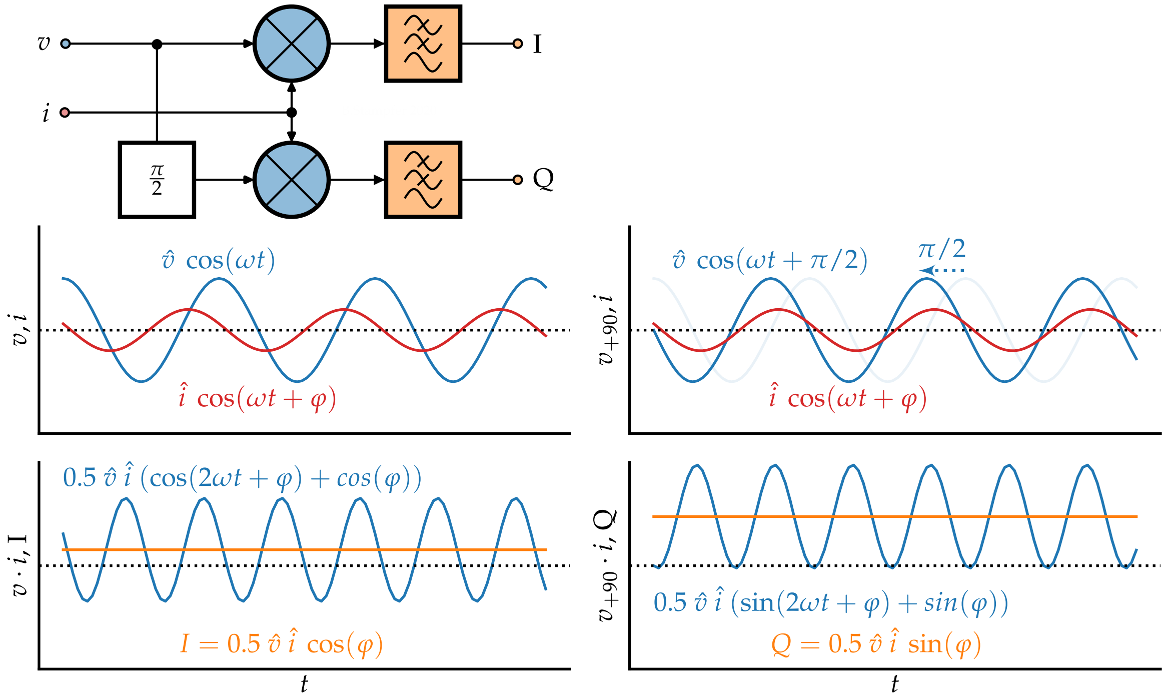


Figure 4.18. Extraction of the phase difference from a CV measurement. The current signal is mixed with the original voltage signal **(left)** and a phase-shifted signal **(right)**. This gives signals with their sum and difference frequencies, i.e. $2\omega_0$ and $0\omega_0$. After low-pass filtering or integrating, the DC I and Q signals which are proportional to the imaginary and real parts of the current remain.

The amplitudes and phase difference necessary for the calculation can be obtained from the sinusoidal signals either using analog circuitry or in the digital domain. In either case, a common approach to extract the phase shift is to employ I-Q demodulation as shown in Figure 4.18. For this, the current signal is mixed with the voltage signal and successively low pass filtered or integrated to obtain an in-phase signal I which is proportional to the conductance—the real part of the admittance. By mixing the current signal to the $\pi/2$ phase-shifted voltage signal, the quadrature-phase signal Q is obtained, which is proportional to the susceptance—the imaginary part of the admittance. From these signals the relative phase shift can be calculated as $\varphi = \tan^{-1}(Q/I)$. Together with the measured amplitudes, the impedance of the device can then be calculated. It should be noted that commercial impedance analyzers or CV meters often give the impedance as values of an effective parallel or series R-C circuit, i.e.:

$$Z = R_\mathrm{s} + \frac{1}{j\omega C_\mathrm{s}}, \qquad \text{for the series R-C circuit} \tag{4.10}$$

$$1/Z = 1/R_\mathrm{p} + j\omega C_\mathrm{p}, \qquad \text{for the parallel R-C circuit.} \tag{4.11}$$

Since in general neither of these equivalent circuits will correctly represent the situation, it is preferable to obtain the complex impedance and then calculate the required quantities using the equivalent circuit most suitable for the combination of experimental setup, operating condition and device.

Defect Parameter Extraction

From the measured impedances, a variety of device parameters, including the oxide thickness, doping concentration, flat band and threshold voltage can be extracted. To find the contribution of the defects to a measurement, it has to be compared to some reference. There are several options on what to use as a reference.

One option is the high-low frequency method [100], where a low frequency measurement is compared to a high frequency measurement. The difficulty of this method comes from performing measurements at frequencies high enough for most defects to be inactive, which can require frequenices of up to 100 MHz. Another pitfall of this method can be additional effects unrelated to defects, which cause the high frequency curves to look different from low frequency curves. Another option is to measure at a single frequency and compare the measured curve to one obtained from simulation or calculation [101, 102]. While this seems easy from the measurement perspective, obtaining theoretical results which accurately represent the defect free device is difficult and requires exact knowledge of the device properties. Finally, measurements may be done at a single frequency in a MSM-like manner. The difference between subsequent measurements can then be linked to changes in the defect populations of the device. A more detailed discussion on how to obtain defect parameters from CV measurements can be found in the following chapter.

4.2.2 Charge Pumping (CP)

Charge pumping is a characterization technique first performed by Brugler and Jespers in 1969 [103] and later treated from a theoretical perspective by Groeseneken in 1984 [104]. The technique is most sensitive to fast defects located at or close to the interface. Unlike CV measurements, where defects are charged using a small sinusoidal AC voltage, CP uses a large rectangular AC voltage signal. The energetic distribution of defects can be probed by modifying the properties of the AC signal.

For the measurement, the gate terminal of the MOSFET is connected to the signal source while the bulk is forced to ground and the resulting bulk current is measured or integrated. The source and drain terminals are shorted to ground or slightly reverse biased to allow injection of minority carriers during inversion. This is illustrated in Figure 4.19a. The gate of the device is then repeatedly pulsed between a high and a low voltage, $V_{\mathrm{G,h}}$ and $V_{\mathrm{G,l}}$ respectively, as shown in Figure 4.19b. The frequency used for CP measurements is commonly chosen in the 10 to 1000 kHz range, the steepness of the edges of the signal is critical to reproduce recorded measurements. During the high bias phase, the defects charge with minorities originating from the source/drain regions. In the following low bias phase, the defects then discharge and the released minorities recombine with majorities in the bulk, which have to be replenished from the bulk contact. These charges are measured as bulk current, and their amount depends on the number of defects which have been charged and discharged under the chosen measurement conditions.

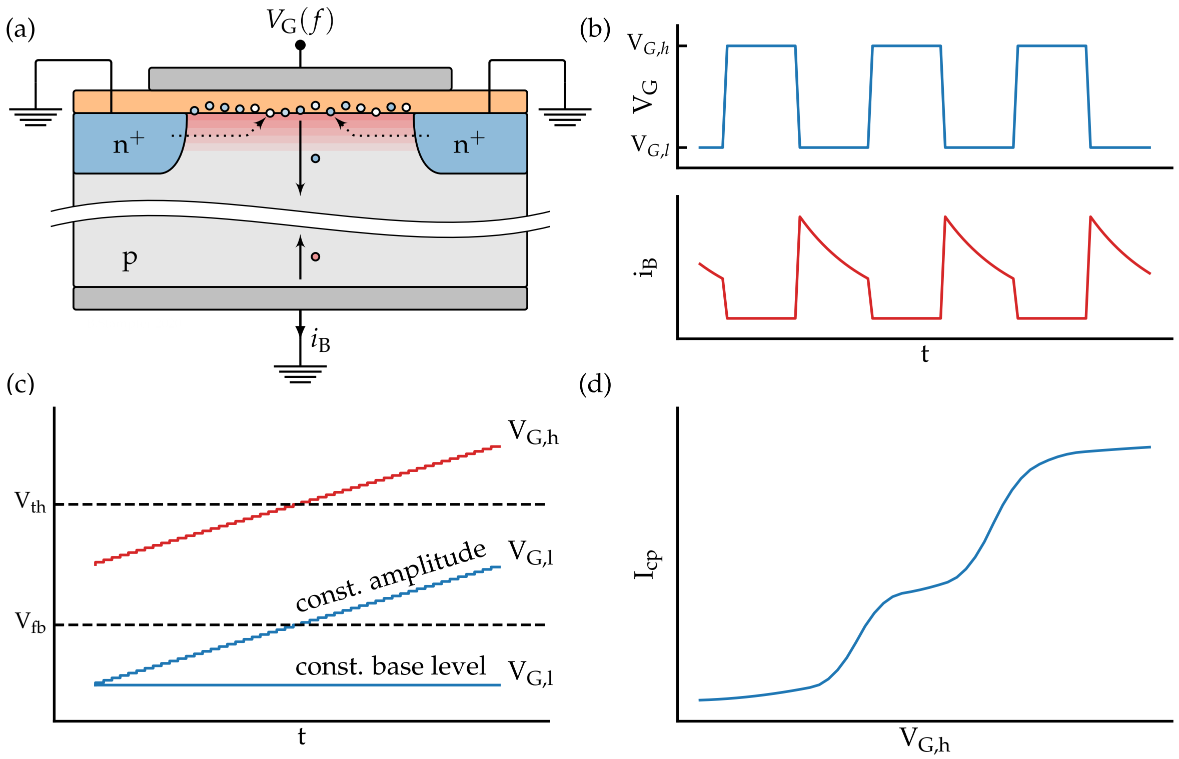


Figure 4.19. The CP technique. **(a)** DUT during measurement. During the high bias phase defects charge with minorities supplied from source and drain. In the following low bias phase, the minorities are emitted and recombine with majorities sourced from the bulk. Their amount is measured as the bulk current. **(b)** Gate voltage and bulk current during measurement. **(c)** Gate voltage parameters for the two measurement schemes. **(d)** Illustration of a measurement result.

To probe the energetic distribution of the defects, three basic measurement schemes are available. In the constant-amplitude scheme, the amplitude of the AC signal is held constant and its offset is swept, while in the constant-base-level scheme, $V_{\mathrm{G,l}}$ is kept constant while $V_{\mathrm{G,h}}$ is varied, as shown in Figure 4.19c. Likewise, in the constant-high-level scheme $V_{\mathrm{G,h}}$ is kept constant while $V_{\mathrm{G,l}}$ is swept. The total sweep range is commonly chosen from strong accumulation to inversion.

The resulting CP current looks as depicted in Figure 4.19d. The current depends directly on the interface trap density, the frequency, and the gate area. The defect densities can then be calculated from measurements with varying t_{r} and t_{f} using [104, 105]

$$I_{\mathrm{cp}} = 2qkTf\overline{D_{\mathrm{it}}}A_{\mathrm{G}}\ln\left(v_{\mathrm{th}}n_{\mathrm{i}}\sqrt{\sigma_{\mathrm{n}}\sigma_{\mathrm{p}}}\sqrt{t_{\mathrm{r}}t_{\mathrm{f}}}\,\frac{|V_{\mathrm{fb}}-V_{\mathrm{th}}|}{|V_{\mathrm{G,h}}-V_{\mathrm{G,l}}|}\right). \tag{4.12}$$

Here, I_{cp} is the charge pumping current, A_{G} the gate area, $\sigma_{\mathrm{n,p}}$ the capture cross sections, n_{i} the intrinsic carrier concentration, and $t_{\mathrm{r,f}}$ the rise and fall times of the trapezoidal waveform.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN

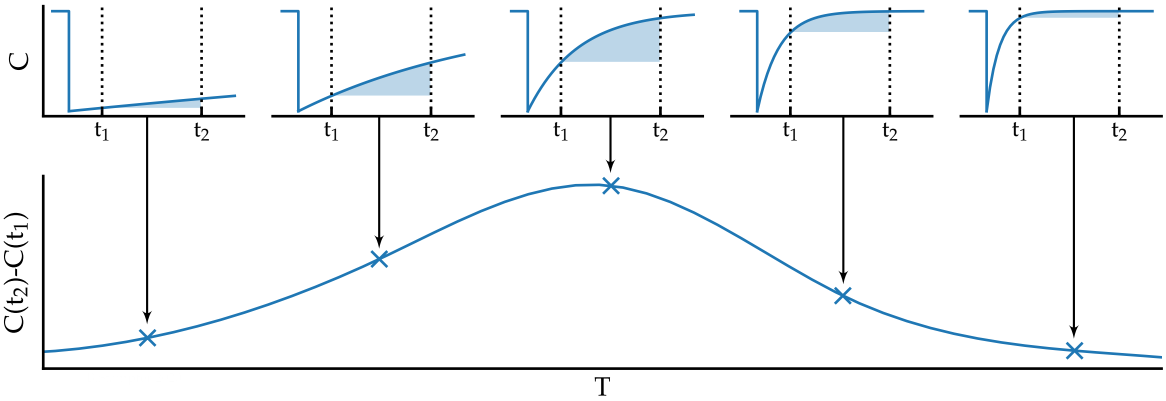


Figure 4.20. Illustration of deep-level transient spectroscopy using rate windows. As the temperature of the sample is increased, the time constants of defects in the device decrease. Once the response time of a defect band coincides with the chosen rate window given by $t_{1,2}$, the difference between the capacitance at these points—the DLTS signal—increases. At even higher temperatures the defects responses get too fast and the signal decreases again.

4.2.3 Deep-Level Transient Spectroscopy (DLTS)

DLTS is a transient capacitance thermal scanning technique introduced by Lang in 1974 [106]. It was first used to characterize bulk defects in the depletion regions of p-n junctions. For this method, the device is biased negatively to form a depletion region below the junction. This bias is repeatedly interrupted by pulses of more positive bias, which cause non-equilibrium conditions in the device. After the bias has returned to its initial value, the defects charged during the pulse return to their equilibrium state. The transient change of the capacitance during this time is evaluated, in the original version using a double boxcar integrator, which measures the change in capacitance during a given time frame after the pulse. In addition, the temperature of the sample is slowly swept, and each defect band in the device will exhibit a maximum of the measured change in capacitance at a specific temperature, at which the time constants of the defects are in the range of the rate window, as shown in Figure 4.20 for one such peak. By changing the rate window, these temperatures slightly change. This allows to draw an Arrhenius plot for each peak and subsequent extraction of the corresponding defect energy. The method is able to distinguish between majority and minority carrier emission by choosing the level of the bias pulse in the reverse or forward regime.

With regard to MOS devices, interface trapped charge DLTS [107] was developed to characterize interface defects. In this variant, an accumulation pulse causes traps to be filled by majorities. After the pulse, the device is switched to deep depletion. This stimulates the traps to emit, which can be observed as a capacitance transient. While this is possible on MOSCAPs, the source/drain terminals on MOSFETs allow control of the minorities which extends the applicability of DLTS. With the source/drain diodes reverse biased, no minorities are injected from this regions which can interfere with the majorities. This allows characterization of the majority half of the band-gap. With the source/drain diodes forward biased, an inversion layer forms. This allows the characterization of the minority carriers in the other half of the band-gap.

A number of variants of DLTS have been developed over the years, including constant capacitance DLTS, lock-in amplifier DLTS, correlation DLTS, isothermal DLTS, computer-, Laplace-DLTS, and optical and scanning DLTS [105]. All the variants are targeted for the characterization of large structures, as a sufficient number of trapped defects are required to cause a recombination current of measurable size.

4.2.4 Direct-Current IV (DCIV)

DCIV is a relatively simple method to characterize the number of charged interface or near interface defects. The technique was first demonstrated by Neugroschel et al. [108]. For this measurement technique, the gate bias is swept from accumulation to inversion while the bulk current is recorded. At least one of the source/drain contacts has to be slightly forward biased to allow for minority carrier injection, as shown in Figure 4.21. During the sweep, minority carriers will be injected into the space charge region under the gate oxide. These minorities then combine at the interface with defect charges, causing majorities to drift towards the space charge region. To fulfill charge neutrality over the gate stack, charges have to be supplied by the bulk, which can be measured as a bulk current.

Note the dissimilarities between DCIV and CP. In CP, charges are not actively injected using source/drain biases as in DCIV but are made available by channel inversion. These minorities can then charge defects close to the interface, which are later emitted during accumulation to produce the CP current. In DCIV, the injected charges recombine with previously charged defects, causing the bulk current. Compared to CP, the gate bias range for DCIV is smaller.

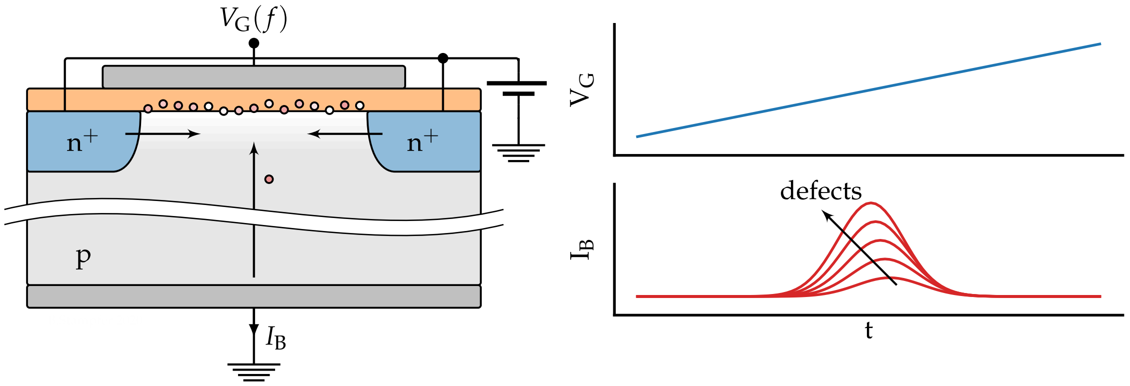


Figure 4.21. DCIV measurement in the top emitter configuration. The drain and source diodes are slightly forward biased, causing minority injection into the space charge region. As the voltage is swept from accumulation to inversion, defect charges recombine with the injected minorities. This causes majorities to move towards the space charge region, in turn creating a measurable bulk current. The scale of this current depends on the number of defects discharging, as illustrated on the right.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.
TU Bibliothek WIEN Your knowledge hub

4.2.5 Thermal Dielectric Relaxation Current (TDRC)

The last method presented here is the TDRC method. It was developed by Simmons and Mar in 1973 [109], is sensible to interface defects or near interface defects and can determine trap levels and lifetimes. The technique is based on the measurement of the non-steady-state generation or emission current over temperature. To measure the charge emission current of the defects, a MOSCAP is biased in accumulation and cooled to a cryogenic temperature. After cooling, the device is reverse biased, which leads to deep depletion as the thermal generation and recombination rates for carrier are negligible, which suppresses the formation of an inversion layer. Charge emission from the defects is negligible at this temperature. The device is then heated at a constant rate, which causes traps in the depletion region to emit, in turn causing a measurable emission current. Similarly, to measure the generation current, the device is first biased in inversion. After cooling, the reverse voltage is increased further, which again causes deep depletion. With the increasing temperature, additional carriers are generated, allowing measurement of the generation current. From the emission characteristics, the interface trap distribution in the upper (lower) half of the band gap can be determined for a n-type (p-type) semiconductor, while the generation current allows extraction of the carrier lifetimes.

4.3 Physical Characterization Methods

Aside from the electrical methods of defect characterization discussed in the previous parts of this chapter, there is a number of experimental methods which give insights to the physical, structural and stochiometric properties of the gate stack and thus allow to draw conclusions as to the nature of the defects present in a device. While these methods were not accessible to the author using the measurement equipment at the institute for microelectronics, for the sake of completeness and their relevance to the topic of this work the most popular of these methods and their use in defect characterization are briefly presented in this section.

4.3.1 Electron Paramagnetic Resonance (EPR)

ESR or EPR spectroscopy was pioneered by Zavoisky and Bleaney in 1944. It is sensitive to unpaired electrons present in a sample. In the context of MOS reliability, EPR was used to show the presence of P_b centers at the Si/SiO_2 interface [33, 110]. For this technique, a sample is placed inside a homogeneous magnetic field. The magnetic field can be generated using a Helmholtz coil as shown in Figure 4.22a. The coil is supplied with a DC current which is sweeped during the measurement. This magnetic field causes the magnetic moments of unpaired electrons to align with or against the field due to the Zeeman effect. The two alignment configurations have different energies depending on the strength of the external field, as shown in Figure 4.22b. In addition,

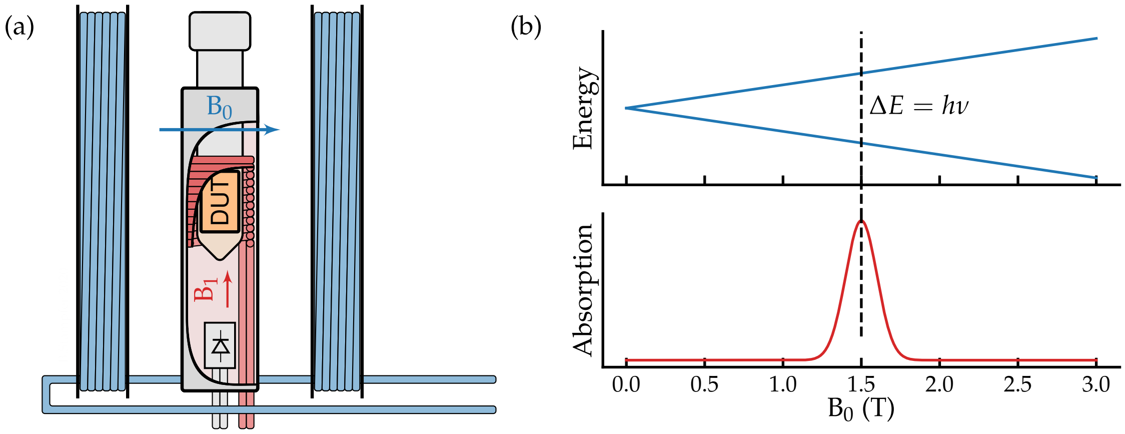


Figure 4.22. Principle of EPR spectroscopy. **(a)** A sample is placed in a homogeneous magnetic field and irradiated by microwave radiation. **(b)** The magnetic field causes the energy levels of unpaired electrons to split due to the Zeeman effect. When the split caused by the magnetic field coincides with the microwave energy, the absorption peaks.

the sample is irradiated with microwaves, and the microwave absorption is measured. When the energy of the incident microwaves is close to the split between the low and high energy states, resonance occurs and the microwave absorption of the sample peaks, see Figure 4.22c. The condition for this is given by

$$h\nu = g_e \mu_B B_0. \tag{4.13}$$

Here, g_e is the electron's g-factor and μ_B the Bohr magneton.

4.3.2 X-Ray Photoelectron Spectroscopy (XPS)

The X-ray photoelectron spectroscopy (XPS) method, developed by Siegbahn et al. in the 1950s can be used to study the surface chemistry of a material. In the context of MOSFETs, it is used to study the structure of oxides and interfaces [111].

XPS is based on the photoelectric effect. The sample is irradiated with X-rays. Upon collision between an X-ray photon and the sample, electrons are ejected from inner orbitals of the irradiated species. Part of them then escape the material and emit from the surface. The difference between the energy of the X-ray photon, the kinetic energy of the detected electron, and the work function of the material and detector gives the binding energy of the electron. From the amount of detected electrons at a specific binding energy the irradiated atomic species can be calculated. By adjusting the angle of the detector, the probed depth can be controlled, see Figure 4.23 for an illustration.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.
TU Bibliothek WIEN Your knowledge hub

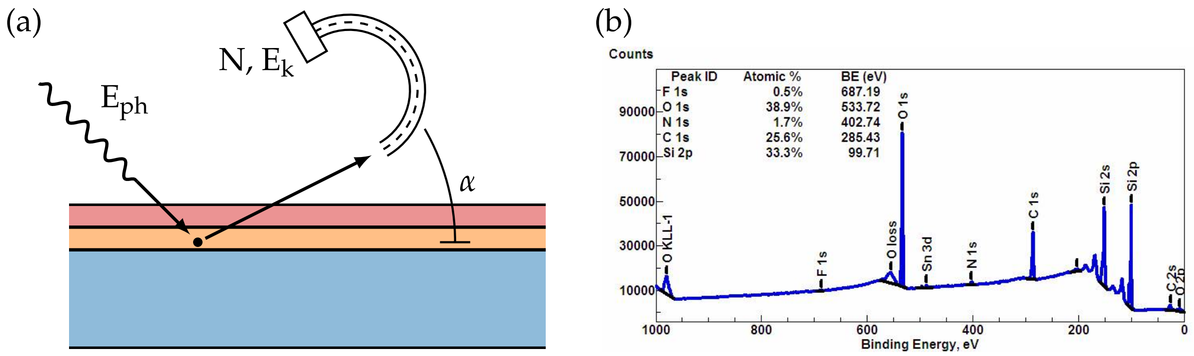


Figure 4.23. Principle of XPS. **(a)** X-ray photons are directed at the target. Upon collision they remove an electron from an atom. The electron moves to the sample surface and is emitted with an kinetic energy E_k. The binding energy of the electron can then be calculated from the detected energy and the work functions of the sample and the detector. The probed depth can be influenced by the angle of the detector. **(b)** The resulting peaks in the observed binding energies can be linked to the targeted species and their concentrations. Graph from [112]

4.3.3 Secondary Ion Mass Spectroscopy (SIMS)

Secondary ion mass spectroscopy (SIMS) is a technique which allows to characterize the surface composition of a sample by sputtering it with ions. Historically, emission of atoms and ions from a surface targeted with ions was first observed by J. J. Thomson in 1910 [114]. In 1949, the first prototype instrument was built by Herzog and Viehböck [115].

During SIMS measurements the secondary ions removed from the surface of the sample are measured. For this, the sample is targeted with ions which physically remove ions from the surface. Part of these secondary ions will be neutral, while others will be positively or negatively charged. The charged particles are then analyzed in a mass spectrometer to find the concentrations of atomic species on the sample surface. Due to the progressing removal of the sample surface during measurement, the method allows to profile the sample in depth. As a drawback of this method, the sample is destroyed during measurement. See Figure 4.24 for an illustration.

4.3.4 Neutron Activation Analysis (NAA)

Neutron activation analysis can be used to determine the concentration of elements within a sample. First experiments were performed by George de Hevesy and Hilde Levi in 1935 where they tested radiation induced in lanthanides after being irradiated by neutrons, and found that they can detect even small amounts of dysprosium in other lanthanides [116].

For the experiment, the sample is targeted with neutrons, which causes atoms in the sample to become radioactive isotopes. This induced radioactivity can then be measured. Elements can be distinguished by the type and energy of radiation they

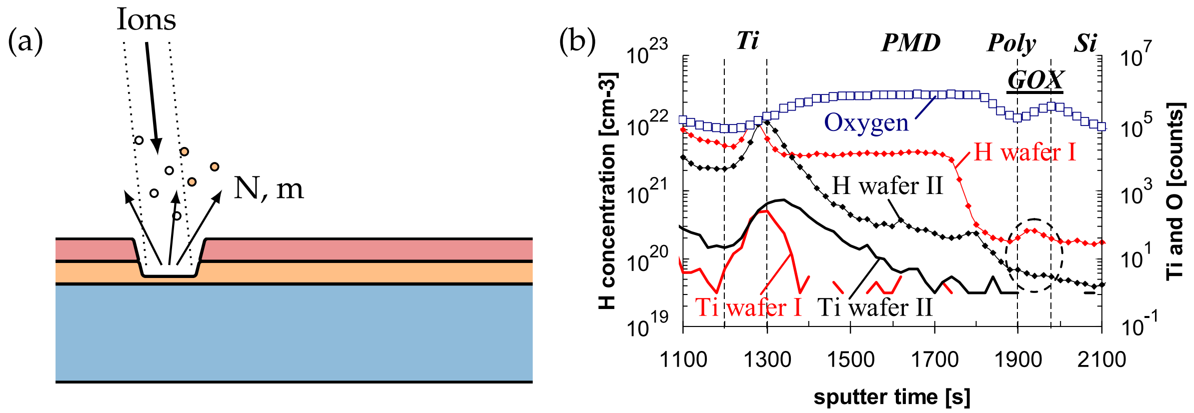


Figure 4.24. Principle of SIMS. **(a)** The sample is sputtered with ions. Upon collision they remove ions from the sample. These secondary ions are then analyzed in a mass spectrometer. **(b)** Exemplary results: the composition of the removed material over time is seen in the output of the mass spectrometer. Measurement results from [113]

emit. The measured spectrum is a mixture of emissions by the contained elements. To characterize the individual components, two methods are possible [116]:

- Destructive/ Radiochemical Neutron Activation Analysis: The sample is chemically separated and the individual fractions are analyzed for radiation.

- Non-destructive/ Instrumental Neutron Activation Analysis: The intact sample is measured a number of times with intervals in-between to measure decay rates.

CHAPTER 5

DEFECT PARAMETER EXTRACTION FROM RTN, TDDS, AND CV MEASUREMENTS

The conduction of the experiments at controlled bias conditions with exact timing and low noise measurement of the typically observed small currents is a challenge throughout the entire characterization of a technology. Another key challenge which will be discussed in this chapter is the evaluation of the resulting data set in order to extract parameters for the defects. Furthermore, these parameters should serve as basis for subsequently performed simulations and lifetime estimations and should lead to accurate results. For RTN and TDDS traces, this means analyzing the drain current traces for step heights and transition times of the individual defects at a specific gate bias, whereas for CV measurements, it means finding the defect densities at a specific frequency and gate bias. In the following the typical measurement signals, their analysis and the meaning for the defect parameters and distributions are discussed.

5.1 Random Telegraph and TDDS Signals

RTN and TDDS signals consist of a sequence of ΔV_{th} or I_{D} values. The data points may be sampled linear in time, i.e. in equidistant time intervals, or non-linear, e.g. with the sampling time increasing logarithmically. Some of the analysis methods presented in this section will work on both types of data, while some will require pre-processing or weighting of the data to yield correct results for non-linear sampled data.

For RTN data, the method of choice depends on the signal-to-noise ratio (SNR) of the recorded data, the parameters to be extracted from the measurements and the amount of data which has to be processed. In contrast to RTN data sets, TDDS signals from stress-recovery measurements have the advantages that they have a defined point where the recovery starts, and thus where the emission time can be measured from, and usually contain only a single emission step per defect. This allows for a much simpler data extraction using an edge-detection based method, as will be discussed at the end of this section.

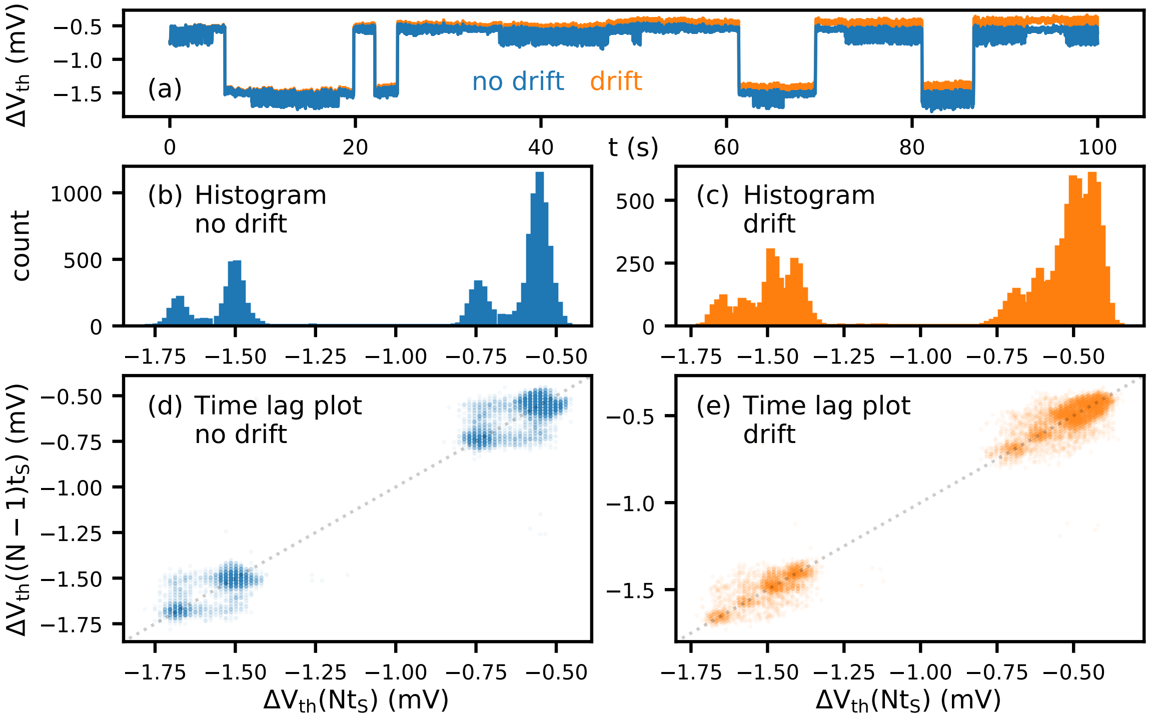


Figure 5.1. Exemplary RTN signals **(a)** extracted using histograms **(b,c)** and time lag plots **(d,e)**. The original RTN signal is shown in blue. The corresponding histogram **(b)** shows four partially overlapping peaks, indicating two defects. The time lag plot **(d)** shows the same peaks, but with better separation. In addition, the transitions between the states are shown in the off-diagonals. For the orange signal, a linear drift of $1.5\,\mu\mathrm{V/s}$ was added to show the influence of drift and low-frequency noise. Due to the drift, both the corresponding histogram **(c)** and time lag plot **(e)** show deformed peaks and it becomes difficult to distinguish the small defect. Originally published in [BSJ1]

5.1.1 Histogram and Lag Methods

Two of the most accessible methods for RTN and TDDS defect parameter extraction are the *histogram method* [117] and the slightly more advanced *time-lag method* [118]. The idea behind these methods is to analyze the occurrence of ΔV_{th} values in the recorded signals.

For the histogram method, as the name suggests, a histogram is created from the recorded data. Given a low enough SNR, the individual levels in the signal appear as peaks in the histogram. The peaks themselves typically have a Gaussian shape and their width is determined by the noise of the signal. If the method is applied to RTN signals, as is the most common use-case, a total of 2^N peaks will be present for N active defects in the signal. For the signal shown in Figure 5.1a with $N = 2$ defects, this relation is visible in Figure 5.1b. The height of the peaks is determined by the ratio of capture and emission times for the defects.

The more advanced time-lag method works by creating a two-dimensional histogram or scatter plot of each two consecutive ΔV_{th} points. This results in a better separation between the peaks, as the distance between the peaks is increased by a factor of $\sqrt{2}$ and points containing intermediate values of ΔV_{th} because a switch occurred during the sampling period are located on the off-diagonals. This is illustrated in Figure 5.1d. Four separate clusters are visible on the diagonal, and the most common transitions occur between the two lower, and between the two upper clusters. From this it follows that the average time constant of the smaller defect is much shorter than the one from the larger defect.

From the positions of the peaks in the histogram and the time-lag plots, the step heights of the defects visible in the signal can be extracted. If the peaks can be clearly separated a combination of defect states can be assigned to each peak, and the corresponding dwelling times can be extracted. If this is not possible, defect occupancies can be estimated by measuring the approximate areas of each peak.

One of the main drawbacks of these methods is their reliance on absolute values of the signal for obtaining defect states. This is shown for the orange signal in Figure 5.1a, c, and e. A small drift of 1.5 µV/s was added to the 100 s long signal. For the noise in the measurement, this is enough to obfuscate the number of defects in both the histogram and time-lag plots. Another drawback is the relatively high SNR required for these methods to work. This can be mitigated using a recent improvement by Martin-Martinez et al., the weighted time-lag method [119]. Despite their drawbacks, these methods are commonly used for extraction of parameters from RTN signals and fast visual inspection of RTN data.

5.1.2 Edge Detection

The main goal of RTN and TDDS data analysis is often to obtain the positions and amplitudes of the discrete steps in the signal. Thus, a natural approach is to use an edge detection algorithm to extract the steps from the data. Compared to other methods, the advantage of this approach is that edge detection is less susceptible to drift or low frequency noise present in the measurement data. In the following, two edge detection methods—the BCSUM algorithm and the Canny algorithm—will be outlined and benchmarked against each other.

The BCSUM Algorithm

The bootstrapping and cumulative sum (BCSUM) algorithm [120, 121] is a recursive algorithm which can be used to detect changes in the mean of a signal. It can be used on both uniform or non-uniformly sampled data with varying noise power. The general idea is to describe the measurement as two distinct constant signals separated by a step of height d at time τ. This is described by the mean shift model:

$$r(t) = \mu_0 + d\sigma(t - \tau). \tag{5.1}$$

Here, μ_0 is the mean of the first signal, d the step height, σ the step function and τ the change point. The discrete measurement signal is described by a vector $\vec{r}$ =

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.
TU Bibliothek WIEN Your knowledge hub

$[r_0, r_1, ..., r_{N-1}]$ of length N. Assuming that the samples in the signal are normally distributed around the mean value in each of the two partial signals, the distribution for the two parts are given by $p_0(r)$ and $p_1(r)$ with means μ_0 and $\mu_1 = \mu_0 + d$ and variances $\sigma_0 = \sigma_1 = \sigma$. With this, a log-likelihood ratio can be defined which changes its sign depending on which part of the signal the sample more likely belongs to:

$$s_i = \ln \frac{p_1(r_i)}{p_0(r_i)}. \tag{5.2}$$

The cumulative log-likelihood for any slice of a signal is defined as

$$S_m^n = \sum_{i=m}^{n} s_i. \tag{5.3}$$

With the assumed Gaussian distributions of the partial signals

$$p(r_i) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(r_i - \mu)^2}{2\sigma^2}\right), \tag{5.4}$$

the log-likelihoods ratio can be written as

$$s_i = d\frac{r_i - \mu_0 - d/2}{\sigma^2}, \tag{5.5}$$

with an expectation values of $s_i = \mp d^2/(2\sigma^2)$ before and after the step respectively. The change point index is thus found as

$$i_\mathrm{c} = \{k : \min |_k S_0^k\}. \tag{5.6}$$

As the underlying statistical distributions of the signals are unknown a priori, bootstrapping—a statistical method based on random sampling with repetition—is used for parameter estimation. For this, a cumulative sum $\vec{C}$ is defined in lieu of S:

$$C_n^* = \sum_{i=0}^{n} (r_i^* - \nu) \text{ with } \nu = \frac{1}{N}\sum_{N} r_i^*, \tag{5.7}$$

where $\vec{r^*} = [r_0^*, r_1^*, ..., r_{N-1}^*]$ is a signal randomly sampled from the original signal. From $\vec{C^*}$,

$$\gamma^* = \max \vec{C^*} - \min \vec{C^*} \tag{5.8}$$

is obtained for a number B of bootstraps. A sensitivity parameter $\epsilon \in [0, 1]$ is further defined for the algorithm. If then γ from the original signal is above the ϵ quantile of the distribution of γ^*, this is taken as indication that a changepoint occurred in the signal at $\max |\vec{C}|$ of the original signal. In this case the algorithm is recursively applied on the two split parts of the signal.

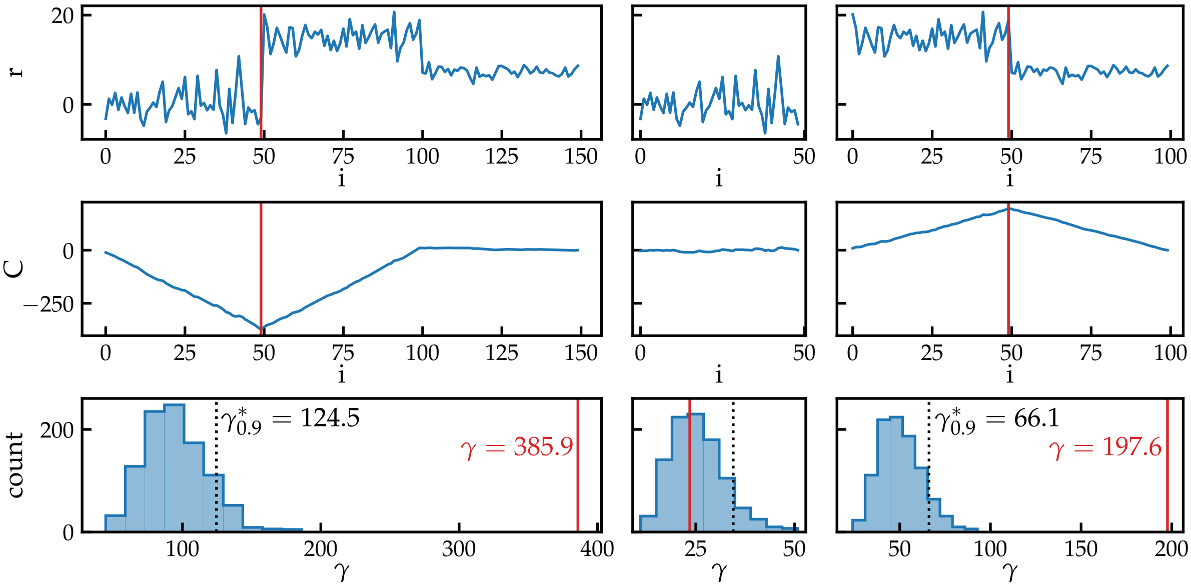


Figure 5.2. The BCSUM algorithm applied to an exemplary signal r. The columns show the initial step **(left)**, and the first recursion for the first part **(center)** and second part **(right)** of the signal. **Top:** Signal used during recursion. **Middle:** CSUM signal of the original data. **Bottom:** Histogram of γ^* values obtained from the bootstrapped data, together with the $\epsilon = 0.9$ quantile value (Black), and γ from the original data (Red). Recursion ends when $\gamma < \gamma^*_\epsilon$, as is the case in the bottom center plot.

An example is given in Figure 5.2. The BCSUM algorithm is applied to on the signal on the left. A step is found at $i \approx 50$, and the signal is split into the parts shown in the center and on the right. For the part shown in the center, no step is detected and the algorithm stops. For the part shown on the right on the other hand, another step is found and the algorithm continues by analyzing the two splits of this signal.

The Canny Algorithm

The Canny filter is an algorithm commonly used in computer vision to detect edges in recorded images, published in 1987 [122]. Unlike 2D images analyzed in computer vision, or even 3D images analyzed in medical applications, RTN data are 1D signals which considerably simplifies the filter. In his work, Canny finds the optimal filter for step-edges based on three criteria. The criteria he defines for the filter are a low probability of error—equivalent to the SNR of the filter, a good localization of the edge position and the suppression of multiple detections per edge. Based on this, he found that there is a trade-off between the SNR and the localization, with the SNR increasing proportional to the square root of the filter width and the localization increasing inversely proportional to it. He further showed that the optimal filter kernel can be closely represented by the derivative of a Gaussian, which leads to improved performance. The first derivative of a Gaussian is given by

$$G'(t) = A\frac{\mathrm{d}}{\mathrm{d}t}\exp\left(-\frac{t^2}{2\sigma^2}\right) = -A\frac{t}{\sigma^2}\exp\left(-\frac{t^2}{2\sigma^2}\right), \tag{5.9}$$

with some normalization constant A. The filter response is the convolution of the kernel—truncated to a width W and discretized as $G'^{*}(n)$—and the signal $f(n)$:

$$H(n) = \sum_{i=-W}^{W} G'^{*}(i)f(n-i). \tag{5.10}$$

In the case of a 1D signal, non-maximum suppression, i.e. finding only the strongest edges in the signal, can be performed simply by finding the local extrema in the response:

$$H_\mathrm{m}(n) = \begin{cases} H(n) & (|H(n)| > H(n+1)) \wedge (|H(n)| > H(n-1)) \\ 0 & \text{else.} \end{cases} \tag{5.11}$$

Finally, thresholding has to be applied to suppress responses caused by noise. For this, a threshold H_th has to be applied, selected either by a noise estimation from evaluation of the filter response or manually by the user:

$$H_\mathrm{t}(n) = \begin{cases} H(n) & |H_\mathrm{m}(n)| > H_\mathrm{th} \\ 0 & \text{else.} \end{cases} \tag{5.12}$$

Positive values in $H_t(n)$ now correspond to positive edges, negative values to negative edges. The heights of the edges can be determined from the original signal or from the peaks in $H(n)$. An illustration of the method can be found in Figure 5.3.

Step Heights and Time Constants

Once the steps in the signal are obtained, they can be binned in step height to assign them to individual defects. If all steps in a bin are from one defect, the step height is simply the mean over all contributors. The charge capture time can be obtained by calculating the average time between positive and successive negative steps and vice versa for the charge emission time. If different defects with similar step heights show in the measurement data, extracting them is—in most cases[1]—not possible using this method. If the defects need to be characterized, extraction may be possible using a hidden Markov model (HMM) based approach.

---

[1] Extraction may be possible if the defects show correlated behavior, in which case a state machine may be used to assign the steps to defects, see [BSC5].

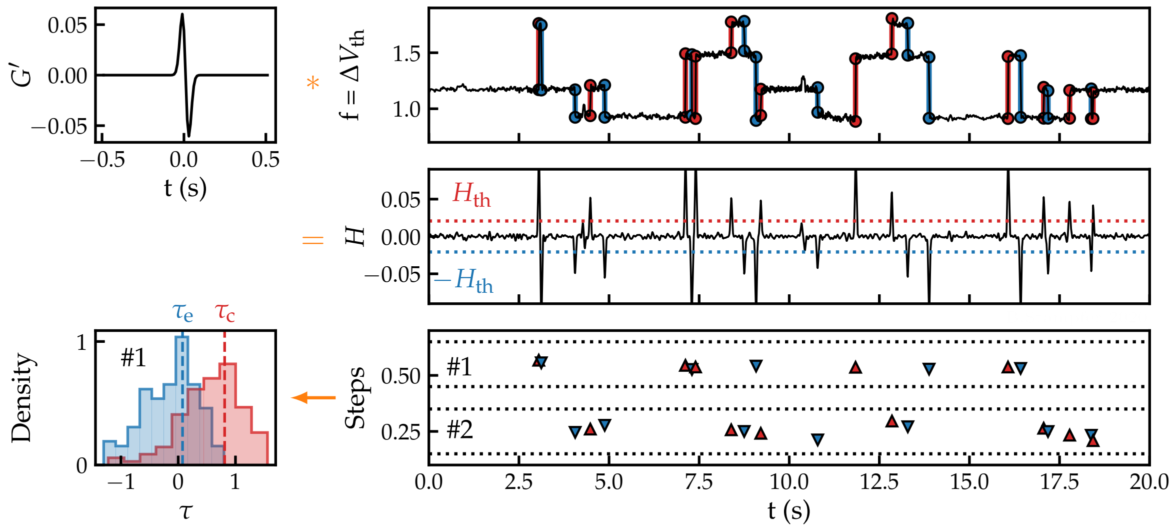


Figure 5.3. The Canny algorithm for edge detection demonstrated for an exemplary measurement trace. The input signal $f(t)$ is convoluted with the first derivative of a Gaussian pulse with a chosen variance G'. This gives a signal $H(t)$, which exhibits peaks corresponding to the steps in the original signal. All peaks above a chosen threshold are recognized and used to mark the positions of the steps in $f(t)$. The height of the peaks can be obtained from the original signal or from the height of the peaks. Modified from [BSJ1]

Comparison of Methods

To give an idea of how the BCSUM and the Canny algorithm perform when confronted with measurement data affected by noise, the algorithms are evaluated considering three test cases. The first test case used original data recorded in a low noise environment while the second and third test cases use the same data but with $\sigma = 0.1\,\mathrm{mV}$ and $\sigma = 0.2\,\mathrm{mV}$ of added Gaussian noise. The test cases are shown in Figure 5.4. The top plot is always the data to be analyzed and the two plots below show the steps extracted using the Canny and the BCSUM algorithm. The small plots to the right are histograms for the dwelling times of the smaller defect. As can be seen, the accuracy for both methods considerably decreases as the noise level approaches the step height of the defects. For the results published in [BSC5], [BSJ2] (shown in Section 6.1), and [BSJ7] (shown in Section 6.3) the author has used the Canny algorithm to evaluate the RTN data.

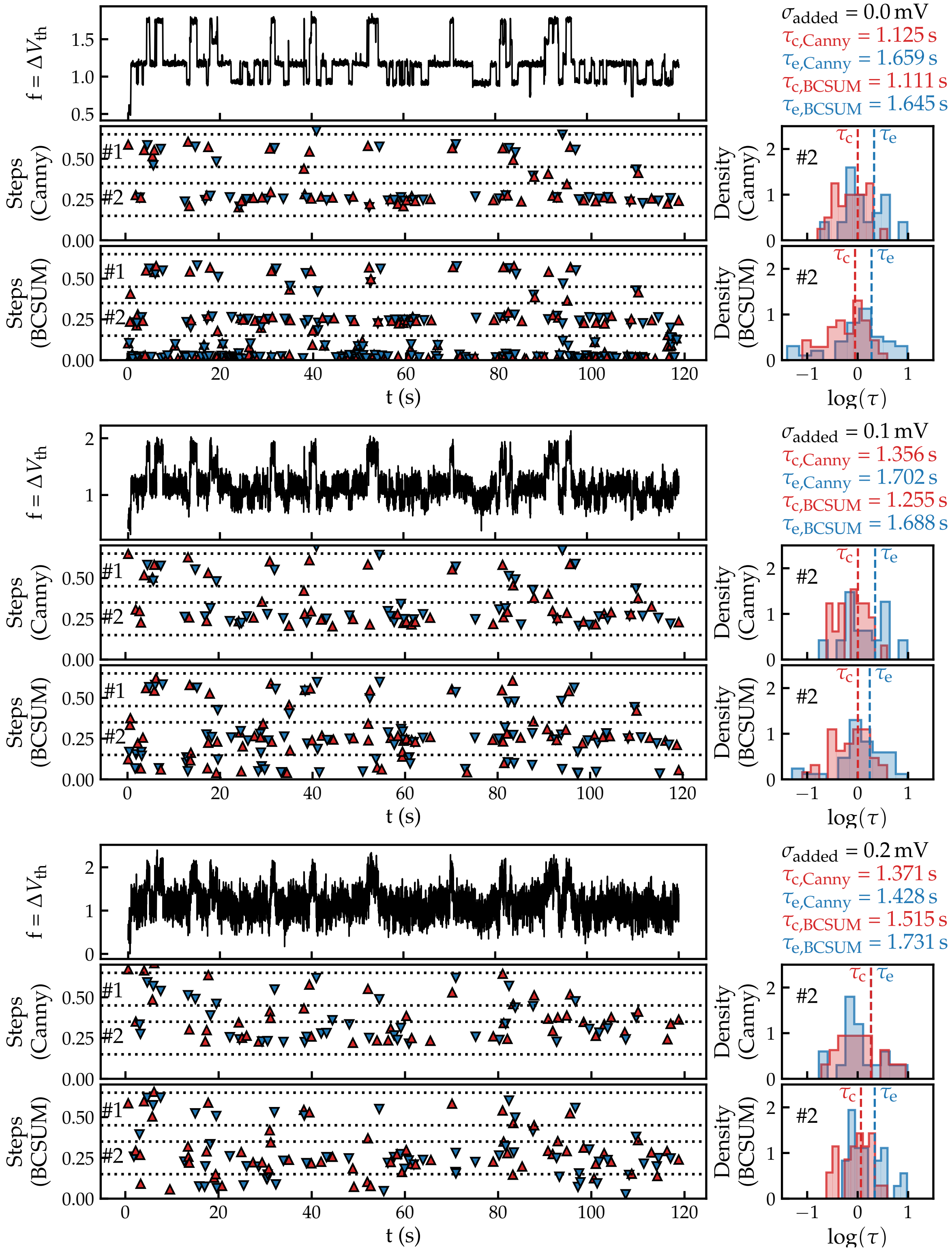


Figure 5.4. Comparison of the Canny edge detector and the BCSUM method for signals with different noise floors. The signal contains two defects with step heights of $\eta_1 \approx 0.25\,\mathrm{mV}$ and $\eta_2 \approx 0.55\,\mathrm{mV}$. **(top)** Original signal, both methods agree within 2% of absolute deviation of the extracted charge transition times. **(middle)** $\sigma = 0.1\,\mathrm{mV}$ of added noise. The extracted capture time increases as the algorithms miss shorter transitions. **(bottom)** $\sigma = 0.2\,\mathrm{mV}$ of added noise. With a noise level similar to the step height of the extracted defect, spurious transitions affect the detections while proper steps are missed, leading to false results.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.
TU Bibliothek WIEN Your knowledge hub

5.1.3 Hidden Markov Models

As stated earlier, defects are commonly described mathematically using Markov models. In the simple case of a two-state defect and negligible measurement noise, each state in the model can be associated with a unique level of the channel current. The effect of the defect being in the charged state on the current can be measured, simply by measuring the height of the observed steps. Neutral defects are assumed to have no influence on the channel current. At any given gate bias, the capture and emission times, i.e. the charge transition times between the charged and neutral state can be measured by averaging the dwelling times in the high and low current states. This gives the transmission rates between the two states of the Markov model. Thus, the sequence of states of the Markov process is fully observable[2] from the measurement and the parameters of the Markov model can be found.

For multi-state defect models, such as the 4-state model, this is generally not true. Since multiple states of the defect will have the same effect on the channel current, the states of the defect can not be known from the measurement—they are hidden. Even for the two-state model, when considering real data with measurement noise, associating a point in the measurement to a defect state is not possible.

Assuming the measurement noise as Gaussian, this scenario can be described using a hidden Markov model with Gaussian observations. Such a model is defined by the following parameters:

- The state vector of the Markov model $\vec{I}$ of length N
- The $N \times N$ matrix of transition probabilities $\mathbf{P}$
- The vector of mean observations for each state $\vec{\mu}$ of length N
- The vector of variances of observation for each state $\vec{\sigma^2}$ of length N
- The sequence of states $\vec{x}$ of length T
- The sequence of observations $\vec{y}$ of length T—i.e. the measured data

Given the measured data and the underlying Markov chain, the hidden Markov model can be trained in order to provide the parameters of the model and the probabilities of measuring a certain ΔV_{th} or I_{D} for each state[3]. This can be done by using the Baum-Welch algorithm [123], which finds the local maximum-likelihood solution of the problem. The Baum-Welch algorithm is an expectation-maximization algorithm using the forward-backward algorithm. Apart from defect characterization, this method is used for DNA sequencing, speech and text recognition. The algorithm iterates between an expectation step, where statistics are collected from the current model parameters and

---

[2] *Observable* in the context of HMMs refers to being able to measure the state the model is in, this is different from *observable* in the context of control theory.

[3] This is called *emission probability* in the context of hidden Markov models—not to be confused with the charge emission probability of a defect

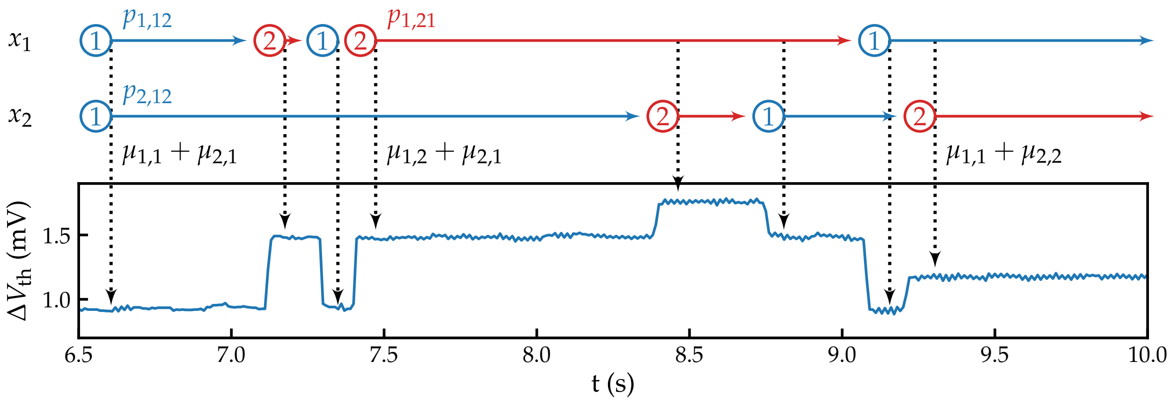


Figure 5.5. The measured drain current can be described by the evolution of the underlying Markov chain(s) of the defects. The reverse, however, may not be true. If reconstruction of the individual defect states is not possible from the recorded drain current data, a hidden Markov model in conjunction with the Baum-Welch algorithm may be used to estimate the model parameters, i.e. transistion probabilities **P** and emissions $\vec{\mu}$.

the observations, and a maximization step where the model parameters are improved using those statistics.

Initialization: The parameters $\vec{\mu}$, $\vec{\sigma^2}$ and **P** are initialized, either using estimates or by using random values within certain bounds. Prior to calculation the following quantities are introduced:

- The transition probability matrix **P** is discretized in time to yield the state transition probability distribution **A**. For equal time steps this is a $N \times N$ matrix with elements:

$$a_{ij} = \mathcal{P}(x_{t+1} = j | x_t = i). \tag{5.13}$$

- From the vectors of the mean and variance values for the Gaussian distribution assigned to each state, and the observations, the observation probability distribution **B** is calculated. These are the probabilities of the observations being caused by each of the model states:

$$b_{it} = \mathcal{P}(y_t | x_t = i). \tag{5.14}$$

  As we assume Gaussian noise on our observations, they are calculated from the respective PDFs with means μ_i and variances σ_i^2.

- The initial state distribution $\vec{\pi}$ gives the initial probabilities per state

$$\pi_i = \mathcal{P}(x_1 = i). \tag{5.15}$$

  $\vec{\pi}$ is commonly initialized to equal numbers, meaning that there is no preferred initial state for the model.

Expectation step: The expectation step is performed using the forward-backward algorithm [124]:

- In the forward step, vectors $\vec{\alpha}_i$ with the probabilities of arriving in state i at a certain time, given the observations and current parameters are obtained:

  $$\alpha_{i,t} = \mathcal{P}(y_1, y_2, ..., y_t, x_t = i | \mathbf{A}, \mathbf{B}, \vec{\pi}). \tag{5.16}$$

  The values α_i are recursively calculated from the start of the sequence:

  $$\alpha_{i,1} = \pi_i b_{i,1} \tag{5.17}$$
  $$\alpha_{i,t+1} = b_{i,t+1} \sum_j \alpha_{jt} a_{ji}. \tag{5.18}$$

- In the backward step, vectors $\vec{\beta}_i$ with the probabilities of starting from a state i at a certain time, given again the observations and parameters are obtained:

  $$\beta_i t = \mathcal{P}(y_{t+1}, y_{t+2}, ..., y_T, x_t = i | \mathbf{A}, \mathbf{B}, \vec{\pi}). \tag{5.19}$$

  The values β_i are recursively calculated from the end of the sequence:

  $$\beta_{i,T} = 1 \tag{5.20}$$
  $$\beta_{i,t} = \sum_j \beta_{j,t+1} a_{i,j} b_{j,t+1}. \tag{5.21}$$

Maximization step: To find the model parameters that maximize the probability of observation:

- The probability of being in state i at time t for the model and the observation is defined as

  $$\gamma_{i,t} = \mathcal{P}(x_t = i | \vec{y}, \mathbf{A}, \mathbf{B}, \vec{\pi}). \tag{5.22}$$

  This is calculated from the product of forward and backward probabilities to and from each state:

  $$\gamma_{i,t} = \frac{\alpha_{i,t}\beta_{i,t}}{\sum_j \alpha_{j,t}\beta_{j,t}}. \tag{5.23}$$

- The probabilities of transitions between all states at all times are defined as

  $$\xi_{i,j,t} = \mathcal{P}(x_t = i, x_{t+1} = j | \vec{y}, \mathbf{A}, \mathbf{B}, \vec{\pi}), \tag{5.24}$$

  which are calculated from the forward probability to the first state, the backward probability from the second state, and $a_{i,j} b_{j,t+1}$—the probability of the transition and observation after the transition:

$$\xi_{i,j,t} = \frac{\alpha_{i,t} a_{i,j} b_{j,t+1} \beta_{j,t+1}}{\sum_k \sum_l \alpha_{k,t} a_{k,l} b_{l,t+1} \beta_{l,t+1}}. \tag{5.25}$$

- From this, the new parameters of the model can be calculated using the Baum-Welch reestimation equations:

  $$\pi_i^* = \gamma_{i,1} \tag{5.26}$$

  $$a_{i,j}^* = \frac{\sum_t \xi_{i,j,t}}{\sum_t \gamma_{i,t}}. \tag{5.27}$$

  For Gaussian observations the observation probabilities **B** are not reestimated directly, but instead $\vec{\mu}$ and $\vec{\sigma^2}$ are calculated from the weighted observation sequence for each state:

  $$\mu_i^* = \frac{\sum_t \gamma_{i,t} y_t}{\sum_t \gamma_{i,t}} \tag{5.28}$$

  $$\sigma_i^{2*} = \frac{\sum_t \gamma_{i,t} (y_t - \mu_{i,t})^2}{\sum_t \gamma_{i,t}}. \tag{5.29}$$

  Afterwards, $\mathbf{B}^*$ is recalculated as in Equation (5.14).

End: The algorithm always converges to a local maximum in the likelihood, whose absolute or relative change between iterations is typically used as the convergence criterion. Once the algorithm converged, the transition rates of the defect **P** can be recovered from **A**. The method may be used to fit multiple defects by combining them into a single Markov chain. In this case, $\mathbf{A}^*$, $\vec{\mu}^*$, and $\vec{\sigma^2}^*$ need to be calculated in a way to preserve the independence of the defects, see [125].

As the computational cost of the calculation increases exponentially with the number of states, a better approach might be to use a factorial HMM (FHMM) [126] to extract the defect parameters, as proposed by Puglisi *et al.* [127]. Implementations of HMMs are readily available in many programming languages, e.g., for python [128, 129].

Similar to the histogram and time-lag methods, HMMs depend on the absolute values of the drain current. Such approaches are always affected by slow drift or low frequency noise of the measurement data. To mitigate this, the iterative solver can be extended by a method for baseline correction. Suitable methods include spline smoothing, asymmetric least squares [130], and local regression [131]. HMMs have been used by the author in [BSJ7] (shown in Section 6.3) to characterize a multi-state defect which produced aRTN and in [BSJ5] to investigate correlated RTN signals.

5.1.4 Estimations for Trap Level and Position

Once the charge capture and emission times have been extracted at a number of gate voltages, the vertical position as well as the energy of a defect can be calculated. For this, the gate bias at the intersection point where $\tau_\mathrm{c} = \tau_\mathrm{e}$ and the slope $\partial \ln(\tau_\mathrm{c}/\tau_\mathrm{e})/\partial V_\mathrm{G}$ needs to be calculated from the measurement data.

We start from general rates for capture and emission, which can be given analytically considering a two-state model by:

$$k_\mathrm{c,e} = k_\mathrm{0c,e} \exp\left(-\frac{\mathcal{E}_\mathrm{c,e}}{k_\mathrm{B}T}\right). \tag{5.30}$$

Here, $k_\mathrm{0c,e}$ are prefactors and $\mathcal{E}_\mathrm{c,e}$ are energy barriers for capture and emission. Taking the logarithm and the first derivative after V_G yields

$$\frac{\partial \ln k_\mathrm{c,e}}{\partial V_\mathrm{G}} = \frac{\partial \ln k_\mathrm{0c,e}}{\partial V_\mathrm{G}} - \frac{1}{k_\mathrm{B}T}\frac{\partial \mathcal{E}_\mathrm{c,e}}{\partial V_\mathrm{G}}. \tag{5.31}$$

For this estimation, the bias dependence of the prefactors is neglected. The subtraction of the equations for capture and emission rates yields

$$\frac{\partial \ln k_\mathrm{c}/k_\mathrm{e}}{\partial V_\mathrm{G}} = -\frac{1}{k_\mathrm{B}T}\frac{\partial(\mathcal{E}_\mathrm{c} - \mathcal{E}_\mathrm{e})}{\partial V_\mathrm{G}} = -\frac{1}{k_\mathrm{B}T}\frac{\partial E_\mathrm{t}}{\partial V_\mathrm{G}}, \tag{5.32}$$

with $\partial(\mathcal{E}_\mathrm{c} - \mathcal{E}_\mathrm{e})$ replaced by the change in defect level ∂E_t. Assuming inversion and approximating the oxide free of charge, any change in gate bias should influence the defect level proportional to the relative depth of the defect in the oxide:

$$\frac{\partial E_\mathrm{t}}{\partial V_\mathrm{G}} = -q\frac{d}{t_\mathrm{ox}}. \tag{5.33}$$

Now we can express the relative depth of the defect using the steepness of the transition rates or time constants:

$$\frac{d}{t_\mathrm{ox}} = -\frac{k_\mathrm{B}T}{q}\left(\frac{\partial \ln \tau_\mathrm{c}/\tau_\mathrm{e}}{\partial V_\mathrm{G}}\right). \tag{5.34}$$

Integration of Equation (5.33) further yields the trap energy

$$E_\mathrm{t} = -q\frac{d}{t_\mathrm{ox}}V_\mathrm{G} + C. \tag{5.35}$$

Finally, the integration constant can be found by evaluating Equation (5.35) at the voltage $V_\mathrm{G,i}$ of intersection of the charge capture and emission time characteristics, i.e. $\tau_\mathrm{c}(V_\mathrm{G,i}) = \tau_\mathrm{e}(V_\mathrm{G,i})$, where $E_\mathrm{t} = E_\mathrm{F}$. This allows to calculate the defect energy at zero field

$E_{\mathrm{t,0}}$, i.e. at $V_{\mathrm{G}} = -V_{\mathrm{FB}} - \phi_{\mathrm{s}}$ [132]:

$$E_{\mathrm{t,0}} = E_{\mathrm{F,i}} + q\frac{d}{t_{\mathrm{ox}}}(V_{\mathrm{G,i}} - \phi_{\mathrm{s}} - V_{\mathrm{FB}}), \tag{5.36}$$

with the surface potential ϕ_{s} and the flatband voltage V_{FB}. It has to be noted that this estimation has a number of shortcomings:

- It occasionally yields negative distances for defects interacting with the gate. In this case the defect might have been interacting with the gate, in which case the distance and Fermi level should be considered from the gate.
- It neglects the bias dependence of the prefactors. This usually leads to an overestimation of the distances and energies.
- It assumes a single transition barrier, i.e. it does not work for defects which cannot be modeled using a two state model.

Nonetheless, in many cases it allows for a quick and reasonable estimation of defect parameters, which may in turn provide a suitable initial guess for more accurate numerical simulations. The author has employed the method in this fashion in [BSJ7] (shown in Section 6.3) and [BSJ2] (shown in Section 6.1).

5.1.5 Frequency Domain Methods

As described in Section 4.1.1, RTN caused by a single defect will appear as a Lorentzian in the frequency spectrum. By characterizing the individual Lorentzians of a full PSD, the average transition rates of the defects can be obtained. Variation of the measurement voltage further allows to find the point where the noise energy of a defect is at a maximum, which means that the occupancy of the defect is at 50%. At this point, the step height of the defect can be calculated, which allows calculation of the occupancy at other voltages.

In devices with many active defects, the relation between the noise power calculated from the drain current and the number of defect states close to the Fermi level allows to obtain the defect density over the energy of the band gap [133]. The basic equation which describes this relation is given as [134]

$$\frac{S_{I_{\mathrm{D}}}(f)}{I_{\mathrm{D}}^2} = \left[\frac{R}{N_{\mathrm{s}}} + \alpha\mu\right]^2 \cdot \frac{N_{\mathrm{ot}}(E_{\mathrm{F}})k_{\mathrm{B}}T}{WL\gamma} \cdot \frac{1}{f}. \tag{5.37}$$

Here, $S_{I_{\mathrm{D}}}(f)$ is the noise spectral density, I_{D} is the average drain current, R is the variation of free-carrier concentration due to carrier trapping, N_{s} is the density of free channel electrons, α gives the influence of the channel mobility on the number of trapped carriers, μ is the carrier mobility, N_{ot} the density of oxide traps, k_{B} the Boltzmann constant, T the temperature, WL the channel area, γ the tunneling attenuation length and f the frequency.

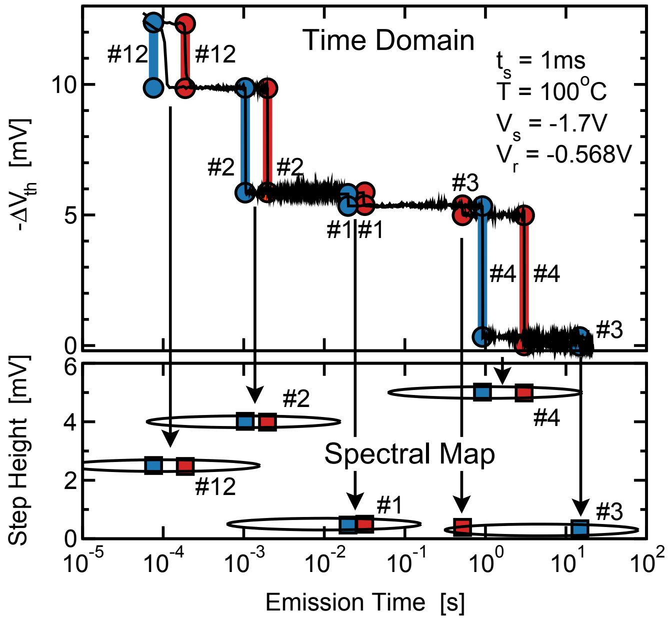


Figure 5.6. Defects found from TDDS measurements. Recovery traces **(top)** are analyzed using a step detection algorithm and the steps are plotted in the step height–emission time plane **(bottom)**. If enough traces are plotted, clusters will form for each defect which emitted during recovery. The clusters are distributed exponentially in time and normally in step height. To obtain the capture time of the defects, measurements at varying stress times have to be performed. From [135]

5.1.6 Time-Dependent Defect Spectroscopy

For the extraction of TDDS data an edge detection method as outlined in Section 5.1.2 is applied to the recovery data. As the charge emission probabilities are strongly modulated by the applied bias, the dwelling times of the defects at $V_{\mathrm{G,rec}}$ are practically identical to the point in time in the recovery trace where the charge is emitted on average. The extracted steps are plotted in a scatter map with the step height and the emission time as axes as shown in Figure 5.6. Plotting multiple repetitions of the measurement leads to the formation of clusters in this so-called spectral map [135]. Each of the clusters corresponds to a single defect. The points in the clusters are distributed normally in step height and exponentially in time around the emission time of the defect. By dividing the number of emissions in each cluster by the number of repetitions, the capture probability of the defects can be obtained for the applied stress conditions. The capture time at $V_{\mathrm{G,str}}$ can then be obtained by plotting the capture probability of the defect over the stress time which is observed to follow an exponential distribution, as expected for a Markov process and shown in Figure 5.7.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.
TU Bibliothek WIEN Your knowledge hub

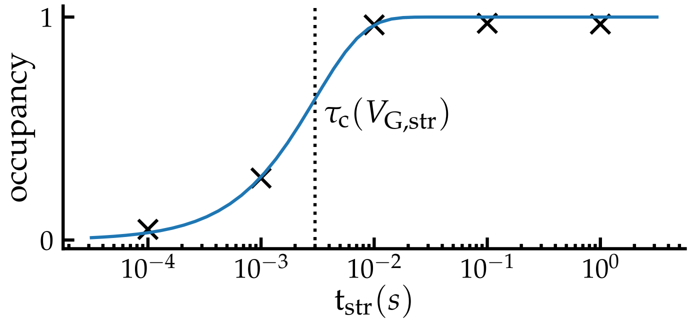


Figure 5.7. Illustration of charge capture time extraction from a TDDS measurement. Assuming a low occupancy of the defect at recovery conditions and a sufficient relaxation time, the average occupation of the defect at the end of the stress phase equals the probability of observing the defect discharging during recovery. By plotting this occupancy, i.e. the ratio of the number of emission events over the number of measurement repetitions, over the stress time, the exponential CDF for charge capture can be obtained. The charge capture time of the defect at the stress condition can then be extracted by fitting the theoretical distribution.

TU Bibliothek Your knowledge hub WIEN
Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

5.2 Capacitance-Voltage Measurements

Depending on the type of measurements that can be performed on the DUT and setup, different methods of defect parameter extraction are available. In general, trap densities can be extracted both from the capacitance and the conductance. In fact, they are related to each other by the Kronig-Kramers relation [4]. The effect of defects at the interface or in the oxide on CV measurements depends on the charge transition times of the defects in relation to the temporal parameters of the measurement. The following cases can be distinguished and are illustrated in Figure 5.8:

- **Fixed charges or very slow defects**, which do not change their charge state during the measurement will add a constant voltage offset to the theoretical defect-free CV curve. This is due to the additional voltage which has to be supplied to the gate to compensate these charges.
- **Defects fast enough to react to the gate bias sweep**, but too slow to react to the AC frequency. These defects will charge or discharge during the measurement, at which point part of the change in gate bias is required to supply the opposing charge to the gate. This modifies the relation between gate voltage and surface potential, which appears as a stretch-out in the CV curve where such defects (dis)charge.
- **Defects fast enough to react to the AC measurement signal**. In addition to the stretch-out of the CV curve, these defects cause additional small-signal current and thus add to the measured capacitance.

5.2.1 Capacitance Methods

The general shape of the differential capacitance obtained from a MOSCAP or MOSFET depends on measurement conditions. To understand why, consider the differential capacitance of the device as viewed from the gate:

$$C = \frac{\mathrm{d}Q_\mathrm{G}}{\mathrm{d}V_\mathrm{G}}. \tag{5.38}$$

We can replace the gate charge with the opposing semiconductor (Q_s) and interface (Q_it) charges. Further, the gate voltage can be written as the sum of the flatband (V_FB) and oxide (V_ox) voltages and the surface potential (ϕ_s) [105]:

$$C = \frac{-\,\mathrm{d}(Q_\mathrm{s} + Q_\mathrm{it})}{\mathrm{d}(V_\mathrm{FB} + V_\mathrm{ox} + \phi_\mathrm{s})} \tag{5.39}$$

By rearranging the equation and replacing the semiconductor charge by the sum of hole charge (Q_p), bulk charge in the space charge region (Q_b) and electron charge (Q_n), we

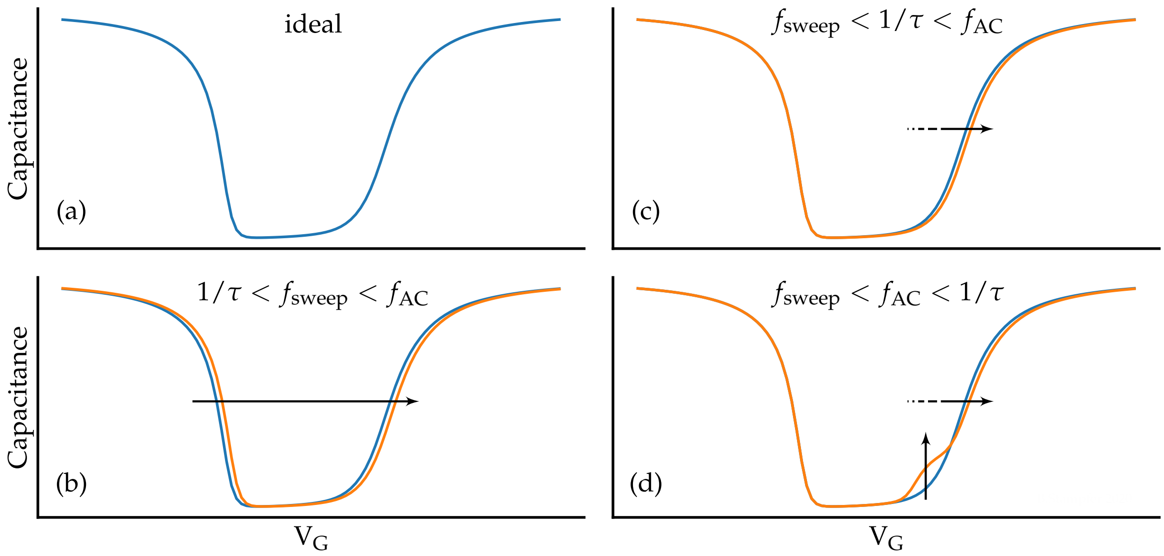


Figure 5.8. Schematic CV curves of a pMOS transistor, affected by defects located close to the conduction band. **(a)** Defect free CV curve. **(b)** Fixed charges or defects with $\tau f_{\text{sweep}} > 1$ shift the CV curve along the voltage axis. **(c)** Defects fast enough to react to the DC sweep but too slow to react to the AC frequency cause a stretch-out of the CV curve. **(d)** Defects fast enough to react to the AC frequency cause additional capacitance in the CV curve. In general, a superposition of these effects will be obtained, depending on the chosen measurement parameters.

obtain:

$$\begin{aligned} C &= -\left(\frac{\mathrm{d}V_{\text{ox}}}{\mathrm{d}Q_{\text{s}} + \mathrm{d}Q_{\text{it}}} + \frac{\mathrm{d}\phi_{\text{s}}}{\mathrm{d}Q_{\text{p}} + \mathrm{d}Q_{\text{b}} + \mathrm{d}Q_{\text{n}} + \mathrm{d}Q_{\text{it}}}\right)^{-1} \\ &= \left(\frac{1}{C_{\text{ox}}} + \frac{1}{C_{\text{p}} + C_{\text{b}} + C_{\text{n}} + C_{\text{it}}}\right)^{-1}. \end{aligned} \tag{5.40}$$

This is shown as an equivalent circuit in Figure 5.9a for an nMOS device. As can be seen from the figure, depending on the measurement conditions, any of three principal types of CV curves can be obtained. Common among all of them are the accumulation and depletion regions. In accumulation, the majority carriers are at the interface and C_{p} is large, leading to $C \approx C_{\text{ox}}$. Towards depletion, a space charge region forms which—together with the interface charges—compensates the gate charge by varying its thickness. At higher voltages, an inversion layer will form if minority carriers can be generated or supplied fast enough to follow the sweep frequency. If this is not the case, the depletion layer will widen and the device will operate in deep depletion. If an inversion layer forms, a low- or high frequency curve is obtained[4]. If the measurement frequency is too high for the inversion layer charge density to follow the AC signal,

---

[4] Note that the meanings of high- and low-frequency curves are ambiguous. Depending on the context, they might refer to measurements where either the interface defects or the inversion charges are (un)able to respond to the measurement frequency.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.
TU Bibliothek WIEN Your knowledge hub

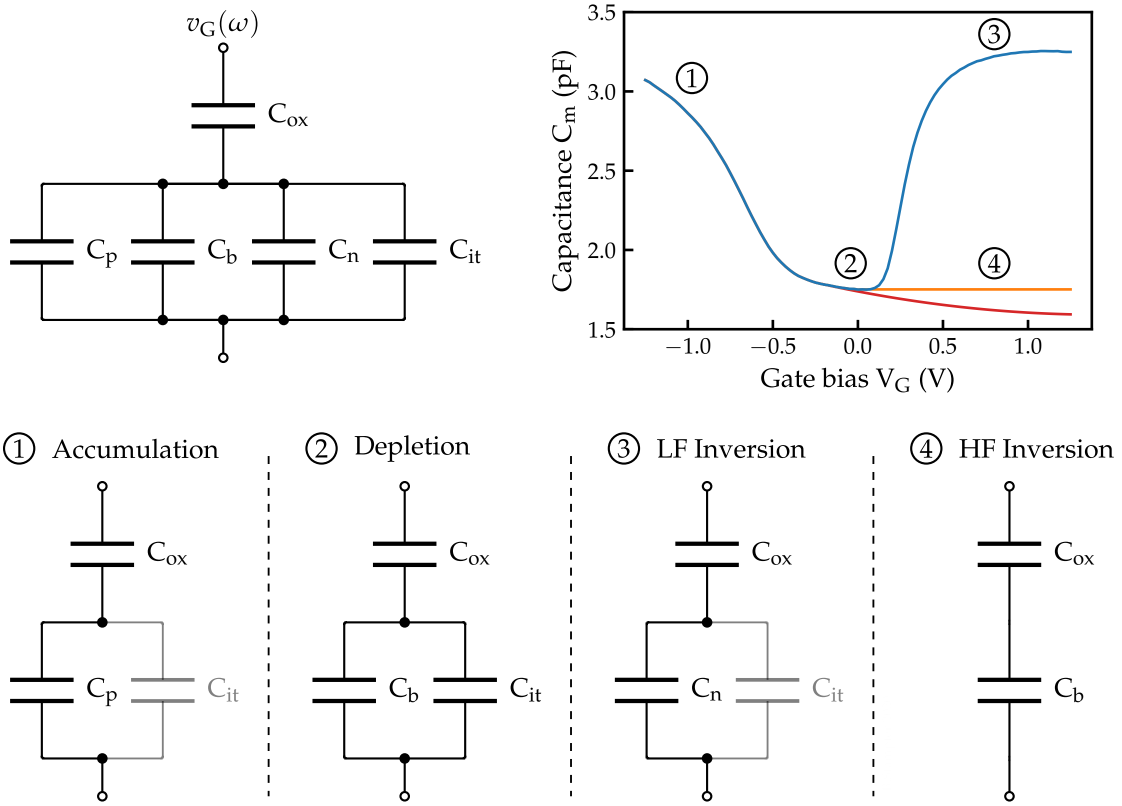


Figure 5.9. Equivalent circuits for CV measurements in the different regimes. **(a)** General equivalent circuit. **(b)** Schematic CV-measurments for varying measurement conditions. **(1–4)** Accumulation, depletion, low-frequency inversion and high-frequency inversion equivalent circuits. To obtain the low- and high-frequency curves (blue and orange), the DC sweep rate has to be slow enough for the minority carriers to follow, otherwise the deep depletion curve (red) is obtained. If in addition the AC frequency is low enough for the minority carriers to respond, the low-frequency curve is obtained.

these charges will be compensated by the width of the space charge region. Since the DC part will be compensated by the inversion layer, the small signal capacitance will stay essentially constant. Only if the AC signal is slow enough for the inversion layer charges to follow, a low-frequency curve will be obtained. This is usually the case for MOSFETs, if the source/drain contacts are connected to ground, as minority carriers can be quickly supplied from there, but usually not for MOSCAPs, as very low measurement frequencies are needed to allow sufficient time for minority carrier generation in the bulk. The capacitance in the low-frequency regime approaches C_{ox}, the situation is analogous to that in accumulation.

Low frequency method

Low frequency here refers to measurements where both inversion charges and interface traps can respond. As shown in Figure 5.9, the low frequency capacitance can be written as

$$C_{\mathrm{lf}} = \left(\frac{1}{C_{\mathrm{ox}}} + \frac{1}{C_{\mathrm{s}} + C_{\mathrm{it}}}\right)^{-1}, \tag{5.41}$$

with $C_{\mathrm{s}} = C_{\mathrm{b}} + C_{\mathrm{n}}$ the semiconductor surface capacitance in depletion and inversion. From this, the interface trap density can be calculated with $D_{\mathrm{it}} = C_{\mathrm{it}}/qA$ as

$$D_{\mathrm{it}} = \frac{1}{qA}\left(\frac{C_{\mathrm{ox}}C_{\mathrm{lf}}}{C_{\mathrm{ox}} - C_{\mathrm{lf}}} - C_{\mathrm{s}}\right). \tag{5.42}$$

The semiconductor capacitance C_{s} necessary to calculate this can be obtained by either TCAD simulation or analytically [105] with the surface potential required for this calculation obtained using Berglund's method [102]:

$$\phi_{\mathrm{s}} = \int_{V_{\mathrm{G}_1}}^{V_{\mathrm{G}_2}} (1 - C_{\mathrm{lf}}/C_{\mathrm{ox}})\,\mathrm{d}V_{\mathrm{G}} + \Delta, \tag{5.43}$$

with Δ the surface potential at $V_{\mathrm{G}} = V_{\mathrm{G}_1}$. The difficulty in this method lies in obtaining the correct semiconductor capacitance, as even small deviations from its real value may lead to relatively large spurious defect densities. If the prime interest of a study lies in finding the changes to the defect density, e.g. due to electrical stress, a variant of this method can be used wherein an initially recorded low frequency curve is used as a reference, e.g.:

$$\Delta D_{\mathrm{it}} = D_{\mathrm{it}} - D_{\mathrm{it,ref}} = \frac{1}{qA}\left(\frac{C_{\mathrm{ox}}C_{\mathrm{lf}}}{C_{\mathrm{ox}} - C_{\mathrm{lf}}} - \frac{C_{\mathrm{ox}}C_{\mathrm{lf,ref}}}{C_{\mathrm{ox}} - C_{\mathrm{lf,ref}}}\right). \tag{5.44}$$

This method has been used by the author in [BSC3] to measure the temporal evolution of near-interface defects under BTI stress to study the permanent component of the degradation using the hydrogen-release model (see Section 3.1.5, [BSC4])

High-low frequency method

Another way of obtaining C_s required for Equation (5.42) is to measure a high-frequency curve at an AC frequency too fast for the interface traps to respond. The interface trap density can then be expressed as:

$$D_{\mathrm{it}} = \frac{1}{qA}\left(\frac{C_{\mathrm{ox}}C_{\mathrm{lf}}}{C_{\mathrm{ox}} - C_{\mathrm{lf}}} - \frac{C_{\mathrm{ox}}C_{\mathrm{hf}}}{C_{\mathrm{ox}} - C_{\mathrm{hf}}}\right). \tag{5.45}$$

This method is limited in range to the onset of inversion [105].

Terman method

A high frequency curve, as measured for the high-low frequency method above still contains information about interface defects. While the defects will not respond to the AC frequency, they will be charged during the DC sweep and cause a stretch out of the curve along the gate voltage axis, as additional charge has to be supplied to the gate in order to reach the same surface potential as in an ideal device.

By comparing the obtained capacitances C_{hf} with the ones from a theoretical curve, the relation between $\Delta V_{\mathrm{G}} = V_{\mathrm{G}} - V_{\mathrm{G,ideal}}$ and ϕ_{s} can be obtained. The interface trap capacitance can then be calculated as [105]

$$D_{\mathrm{it}} = \frac{C_{\mathrm{ox}}}{qA}\left(\frac{dV_{\mathrm{G}}}{d\phi_{\mathrm{s}}}\right) - \frac{C_{\mathrm{s}}}{qA} = \frac{C_{\mathrm{ox}}}{qA}\frac{d\Delta V_{\mathrm{G}}}{d\phi_{\mathrm{s}}}. \tag{5.46}$$

Gray-Brown method

The Gray-Brown method [136] can be used to estimate the interface trap density close to the majority carrier band edge. For this, high-frequency CV curves are measured at various temperatures typically from room temperature to below 100K. The change in temperature causes a shift of the Fermi level between the measurements which in turn changes the interface trap occupancy. By comparing the corresponding flat band voltages, the interface trap capacitance can be extracted in a manner similar to the Terman method. To obtain true high frequency CV curves near flatband, measurement frequencies of more than 200 MHz [4] or even GHz [137] may be required. Lower frequency measurements may be used to obtain qualitative results.

5.2.2 Conductance Method

Unlike the capacitance methods, the conductance method developed by Nicollian and Goetzberger [138, 139] measures the loss occurring from (dis)charging the interface traps. Delayed responses to the AC voltage from interface traps with time constants in the frequency range of the voltage causes a real-valued current. By measuring at varying frequency and gate bias, this loss can be characterized and linked to interface trap density. Figure 5.10a shows the equivalent circuit for the situation with interface trap loss. This is commonly replaced with the circuit from Figure 5.10b, as it allows to write the equations in a symmetric form. For a single interface trap with $\tau_{\mathrm{it}} = R_{\mathrm{it}}C_{\mathrm{it}}$ this yields

$$C_{\mathrm{p}} - C_{\mathrm{b}} = \frac{C_{\mathrm{it}}}{1+(\omega\tau_{\mathrm{it}})^2} \tag{5.47}$$

$$\frac{G_{\mathrm{p}}}{\omega} = \frac{qA\omega\tau_{\mathrm{it}}D_{\mathrm{it}}}{1+(\omega\tau_{\mathrm{it}})^2}, \tag{5.48}$$

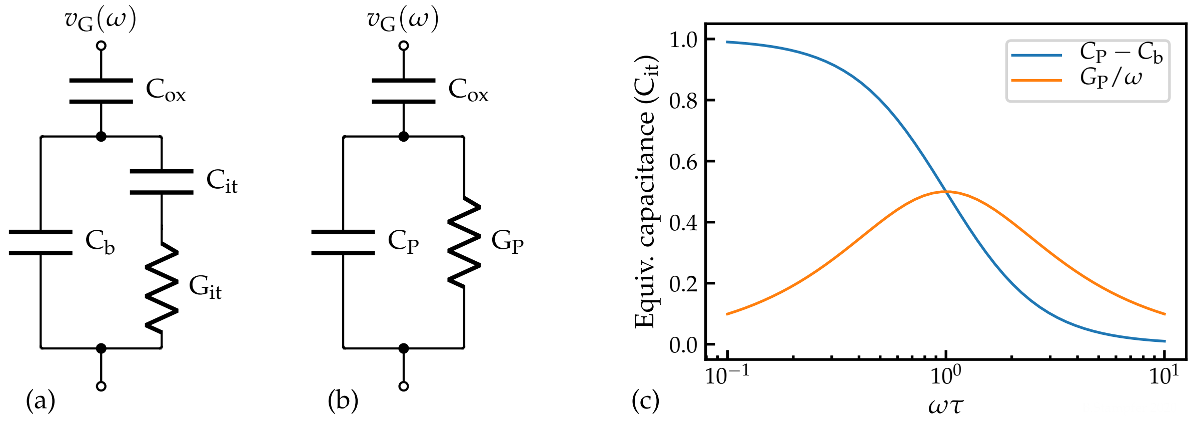


Figure 5.10. MOSCAP in depletion between low and high frequency regimes for a single interface defect. **(a)** Equivalent circuit with lossy interface defect represented by a series RC circuit. **(b)** Simplified circuit with equivalent parallel capacitance and conductance. **(c)** Behavior of C_P and G_P with varying measurement frequency. At low frequencies, $C_P = C_b || C_{it}$ and at high frequencies $C_P = C_b$, in both cases $G_P/\omega \approx 0$. At the corner frequency $\omega\tau = 1$, however, G_P/ω peaks at C_{it}. This behavior is exploited in the conductance method [4].

plotted in Figure 5.10c. In measurements, many interface traps will be present, with their trap levels distributed in energy. With the assumption of traps distributed continuously throughout the band gap, the conductance can be written as [139]

$$\frac{G_p}{\omega} = \frac{qAD_{it}}{2\omega\tau_{it}} \ln\left[1 + (\omega\tau_{it})^2\right]. \tag{5.49}$$

The maximum of this expression is found at $\omega \approx 2/\tau_{it}$, resulting in

$$D_{it} \approx \frac{2.5}{qA}\left(\frac{G_p}{\omega}\right)_{max}. \tag{5.50}$$

The measurement thus works by recording capacitance curves at multiple frequencies to determine G_p/ω for each voltage. The method is typically used from flatband to weak inversion [105]. To simplify the calculation, G_p/ω can also be obtained from the equivalent parallel capacitances and conductances provided by many instruments using:

$$\frac{G_p}{\omega} = \frac{\omega G_m C_{ox}^2}{G_m^2 + \omega^2 (C_{ox} - C_m)^2}. \tag{5.51}$$

The conductance method is generally considered more sensitive than the capacitance methods, but requires measurements over a wide range of frequencies to yield results.

CHAPTER 6

MEASUREMENTS AND RESULTS

In this chapter results obtained using measurement, characterization, and simulation methods as discussed in the previous chapters will be presented. First, in Section 6.1, statistical analysis of single defect noise measurements on Si/SiO$_2$ FETs will be discussed. Following this, in Section 6.2, characterization of BTI and RTN with the defect centric approach and performed on integrated Si/SiO$_2$ pMOS arrays is shown. Finally, in Section 6.3, results of single defect characterizations on novel few-layer MoS$_2$ channel devices are presented.

6.1 Statistical Characterization of Defects Causing RTN in SiO$_2$ Transistors

While the number of oxide defects affecting a MOSFET scales with the gate area, and thus decreases for smaller devices, the defects' individual impact on the threshold voltage follows the opposite trend. In particular, the decreased channel area in scaled devices leads to a higher average impact of each defect on the device threshold voltage. This results in similar total threshold voltage shifts observed on small and large gate

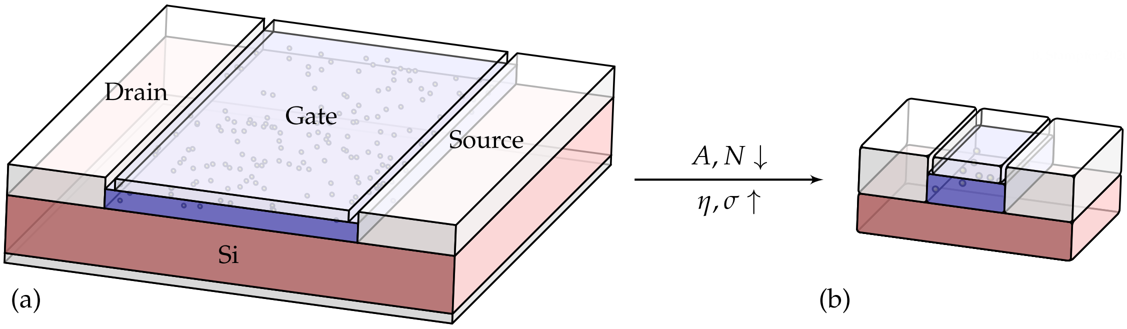


Figure 6.1. Illustration of large area **(a)** and small area **(b)** devices with oxide defects randomly distributed in the SiO$_2$ layer. Devices with a smaller gate area have on average a lower number of defects, but the average impact of the defects is larger. This leads to a similar mean degradation for both types of devices, but a much larger variance among the small devices. Originally published in [BSJ2]

area devices. However, a significant difference between normal and scaled nodes is that the smaller number of defects leads to increased variability among nominally identical devices. This can affect the reliability of integrated circuits using scaled devices and requires statistical characterization of such devices to understand the reliability and variability of a technology and its characteristics under various operating conditions, i.e. biases and temperatures. For this, noise has been studied on more than 300 devices to extract statistical distributions of the defects' impact on the threshold voltage, their location in the oxide, and their trap level.

6.1.1 Devices and Measurements

The devices which have been used in this study are planar Si/SiO$_2$ MOSFETs with a gate area of $\approx 0.15\,\mu\text{m}^2$. For the characterization of the devices a fully shielded Cascade/FormFactor PA300 semi-automatic wafer prober has been used. The electrical measurements have been performed using a dedicated custom designed defect probing instrument (DPI) [39]. The DPI has been equipped with three voltage sources, one for drain, gate, and bulk, and a low noise transimpedance amplifier stage in combination with a high-resolution voltage sampling unit for the source current. All connections have been made with double-shielded TRIAX cables to reduce spurious noise and ensure a high SNR.

To obtain statistics on the step height distribution, a fixed characterization sequence has been applied to each device tested. The sequence started with an initial $I_\text{D}(V_\text{G})$ characterization, which has been used to verify proper operation of the DUT, to determine the gate voltages at which RTN is measured, and to map the recorded drain-source currents to an equivalent threshold voltage shift. The initial $I_\text{D}(V_\text{G})$ measurements are shown in Figure 6.2 together with the bias range which has been used for recording the RTN signals. After the initial $I_\text{D}(V_\text{G})$ curve, the sequence continued with RTN traces recorded at the gate voltages corresponding to $I_\text{D} = 20$, 50, 100, 300, and 1000 nA at a drain bias of $V_\text{D} = -100\,\text{mV}$. At each gate voltage, five short RTN traces have been recorded with a sampling time of $T_\text{s} = 100\,\mu\text{s}$ and a length of $t_\text{r} = 10\,\text{s}$, followed by a longer trace with a sampling time of $T_\text{s} = 10\,\text{ms}$ and a length of $t_\text{r} = 1\,\text{ks}$. The sequence has been chosen to characterize defects with a wide range of charge transition times. Figure 6.3 shows an example of a trace which has been recorded with $T_\text{s} = 100\,\mu\text{s}$.

For the statistics on defect location and trap level, the initial measurements with the fixed characterization sequence have been followed by additional RTN measurements on specific devices to characterize individual defects observed in the initial results in more detail. The measurement parameters for these recordings have been tailored to the individual defect to match their transition times and the gate bias range where they have been most active.

TU Bibliothek WIEN Your knowledge hub
Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

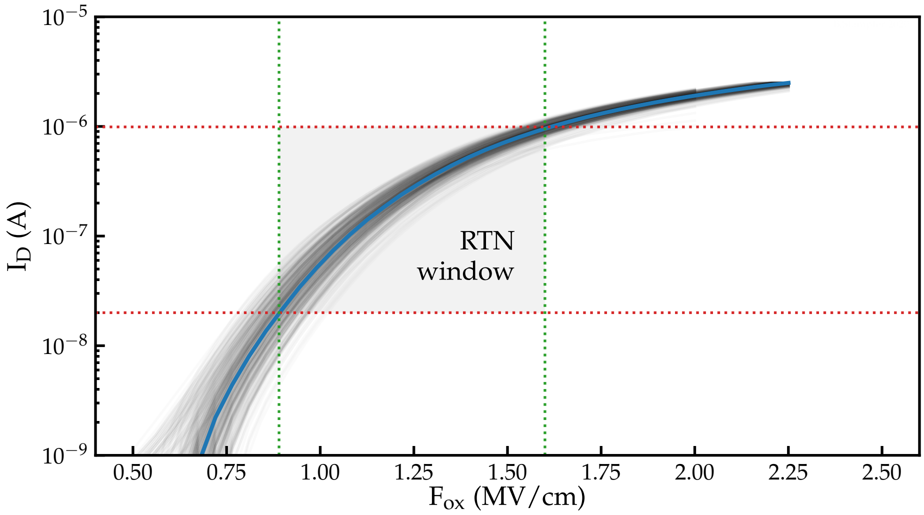


Figure 6.2. Transfer characteristics of the ≈300 devices measured. Shown are the individual transfer characteristics in gray and the average in blue. The bias and current range in which RTN has been measured is highlighted. Originally published in [BSJ2]

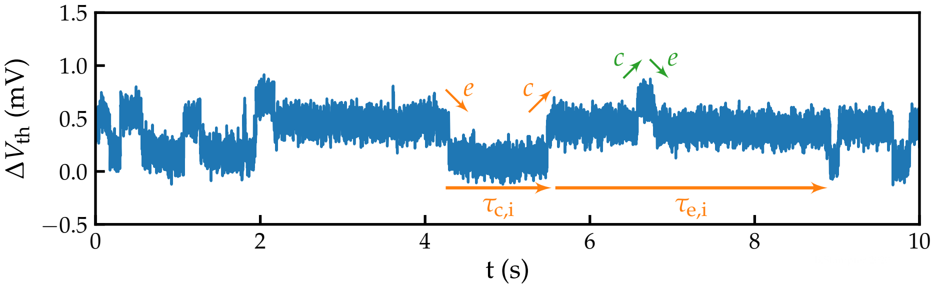


Figure 6.3. An example of a short RTN trace. Even though this is the faster kind of the measurements performed, the noise level at around 0.15 mV is low enough to clearly see the trapping of two defects in the signal. Originally published in [BSJ2]

6.1.2 Parameter Extraction

From the measured RTN data, the noise parameters of the defects producing the signals, i.e. the step heights and the capture/emission times of the discrete charge transitions, have been extracted. This has been done using the Canny edge detection algorithm, as outlined in Section 5.1.2. The Canny method could be chosen because there were only few active defects per device, and the SNR was sufficiently high. The significant advantage of this method is the high level of autonomy it provides. A long RTN trace analyzed with this method is shown in Figure 6.4.

6.1.3 Step Heights

To evaluate the impact of defects on the threshold voltage of the devices, the complementary cumulative density function (CCDF) of the step heights can be analyzed. The CCDF (or 1-CDF) which has been obtained from the measurements is shown in Figure 6.5. The distribution seems to be a bi-modal exponential distribution, with one

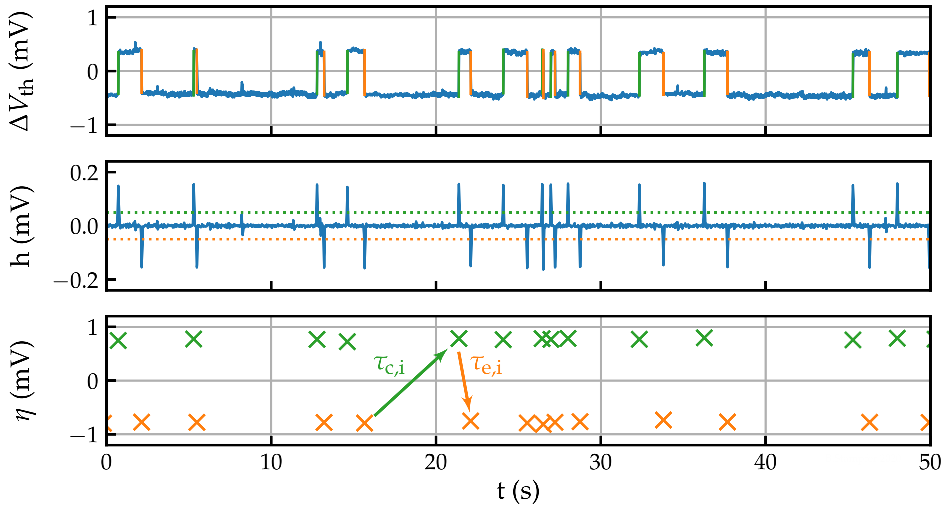


Figure 6.4. Example of the Canny edge detector used on one of the long (1 ks) RTN traces. The original ΔV_{th} signal has been convoluted with the first derivative of a Gaussian pulse to yield a signal h which has peaks at the positions of the steps. Local maxima above a selected threshold give the positions of the steps, their magnitude η has been taken from the original signal. Originally published in [BSJ2]

set of defects exhibiting an average step height of 0.39 mV and a second set of defects showing an average step height of 1.09 mV. This observation is interesting, as for a simple gate stack with only a single oxide, i.e. without a second high-k material layer, only unimodal exponential distributions have been reported in literature so far, which would appear as a straight line in the plot.

6.1.4 Energy and Position

To characterize the trap levels and vertical positions of the defects, more detailed measurements have been performed to extract the charge transitions times for a number of defects as shown in Figure 6.6. This has been done for around 100 defects, and the estimation formulas given in Section 5.1.4 have been used to obtain the distribution of these parameters as shown in Figure 6.7. The mean distance from the interface of the observed defects seems to be around 0.6 t_{ox} and their mean trap level lies 0.4 eV above the Fermi level at the interface, which is close to the Si conduction band at measurement conditions. It should be noted that the depth distribution of the probed defects does not represent the actual distribution of defects in the oxide, as only defects whose trap level is close to the Fermi level in the bias range chosen for the measurements can be characterized. As can be further seen in Figure 6.7, the estimation suggests some defects to be positioned outside the oxide. This is a consequence of the approximations made for this estimation. In particular:

- Gate/defect interaction is neglected. Defects interacting primarily with the gate thus may have inverted capture and emission time behavior which leads to negative distances.

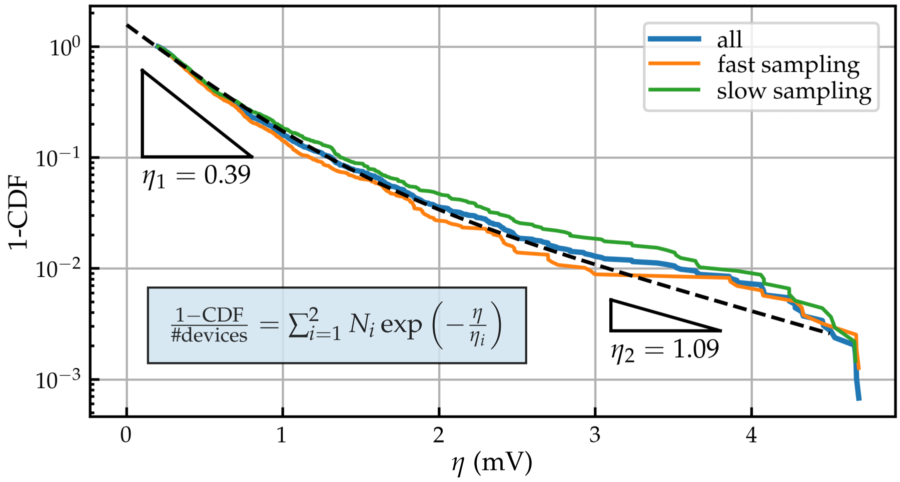


Figure 6.5. Complimentary cumulative density function of the step heights observed in the measurements. The distribution seems to be composed of two separate exponential distributions with mean values of 0.39 mV and 1.09 mV. Originally published in [BSJ2]

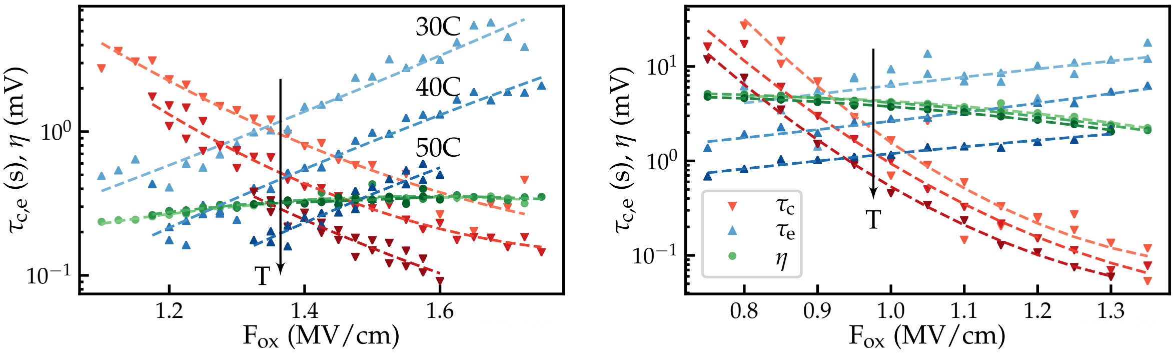


Figure 6.6. Examples of the capture and emission time dependence on gate bias and temperature, for two defects which have been characterized in more detail. Originally published in [BSJ2]

- The prefactors in Equation (5.30) are assumed constant in the estimation, but do change with the channel carrier density, leading to overestimation of the distances.
- The estimation is only valid for two state defects, defects which can not be described using a two state model yield invalid results.

Also, charge transition times of defects have been measured whose intersection point is not visible within the measurement range. The extrapolation of the capture and emission times necessary in this case introduces additional uncertainty and may lead to some defects which have estimated positions close to the interface.

In Figure 6.8, the positions of the extracted defects are shown in a band diagram obtained from TCAD simulation. In addition, defect bands for electron and hole traps in SiO_2 reported in literature [49] are shown for comparison. The peak of the energy extracted from the measurements is 0.1 to 0.2 eV lower than the literature values.

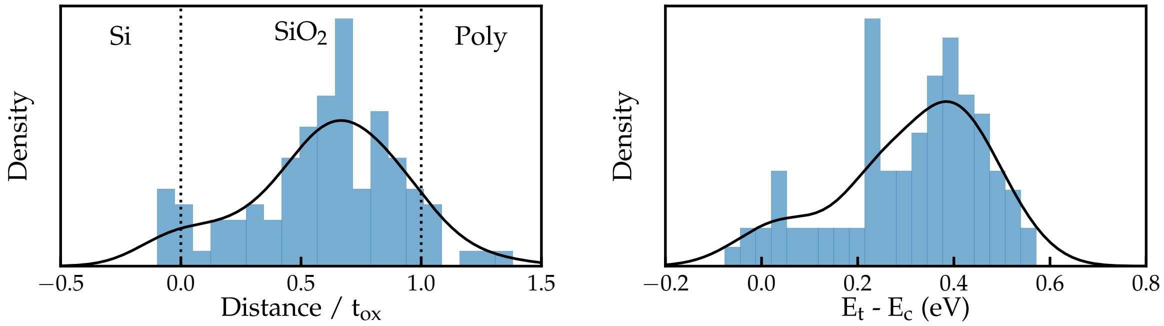


Figure 6.7. Distribution of the extracted vertical positions, measured from the interface, and trap levels for around 100 defects. The extracted energy peaks at around 0.4 eV above the Fermi level, which is close to the conduction band edge during the measurements. The distribution of the depths of the defects in the oxide shows a maximum at 0.6 t_{ox}, which is where the effective trap levels of the defects coincide with the Fermi level. Originally published in [BSJ2]

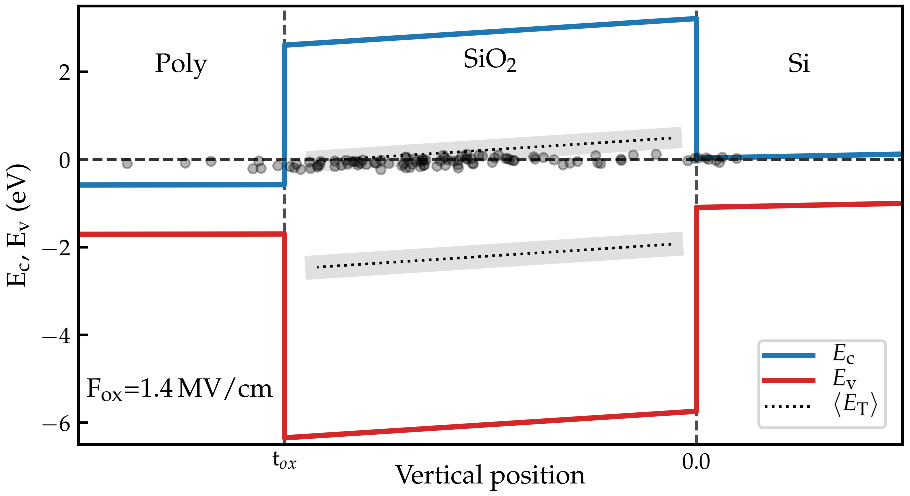


Figure 6.8. Simulated band diagram showing the defects extracted using the estimations for depth and trap level. In addition, defect bands for SiO$_2$ from [49] are shown in gray. Originally published in [BSJ2]

To get more accurate estimations for the trap depth and level, the charge transition times of the single defects have been simulated with *Comphy* [49]. By doing so, the trapping behavior of the single defects has been modeled using a two-state NMP model as outlined in Figure 3.4, together with a calibrated model for the device electrostatics. With this, the parameters for a subset of the defects have been obtained. The resulting distributions for trap level and distance to the interface are shown in Figure 6.9.

As can be seen, the peaks of the distributions for both the distance to the interface and the trap levels are slightly lower in the simulation results. The distribution of trap levels peaks at around 0.85 eV relative to Si midgap. The distribution which has been obtained from measurements on large devices in [49] is also shown in Figure 6.9. The mean energy which has been obtained in this work is lower than the one on large

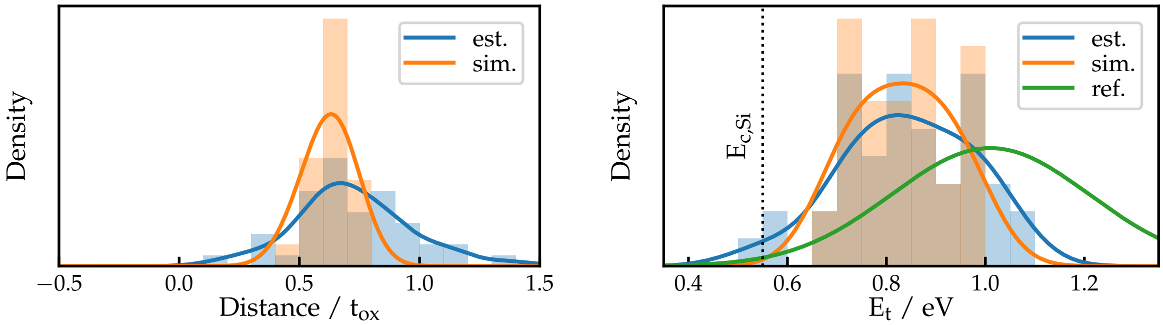


Figure 6.9. Distribution of the vertical positions, measured from the interface, and trap levels, referenced to the Si midgap, extracted using both the estimation approach and TCAD simulations. The simulation results show a narrower distribution in position and indicate that the defects are slightly closer to the interface compared to the estimations. The extracted trap levels are slightly lower as well, peaking ≈0.3 eV above the conduction band edge. The distribution of trap levels which have been measured covers mainly the lower half of the distribution obtained in [49]. Originally published in [BSJ2]

devices, which is at $E_t = 1.01\,\mathrm{eV}$. The energy distribution is narrower and seems to coincide with the lower half of the energy distribution found for the large devices.

6.1.5 Conclusions

In conclusion, we have characterized defects in a Si/SiO$_2$ technology and observed a bi-modal exponential distribution of step heights, which has not been reported so far for a device with a gate stack consisting of a single insulating layer, but is commonly observed on high-k devices [83, 140, 69]. The charge sheet approximation, as outlined in Section 3.2.2 gives values below 0.2 mV for this geometry, which—within the limits of this approximation—agrees with the first branch of the CCDF distribution measured. The origin of the second branch of the distribution indicating larger η, however, remains open at this point and requires further investigation. For the trap depth and trap level distributions, the results obtained from the estimations show good agreement with the simulation results. Compared to the results given in [49], we have only observed defects in the lower half of the energy distribution. This may be due to limits of our measurement range, but it could also indicate that their distribution is broader than the actual distribution of trap levels.

6.2 Defect Centric Evaluation of RTN and BTI using pMOS Arrays

As outlined in the previous section, characterizing oxide defects of scaled technologies requires testing many devices to obtain statistically significant results. In the previous section, this has been done by contacting and probing many individual devices

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN

using a semi-automatic probe station. However, handling such a system can be very error-prone as the positions of the needles can drift when many devices are contacted. This requires frequent manual readjustment, and thus makes the measurement of several hundreds of devices time-consuming and at some point unpractical. Here, a special array chip, purpose-built for this task, has been used instead. This approach allows for reliable and completely unattended measurements.

In this study, defects in high-k pMOS devices have been statistically characterized using a defect-centric approach. This has been done using array structures which contain thousands of nominally identical devices per device geometry. The defect-centric approach, as outlined in Section 3.3, has then been used to describe the statistical distribution of ΔV_{th} after stress over the full set of measured transistors. From this, the average number of defects remaining charged after stress, the average number of defects exhibiting RTN and the average step height of the defects could be extracted. The degradation measured on the devices has further been replicated using TCAD simulations to allow extrapolation of the degradation for lifetime estimation. Finally, the influence of bulk- and drain-bias applied during the stress phase has been investigated too.

6.2.1 Array Chip

The measurements have been performed on a custom designed chip which contains a matrix of MOSFETs of different geometries, as well as on-chip logic to select the individual devices. The selection logic is based on shift registers, which control double transmission gates on the gate and drain lines. The electrical layout of the matrix is shown in Figure 6.10. As can be seen, the bulk and source connections are shared among all transistors, while the gate and drain lines are switched for the individual rows and columns of the matrix to select a single device for measurement. Rows and columns not selected are commonly connected and have been supplied by a gate-off and drain-off bias to assure a defined potential at their terminals. Additional information on the array chips can be found in [141].

This integrated device matrix allows to efficiently characterize many devices without user interaction. However, the matrix layout has the drawback that the leakage currents of all devices on the same drain line will contribute to the measured current. But this does not affect the extracted threshold voltage shifts, as these currents will contribute to both the initial $I_{\mathrm{D}}(V_{\mathrm{G}})$ characteristics and the measured drain current traces, and thus cancel out during mapping of the device current to ΔV_{th}.

The array structures have been produced in a commercial, high-k/metal gate, planar technology. For each of the geometries available on the arrays, around 3000 devices can be addressed. For this study, pMOS transistors with gate widths of 100 nm and lengths of 30 nm and 150 nm have been used, henceforth called short and long devices, respectively.

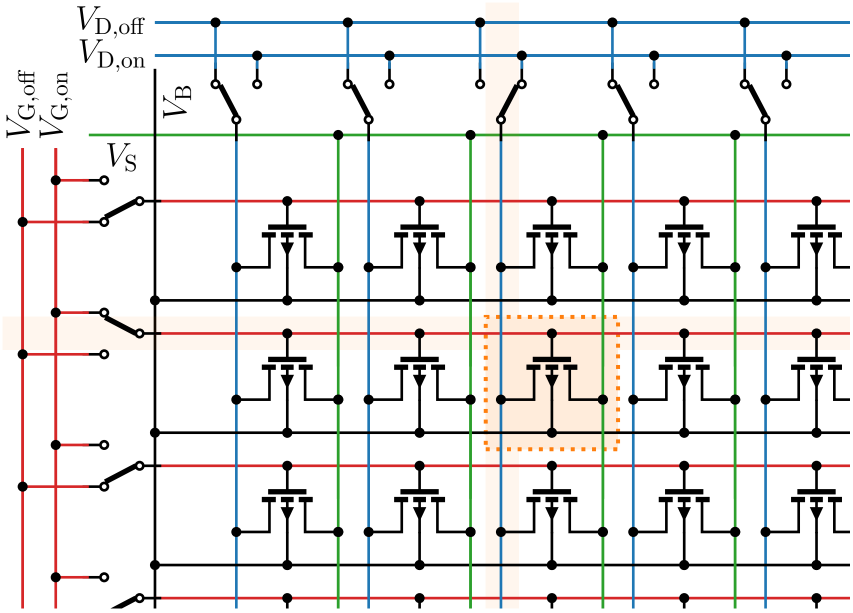


Figure 6.10. Layout of the signal lines of the array which have been used for defect characterization. The gate terminals of the transistors in each row can be switched between externally supplied on- or off-biases using on-chip logic. Likewise, the drain terminals can be switched for each transistor column. This allows to address and thus to characterize each individual device in the array. The bulk and source terminals are common for all devices. More details about the array structures can be found in [141]. Originally published in [BSJ3]

6.2.2 Measurements

All measurements have been performed using the same measurement sequence which has been repeated for each device characterized. The sequence consists of:

1. an initial $I_{\mathrm{D}}(V_{\mathrm{G}})$ curve, used to determine the gate bias during the relaxation phases, and
2. eMSM measurements—as discussed in Section 4.1.3—with stress phases of $t_{\mathrm{s}} = \{2, 10, 100, 1000, 10000\}\mathrm{ms}$ followed by relaxation phases of $t_{\mathrm{r,max}} = 1\,\mathrm{s}$.

Recording of the recovery traces has been started 100 µs after stress release, with the sampling rate varied to achieve 200 samples per decade in time. In total 39 such sets of measurements have been recorded for the study, which resulted in more than half a million recovery traces. In addition to the two geometries, various combinations of gate, drain, and bulk stress biases have been investigated using these sets. All measurements have been performed at 35 °C using a custom-built defect probing instrument [39]. In addition to the voltage sources for gate, drain and bulk, the transimpedance amplifier and the sampling unit for the source current as used in the previous section, the instrument was equipped with two additional voltage sources for the drain-off and gate-off biases, and electrically isolated IO-ports to control the on-chip logic of the arrays.

The gate-off and drain-off lines have been supplied with 0.15 V and 0.0 V, respectively. One set of initial $I_{\mathrm{D}}(V_{\mathrm{G}})$ curves which have been recorded at the beginning of the measurement sequences is shown in Figure 6.11 for the long devices. As can be seen, a significant variation of the threshold voltage of up to ≈30 mV compared to the average value can be observed at $I_D = 10\,\mathrm{nA}$. After the measurements, these $I_D(V_G)$ curves have been used to map the drain current recorded during the relaxation phases of the eMSM measurements to ΔV_{th}. An exemplary set of $\Delta V_{th}(t_r)$ as has been mapped from the $I_D(t_r)$ recovery traces is shown in Figure 6.12. From these recovery curves, distributions of ΔV_{th} have been drawn for the parameter extraction.

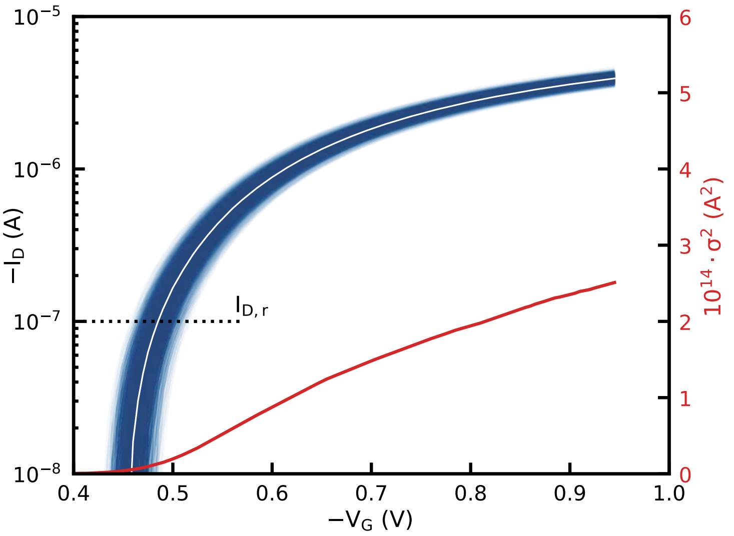


Figure 6.11. A set of $I_D(V_G)$ curves recorded on the long devices (shown in blue). In addition, the mean and variance are given in white and red. From the curves for each device, the gate bias during relaxation has been determined based on a chosen relaxation current $I_{D,r} = 100\,\mathrm{nA}$. Originally published in [BSJ3]

To check the function of the array and the device selection logic, and whether degradation is homogeneous over the area of the array, the degradation extracted $t_r = 1\,\mathrm{ms}$ after $t_s = 10\,\mathrm{s}$ of stress at $V_G = -1.45\,\mathrm{V}$—the most severe stress condition applied—has been plotted over the array in Figure 6.13 (shown for the short geometry). From the top plot, all rows and columns of the matrix can be observed to work properly, and no defective areas can be seen. The bottom plot shows the mean degradation for larger segments of the array. Again, no unexpected behavior can be observed, and degradation appears homogeneous over the entire area of the array.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN

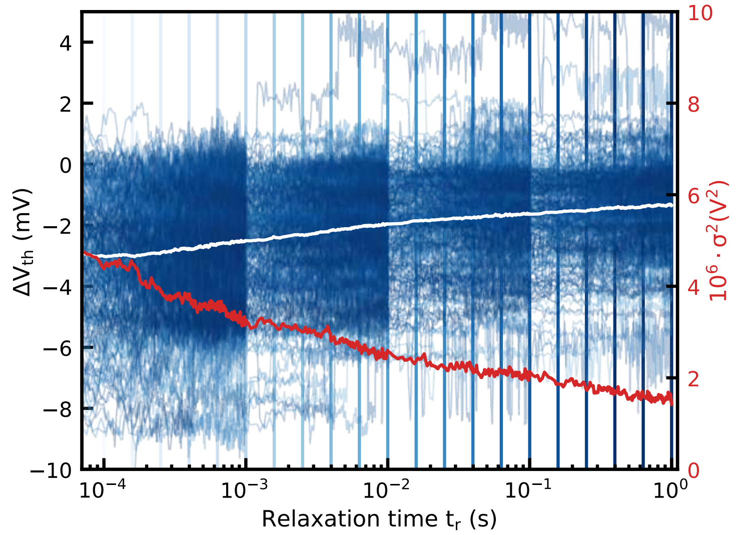


Figure 6.12. A set of $\Delta V_{\mathrm{th}}(t_{\mathrm{r}})$ mapped from $I_{\mathrm{D}}(t_{\mathrm{r}})$ using the corresponding set of $I_{\mathrm{D}}(V_{\mathrm{G}})$ curves (shown in blue). The mean and variance are given in white and red, respectively. The vertical lines indicate moments in time when the distributions of ΔV_{th} have been drawn for further analysis. Originally published in [BSJ3]

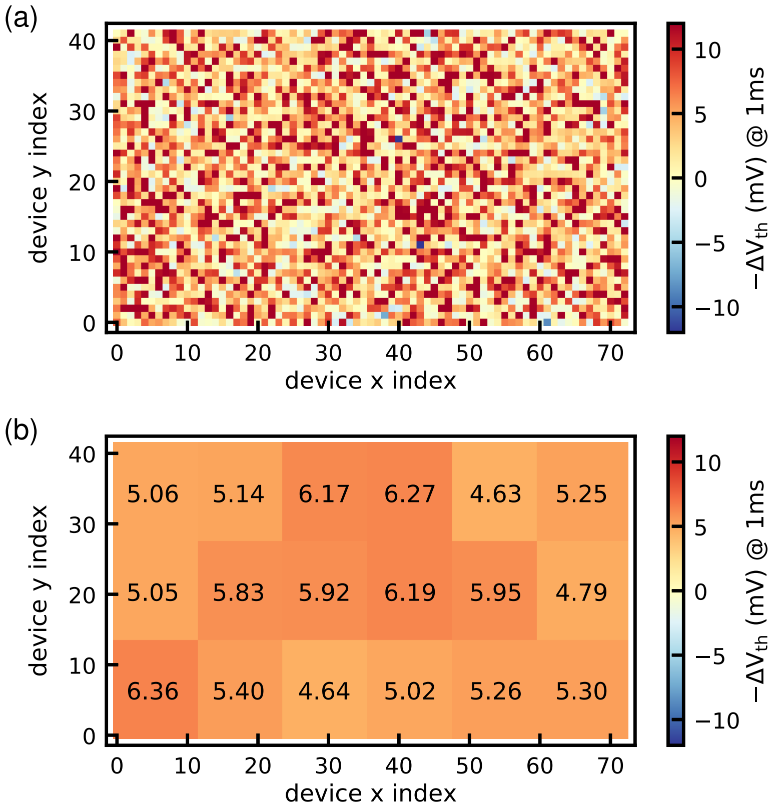


Figure 6.13. Degradation in ΔV_{th} recorded 1 ms after 10 s of stress at $V_{\mathrm{G,str}} = -1.45\,\mathrm{V}$. **Top:** for each device as positioned in the matrix of short devices. **Bottom:** averaged over a number of devices. The plots show neither defective rows nor columns, nor clusters or overall inhomogeneities of the extracted degradation of the threshold voltage. Originally published in [BSJ3]

6.2.3 Parameter Extraction

The large number of recorded traces prohibits the usage of traditional methods for defect parameter extraction, like the ones outlined in Section 5.1. Instead, a statistical approach based on the defect centric model, as described in Section 3.3, has been used. From the measured threshold voltage shifts during relaxation, CDFs can be drawn for any moment in time. These CDFs can be reproduced using the defect-centric model with the equations given in Section 3.3. Examples of measured and calculated CDFs for one measurement set are shown in Figure 6.14. As can be seen, the model can accurately reproduce the measured distributions.

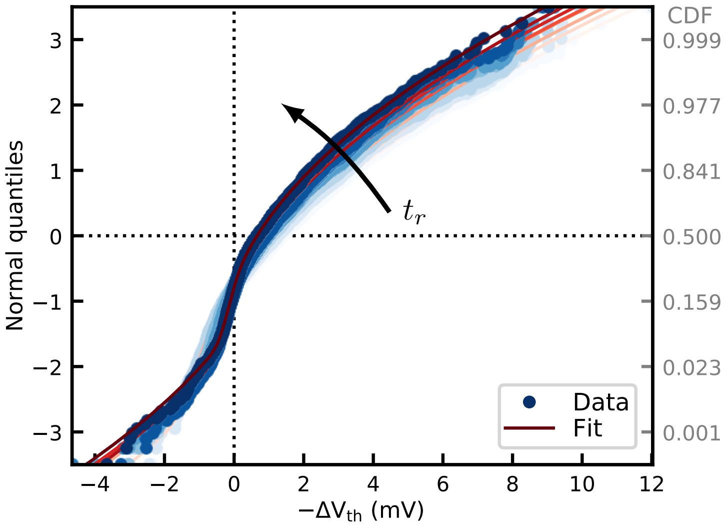


Figure 6.14. CDFs of ΔV_{th} during recovery. Blue: measured, Red: calculated. Most devices exhibit positive degradation ($-\Delta V_{th} > 0$ for pMOS) after stress due to BTI, while for some devices RTN causes the reverse $-\Delta V_{th} < 0$. Originally published in [BSJ3]

From the model, the average number of charged defects N, the average number of active RTN defects N_{RTN}, the average step height of a defect η, and the parameters of the Gaussian measurement noise have been obtained. An overview of extracted parameters for the short devices is shown in Figure 6.15. As expected, N_{RTN} and η remain constant throughout the measurement sets. The number of charged defects on the other hand strongly depends on the stress time and the applied gate and bulk biases. The dependence on gate bias and stress time is shown in Figure 6.16. As can be seen, the number of defects charged during stress shows a strong increase both with gate bias and stress time, as expected for BTI. The average number of charged defects is low, with only around two defects per device for the most severe stress condition and down to 0.03 defects per device, i.e. one defect every 30 devices, for the weakest stress conditions.

A similar plot, which illustrates the relaxation behavior is shown in Figure 6.17. A relatively uniform recovery behavior can be observed along the logarithmic time axis.

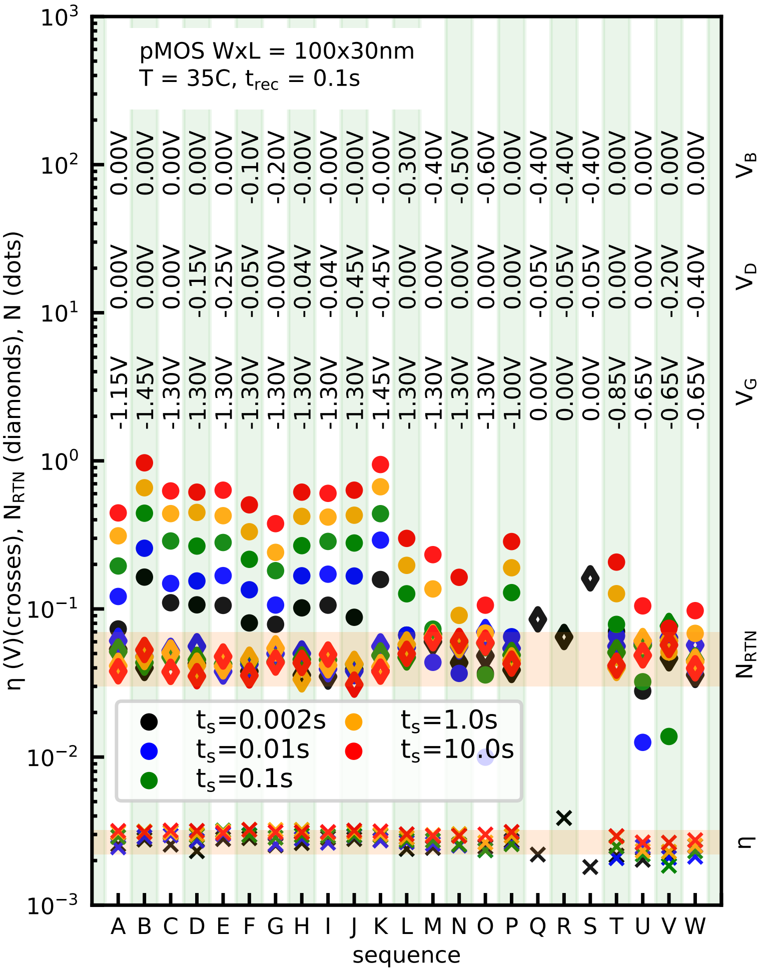


Figure 6.15. Extracted parameters obtained from the short devices after $t_r = 100\,\mathrm{ms}$ of recovery. The average step height η (×), as well as the average number of RTN charges N_{RTN} (◇) remains constant over all sets. The average number of charges captured due to BTI N (●) depends on stress bias and stress time as expected. Originally published in [BSJ3]

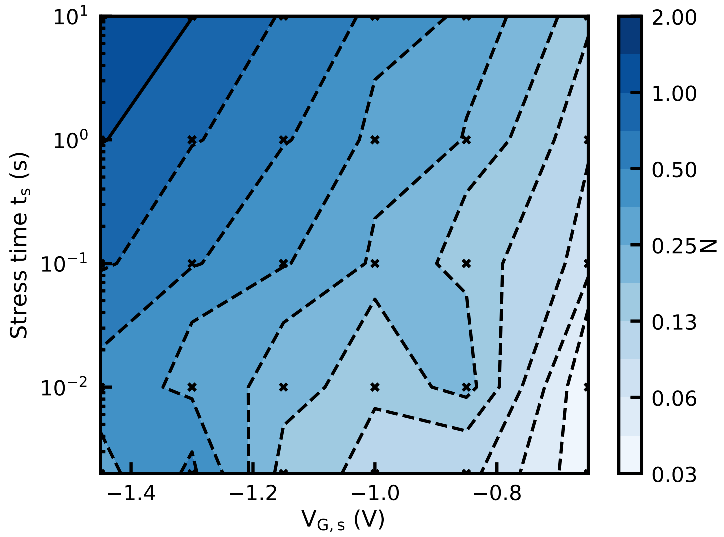


Figure 6.16. Average number of captured defects N (shade) extracted 2 ms after stress release, over stress time and gate bias. The crosses indicate the measurement points, the dashed lines are contour lines separating the iso-surfaces. For weak stress conditions, only a fraction of the devices exhibits a charged defect after stress. Originally published in [BSJ3]

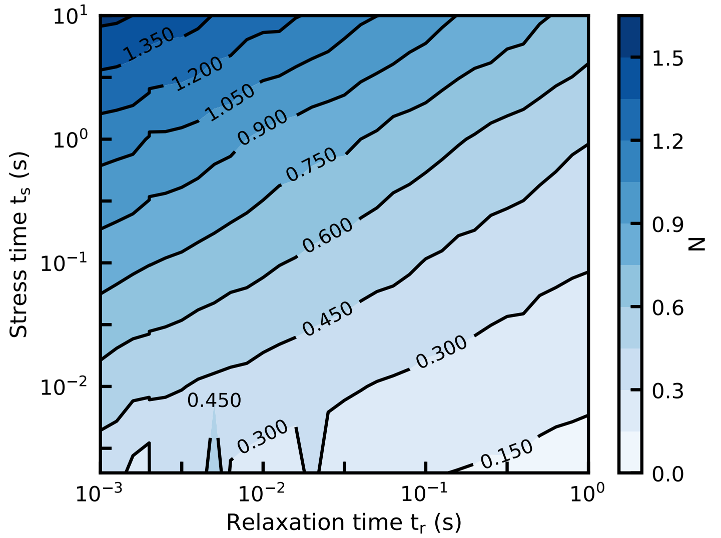


Figure 6.17. Average number of captured defects N (shade) extracted after stress at $V_G = -1.45\,\text{V}$, over stress and relaxation time. The lines are contour lines separating iso-surfaces for the number of charged defects. Only for short stress times, full recovery can be observed in a relatively short relaxation time window. Originally published in [BSJ3]

6.2.4 Simulation and Extrapolation

To explain the device threshold degradation behavior and accurately identify the contributions of the many defects, the mean degradation measured at various gate stress biases has been replicated using TCAD simulation. For this, the open-source TCAD simulator *Comphy* [49] has been used. To calculate the behavior of the defects, this simulator employs an effective two state NMP model to mimic the charge transition kinetics of the defects, and their bias and temperature dependence. More details about this model can be found in Section 3.1.3. The defect bands used in the simulations have been taken from [49], where large devices of the same technology have been characterized and the respective simulation parameters have been extracted.

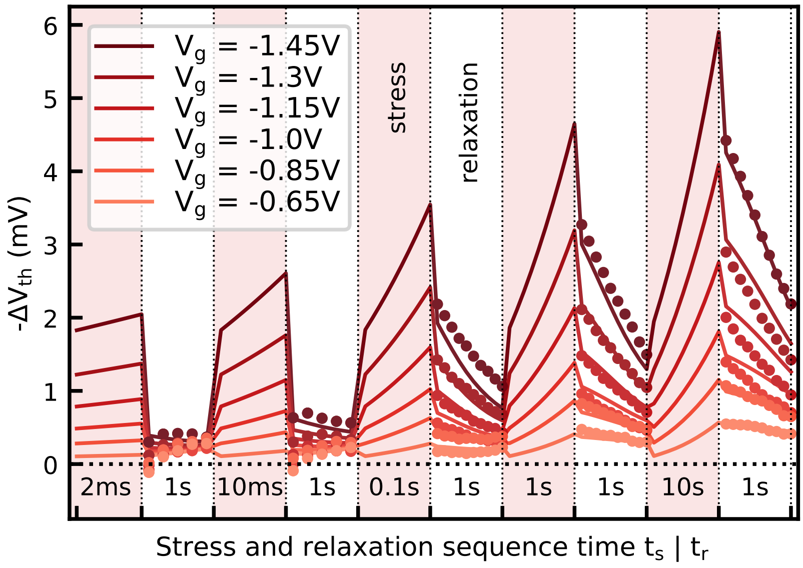


Figure 6.18. Average degradation of the short devices during sets of stress at various gate biases. The simulation data shown as lines and the measurement data given by the points agree well, except for the first moments after short stress which show unusual negative ΔV_{th} shifts in the measurement. The origin of this behavior remains open at this point and requires further experimental and simulation efforts. However, the simulations qualitatively explain the trend of the measurements. Originally published in [BSJ3]

The results, shown in Figure 6.18, show good agreement between the measurements and simulations. For the combination of low stress times and gate biases, unusual negative ΔV_{th} shifts are experienced at very short recovery times. The origin of this behavior is currently unclear and requires further analysis.

Using the simulator, the stress and relaxation behavior shown in Figure 6.17 has been extrapolated to a ten-year time frame, which is a typical margin used in the semiconductor industry. The results are shown in Figure 6.19.

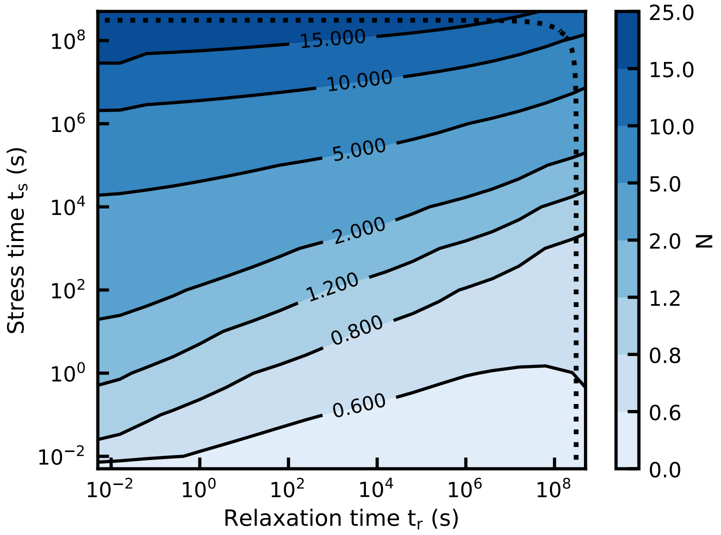


Figure 6.19. Average number of captured defects N (shade) simulated after stress at $V_\mathrm{G} = -1.45\,\mathrm{V}$, over stress and relaxation time. The lines are contour lines separating iso-surfaces for the number of charged defects. The dashed line indicates a ten-year time frame. Originally published in [BSJ3]

6.2.5 Bulk and Drain Bias Dependence

To investigate the influence of bulk and drain biases applied during stress on the device behavior, a number of measurement sequences have been recorded at $V_\mathrm{B,str} \neq 0\,\mathrm{V}$ or $V_\mathrm{D,str} \neq 0\,\mathrm{V}$. In Figure 6.20, the average number of charged defects after $t_\mathrm{r} = 100\,\mathrm{ms}$ of recovery are shown in dependence of V_GB. The measurements labeled with V_G have been recorded with stress biases $V_\mathrm{B} = 0\,\mathrm{V}$ and $V_\mathrm{G} = V_\mathrm{GB}$, while the ones labeled with V_B have been recorded with $V_\mathrm{G} = 1.3\,\mathrm{V}$ and $V_\mathrm{B} \neq 0\,\mathrm{V}$. It can be seen that in both cases the degradation is governed by V_GB only.

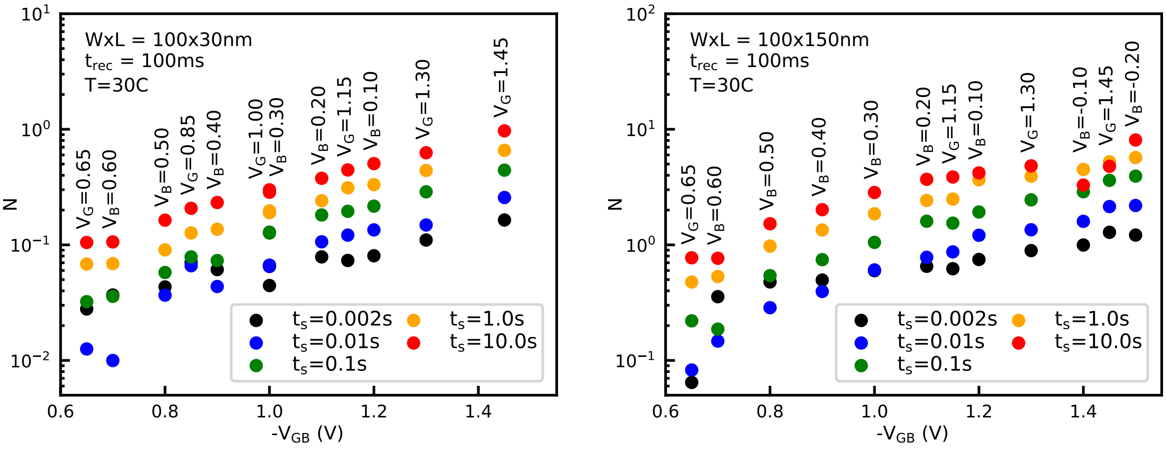


Figure 6.20. Dependence of the average number of charged defects $t_\mathrm{r} = 100\,\mathrm{ms}$ on the stress gate-bulk voltage V_GB for measurements with pure gate stress and measurements with part of the stress voltage applied to the bulk ($V_\mathrm{G} = -1.3\,\mathrm{V}$). Both stress cases have a similar effect on the degradation. Originally published in [BSJ3]

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.
TU Bibliothek WIEN Your knowledge hub

In Figure 6.21, the device relaxation behavior is shown for the same cases. Measurements with pure gate stress are drawn using circles, while measurements with added bulk stress are drawn using crosses. Again, the experienced degradation and recovery seem dependent only on V_{GB}, independent of the potential applied at each contact. However, this result is surprising, as the electric field in the oxide, which drives charge trapping, is largely determined by gate bias and the channel potential, which should be the source bias in inversion. While the gate bias may influence the density of carriers in the channel, the very similar influence of both gate and bulk bias may indicate insufficient control of the device channel during stress.

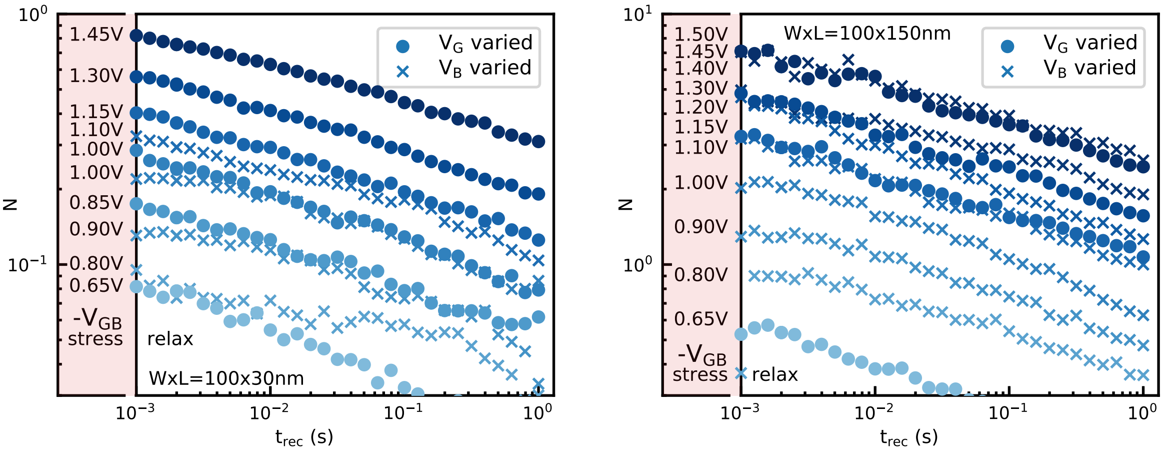


Figure 6.21. Evolution of the average number of charged defects during recovery after $t_{\mathrm{s}} = 100\,\mathrm{ms}$ stress at a number of gate-bulk voltages. Points show the measurements after pure gate stress while crosses show measurements with $V_{\mathrm{G}} = -1.3\,\mathrm{V}$ and varied V_{B}. Again, both stress cases seem to have a similar effect. Originally published in [BSJ3]

The dependence on drain bias stress is shown in Figure 6.22. The results do not show any significant influence of the mild drain stress applied. This seems reasonable, as the applied drain stress is well below the hot-carrier regime.

6.2.6 Conclusions

The custom test chip used in this study allows to efficiently record sufficient statistics for single device characterization and statistical evaluation of the device performance degradation. With this setup, over half a million relaxation traces could be recorded at various stress cases. To analyze this large amount of data, the defect centric model has been used. From this, the behavior of the devices in dependence of gate, bulk, and drain stress, as well as stress time has been investigated. Both gate and bulk stress parameters have severely affected the measured BTI in these devices, consistent with [142] while the elevated drain bias had no effect. The average step heights as well as RTN activity have been independent of the stress parameters. Simulations using defect bands from literature measured on larger devices of the same technology agree with the measured data and allow for lifetime extrapolation employing the measurement data at hand.

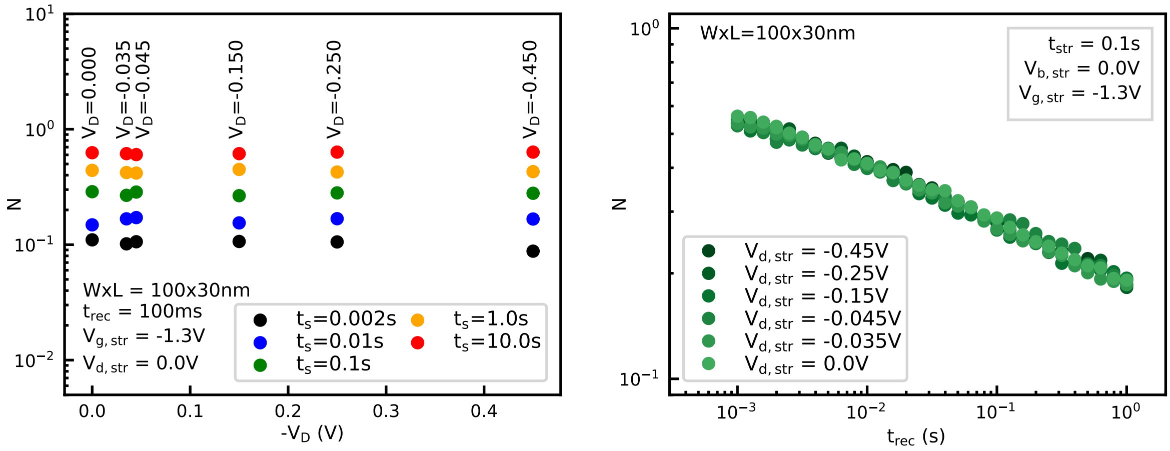


Figure 6.22. Dependence of the average number of charged defects on the drain bias. **Left:** After stress and $t_\mathrm{r} = 100\,\mathrm{ms}$ of recovery. **Right:** During recovery after $t_\mathrm{s} = 100\,\mathrm{ms}$ of stress. The mild drain bias stress applied does not seem to significantly affect the device threshold voltage degradation. Originally published in [BSJ3]

6.3 Single Defects in Few-Layer MoS_2 Devices

As Si seems to approach its scaling limits, many researchers are currently investigating alternative material systems which may be suitable for fabrication of integrated transistors. Some of those emerging technologies are based on devices using 2D materials as channel layers, for instance molybdenum disulfide (MoS_2). In this study, we have investigated single defects using RTN characterization in experimental MoS_2 channel devices. MoS_2, a transition metal dichalcogenide, is a 2D material composed of individual layers with a thickness of about 6.5 Å each. Furthermore, MoS_2 is considered a promising candidate for 2D devices especially due to its relatively large band gap—up to 2.6 eV [143, 144]—compared to other 2D materials. At the moment, however, the performance and reliability of these devices is still severely limited. This is in part due to charge trapping at defects as outlined in this work. Earlier investigations on the reliability of MoS_2 FETs and other 2D devices have revealed oxide traps [145, 146, 147, 148] and trapping sites on top of the channel [145, 149] as possible culprits. All of these studies have been performed on large area devices and thus can not provide detailed information on the trapping characteristics of individual defects. By characterizing the electrical properties of single defects in nano-scale devices, a better insight into their physical nature can be obtained.

6.3.1 Device and Measurements

For the present investigation, around 30 devices have been manufactured by our research partner at Purdue university using an exfoliation method. For this, MoS_2 (SPI Supplies) flakes have first been exfoliated onto SiO_2 substrates using an established scotch tape technique. The used substrates are highly doped Si substrates with 20 nm SiO_2 on top. The flakes have then been structured by e-beam lithography and successive

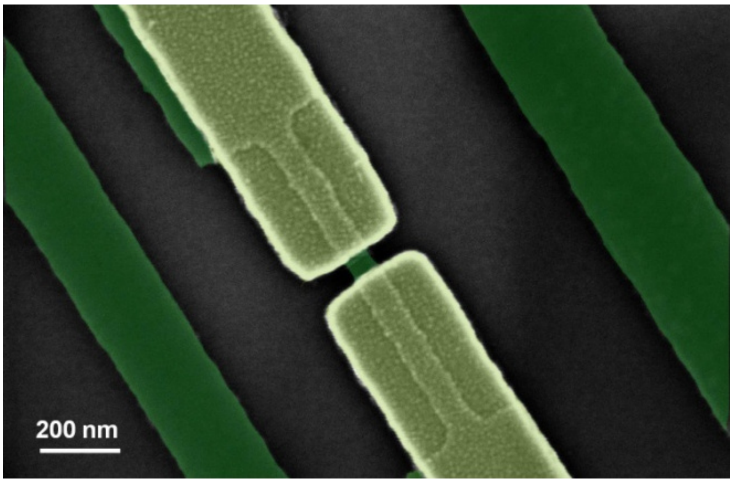


Figure 6.23. Colorized scanning electron microscope picture of one of the studied devices. The source and drain contacts are shown in yellow, below them is the structured MoS_2 layer, shown in green. Originally published in [BSJ7](supporting material)

plasma dry etching. Etching has been performed using 10 sccm of SF_6 and Ar at a pressure of 3 Pa. Both the RF source and bias power have been set to $P = 50\,\mathrm{W}$ and the etching time was 17 s. After this, 65 nm of Ni has been deposited and structured on top to form the source and drain contacts. A SEM image of one of these devices is shown in Figure 6.23. The devices' gate areas are approximately $65 \times 50\,\mathrm{nm}$. This is small enough to enable single defect characterization for this novel technology.

The thickness of the channel material among the available test chips has ranged from 5 nm to 15 nm. This reduces the band gap from 2.6 eV (direct) for single layer material to around 1.29 eV (indirect) for bulk material. The conduction state of the devices has been electrically controlled using the highly doped back gate. A schematic drawing of the devices is shown in Figure 6.24a. The basic characteristics of the devices—recorded in steps of $-0.2\,\mathrm{V}$ with an integration time of 640 µs—are shown in Figure 6.24b-c.

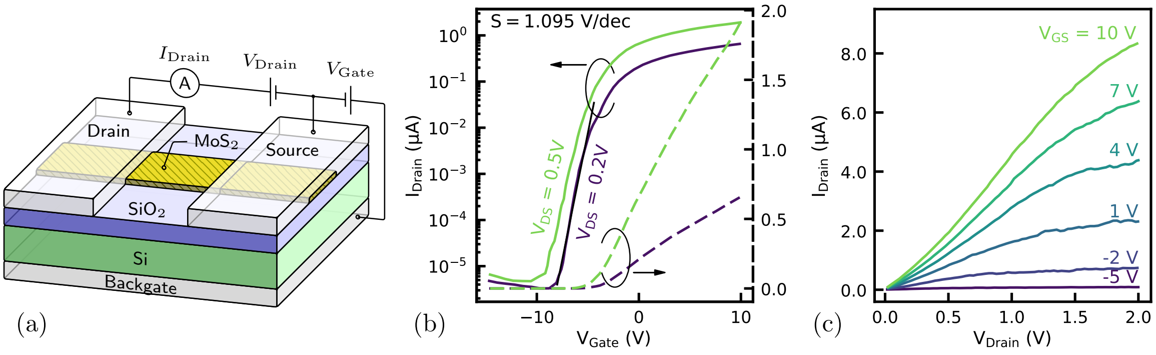


Figure 6.24. (a) Schematics of the devices used during measurement. Few layers of MoS_2 are located on top of a SiO_2 wafer. Source and drain have been contacted on top, while the channel has been controlled using the back gate. **(b)** Transfer characteristics $I_D(V_G)$ of an exemplary device with the subthreshold slope of $S \approx 1.1\,\mathrm{V/dec}$. **(c)** Output characteristics $I_D(V_G)$ for the same device. These devices show normally-on characteristics. Originally published in [BSJ7]

The choice of gate stack and unintentional doping of the channel leads to the devices behaving as normally-on Schottky barrier (SB)-MOSFETs. The pinning of the Fermi level at the Ni-contacts close to the conduction band results in the apparent high electron injection [150, 151]. As can be seen from the figure, the subthreshold slope is around 1.1 V/dec and the on-off ratio is around 10^5. The low subthreshold slope and on-off ratio are due to interface trap densities on the order of $10^{13}\,\mathrm{cm}^{-2}$ and the Schottky barrier operation of the devices. The relatively linear output characteristics are a consequence of the small body thickness of the devices, and not a consequence of "ohmic" contact behavior [151].

For the measurements, the devices have been kept in vacuum at 5×10^{-6} to 10^{-5} Torr and are also shielded from light. At each temperature, an initial $I_\mathrm{D}(V_\mathrm{G})$ has been recorded to allow mapping of the drain current to ΔV_th. Subsequently, RTN traces have been recorded at a number of gate voltages. Once a defect suitable for extraction had been found, additional measurements have been performed in the voltage range where the defect could be seen. Depending on the average charge transition times of the defect and the available measurement ranges, the device temperature has then been changed to characterize the thermal activation energy of the defect. A number of defects have been characterized using this approach, four of which will be discussed here (henceforth called defects A-D).

Figure 6.25 shows a number of exemplary measurements recorded on these MoS_2 devices. Figure 6.25a shows $I_\mathrm{D}(V_\mathrm{G})$ measurements on device D, which has a defect with a relatively large impact introducing a hysteresis-like behavior. Figure 6.25b-c show RTN traces recorded on devices C and D to illustrate the bias dependence of the corresponding defects. It can be seen that defect C is hardly affected by the gate bias, with its average capture and emission times staying relatively constant. For defect D on the other hand, the ratio between the average capture and emission times changes drastically with gate bias. Both trapping behaviors of the defects may be modeled with a two-state model as shown in the inset of Figure 6.25b. A special case is shown in Figure 6.25d, where defect B exhibits anomalous RTN behavior [22]. This defect shows phases of RTN interrupted by phases of inactivity where the defect dwells in its neutral state. The modeling of such charge trapping behavior requires at least three defect states, as shown in the inset.

6.3.2 Charge Transition Times

Once the data had been recorded, the approach based on hidden Markov models, as outlined in Section 5.1.3, has been used to obtain the capture and emission times of the defects. The reason the author has chosen this approach for this study is due to both the comparatively high noise level of the measurement signal and the multi-state defect B. The results are presented in Figure 6.26. Three of the defects—A, C, and D—show two-state RTN behavior, where at any voltage each charge state can be associated with one characteristic time constant. Defect B, on the other hand, shows aRTN, which

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN

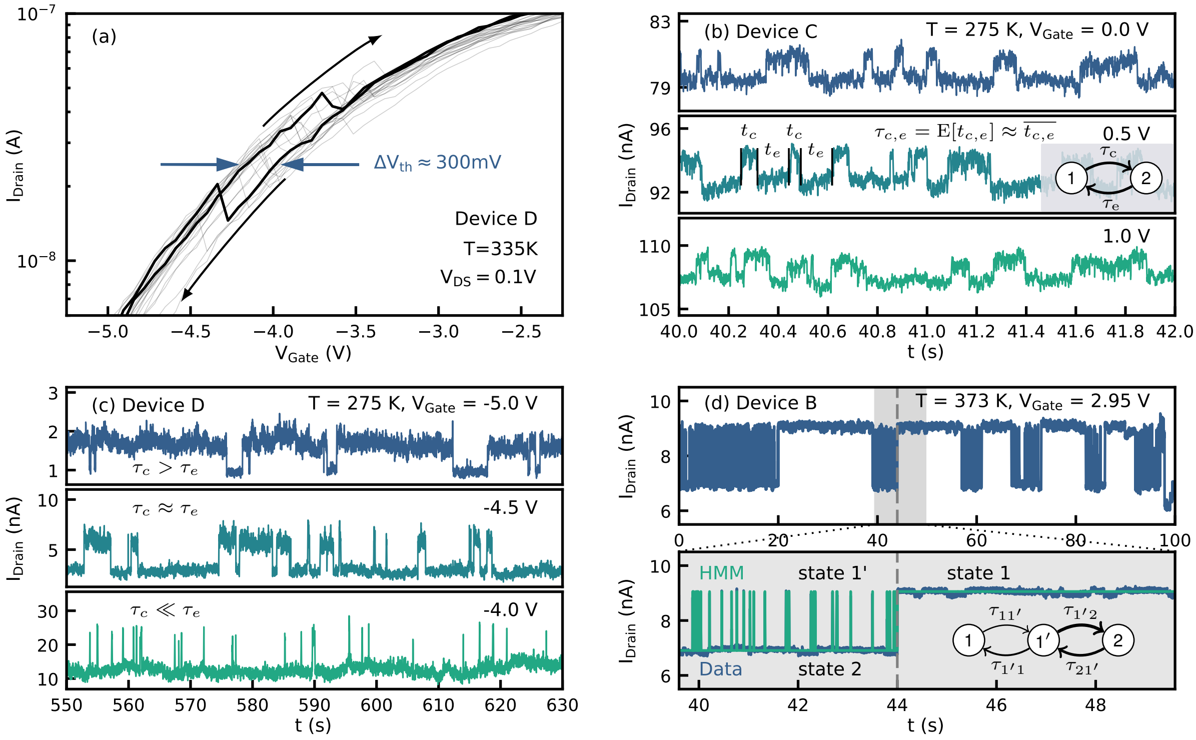


Figure 6.25. Impact of single defects on different device characteristics: **(a)** Steps in the drain current of a transistor during $I_D(V_G)$ sweeps, equivalent to around 300 mV in ΔV_{th}. In large area devices many such defects would be visible as hysteresis, leading to a similar width of the hysteresis, but the individual contributions would not be observable. **(b)** A defect causing RTN. The stochastic behavior of the defect causes a large variety of charge capture and emission times. **(c)** Similar to **(b)**, but this defect shows a significant gate bias dependence. **(d)** A defect showing aRTN, periods of noise are separated by periods of inactivity. The Markov chains necessary to model the respective defect's behavior are shown in the insets **(b)** and **(d)**. aRTN can not be described using a two state model and, in this case, needs an additional neutral state. Originally published in [BSJ7]

requires four transition rates to model, as shown in the inset of Figure 6.25d. The traps A and D show a bias dependence of the charge transition times as has been observed for oxide defects found in regular Si/SiO_2 devices. With increasing bias, their capture time decreases and their emission time increases. Defects B and C exhibit an unusual behavior due to the fact that their capture and emission times stay essentially constant over the entire bias range where they could be characterized. A possible explanation for this is that they are located at a position where their trap level is not affected by the oxide field, i.e. in or on top of the channel. It is important to recall at this point that the devices have been controlled by a backgate contact. For defect D, one can see that the capture time behavior hardly changes with temperature, unlike the emission time. This is an indication that the energy barrier for charge capture is much lower than that for emission.

From the charge transition times, the spatial and energetic positions of the defects could be estimated using electrostatic considerations as outlined in Section 5.1.4.

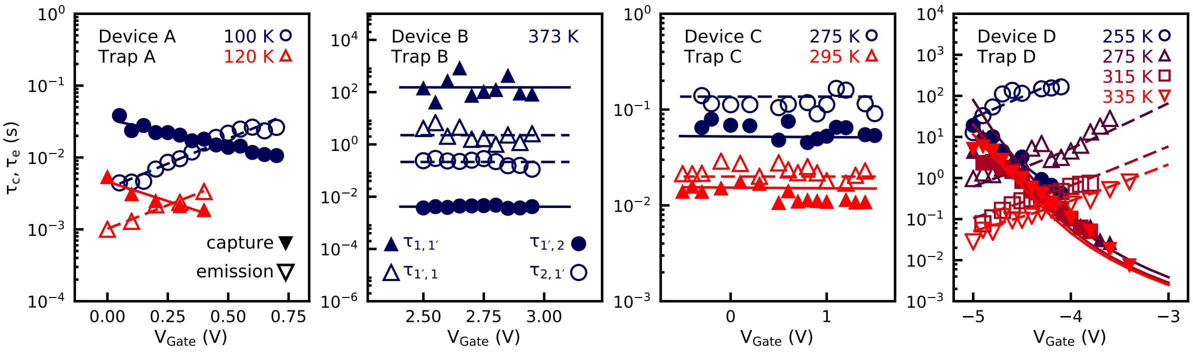


Figure 6.26. Capture and emission times extracted for four defects **(symbols)** and fits **(lines)**. The charge transition times of defects A and D show a significant dependence on the gate bias, while the charge transition times of defects B and C seem to be unaffected by the gate bias. Defect B, which shows aRTN, is described using two sets of transition times due to its additional meta-stable state. The fits for defects B to D are from numerical simulation, while the fit for defect A is a linear fit. This is due to the current inability to simulate these devices at the low temperature at which defect A was measured. Originally published in [BSJ7]

6.3.3 Simulation and Results

To validate the assumptions about the defect behavior in these backgated structures, TCAD simulations have been performed to mimic the charge trapping behavior of the defects. To assess which model to use for the defect simulations, the commonly used SRH and NMP defect models have been compared. For this, simulations have been performed for defects positioned above and below the channel using both the SRH and NMP defect models (see Figure 6.27a). The simulated defects have been positioned at a distance of 2 nm from the channel, at an energy close to the conduction band. The results of these simulations are shown in Figure 6.27b.

From the SRH model below the channel, the symptomatic behavior of this model can be seen. Depending on whether the effective energy of the defect is below or above the Fermi level of the channel, either the capture or emission time will be essentially constant. This is due to the fact that the model does not incorporate a backward barrier—the transition rate for this charge transfer is given by the carrier densities and the effective capture cross section. The same defect modeled using the NMP model yields continuously varying capture and emission times over the whole bias range, in agreement with the results obtained for defects A and D.

For defects simulated on atop the channel the bias dependencies for both defects look similar. The only bias dependence visible is due to the change of carrier density in the channel. The temperature behavior, however, does show a difference. Due to the missing backward barrier in the SRH model, one of the time constants—depending on the energetic position of the defect—will be independent of temperature. This rules out SRH behavior for defect C, where both charge transition times vary with temperature. Finally, defect B could only be characterized at one temperature due to the device failing at the first temperature change. Thus, its data do not allow to distinguish whether this

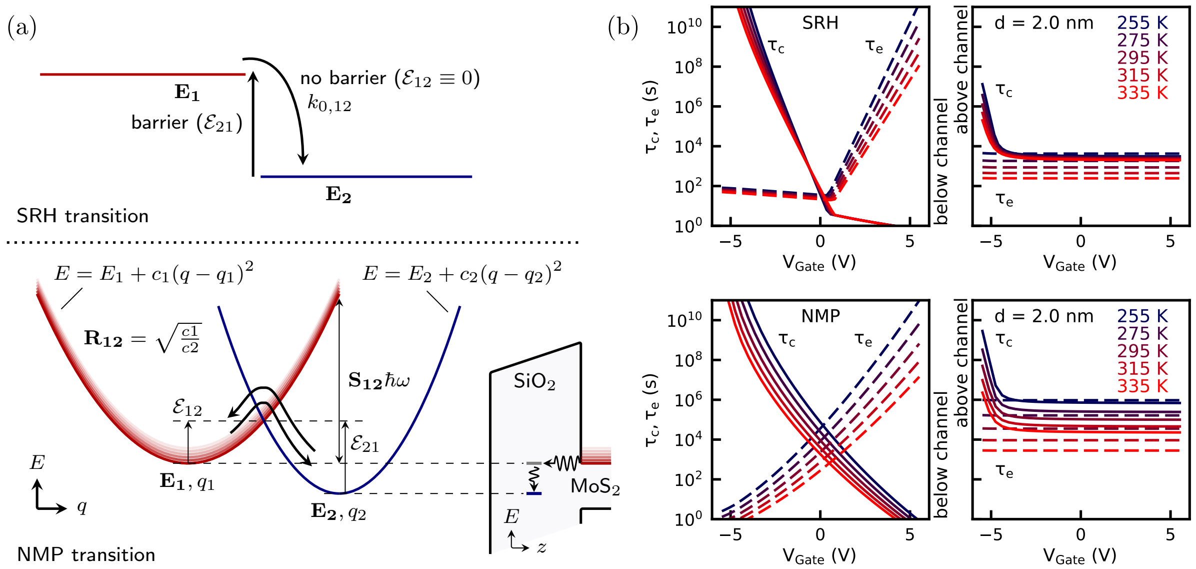


Figure 6.27. The SRH and NMP defect models both widely used for the description of interface and oxide defects. **(a)** The mechanisms for charge transfer as outlined in Section 3.1. **(b)** Simulations of the capture and emission times of defects placed at a distance of 2 nm above and below the channel. It can be seen that due to the lack of a backward energy barrier the SRH model, only one of the charge transition times features a meaningful temperature and bias dependence at any voltage. The SRH model can not describe the strongly voltage dependent capture and emission times as observed for defects A and D. For defect C the temperature dependence of both capture and emission time can not be described by the simple SRH model. Finally, multi state defects such as defect B can not be described at all using the simple SRH model, as it is limited to a defect with only a charged and a neutral state. Originally published in [BSJ7]

defect follows SRH or NMP behavior.

Based on the observed behavior of the defects, simulations have been set up where the defects A and D are placed below the channel and the defects B and C above the channel. The defect parameters have then been optimized to match the extracted time constants, as seen in Figure 6.26.

The results of both the estimation equations and the numerical simulations are shown in Table 6.1. As can be seen from the table, no TCAD results could be obtained for defect A. This is due to the low temperatures (100 K and 120 K) of these measurements, which have led to convergence problems in the simulation. A graphical representation of the results is given in Figure 6.28.

From the characterization it follows that defects A and D reside at some distance (0.8 nm and 3.2 nm) below the channel in the SiO$_2$, and the defects B and C are in direct proximity to the MoS$_2$ above the channel. Possible candidates for defects B and C thus include adsorbed water molecules or processing contaminants. Adsorbed water molecules in particular have already been linked to hysteresis behavior in large area transistors [149] and adsorbates were visible on the wafer at the temperatures defect C was measured at. The more complex behavior of defect B, however, could be caused by

Table 6.1. Extracted trap parameters, calculated with the analytical expressions Equations (5.33) and (5.36) (left), and from TCAD device simulation (right). Energies are relative to the conduction band edge. Originally published in [BSJ7]

Type	**Approximation**		**Fit Parameter**					
Defect	$E_{1',\mathrm{approx}}$	z_approx	E_1[a]	$\varepsilon_{1'1}$[a]	$E_{1'}$	$R_{1'2}$	$S_{1'2}\hbar\omega$	$z_\mathrm{simulated}$
16 (A)[b]	0.003 eV	−0.8 nm	—	—	—	—	—	—
25 (B)	$\lesssim$ 0.010 eV	≈8.0 nm	−0.149 eV	0.99 eV	−0.017 eV	0.80	1.60	8.5 nm
26 (C)	$\lesssim$ −0.018 eV	≈8.0 nm	—	—	−0.016 eV	0.69	1.38	8.0 nm
28 (D)	−0.843 eV	−3.2 nm	—	—	−0.640 eV	0.52	0.33	−1.7 nm

[a] Parameters for three-state defects only [b] No TCAD fit

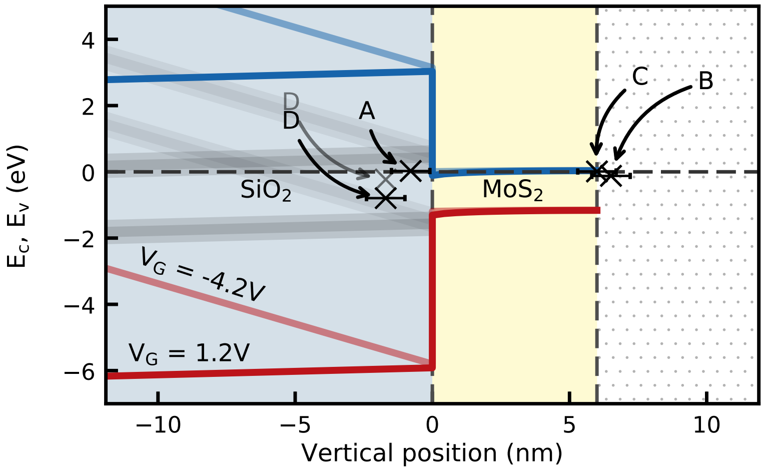


Figure 6.28. Extracted spatial and energetic positions for defects A-D shown in a band diagram simulated for a device considering a 6 nm thin MoS_2 layer. Defects A and D are located below the channel, while defects B and C are on top of it. The gate bias used for simulation was 1.2 V (opaque) and −4.2 V (semi-opaque). The charge transition times for defects A and D change with the gate voltage as their position relative to the Fermi level shifts. The gray lines depict defect bands extracted for SiO_2 [BSC6]. Originally published in [BSJ7]

a more complex type of defect, e.g. an etching related defect. The fast charge transition times in the millisecond range at 100 K for defect A hint at an interface defect, while the behavior of defect D, which is similar to defects measured in regular Si channel transistors, suggests a defect in the oxide. Compared to the estimations, the simulations place defect D closer to the interface, which is likely a result of the approximations made in the estimation equations.

6.3.4 Conclusions

By measuring the responses of single defects in MoS_2 channel transistors, we have been able to extract their characteristic charge capture and emission times in dependence on the gate bias. We have found that part of the defects are affected by the gate bias while others are not. This is most likely due to the positions of the defects in the gate stack. By comparing the SRH and NMP trapping models to the behavior of the defects we have found that the SRH model is not sufficient to replicate their behavior, and thus the NMP model has to be used for simulation—as in silicon channel devices. Using TCAD simulation we have been able to extract physical parameters of the defects, which provides us with some idea as to the physical nature of each defect.

SUMMARY AND OUTLOOK

Summary

The MOSFET is the fundamental building block for many of the technological advances of the last decades. Since its invention, the MOSFET has been continuously improved. The reliable operation of these devices under various operating conditions, e.g. biases and temperatures, is important for the stable operation of the circuits they are part of. Issues limiting the reliability and performance of modern MOSFETs include random telegraph noise (RTN), bias temperature instability (BTI), stress induced leakage current (SILC), and hot carrier (HC) degradation. Responsible for the observed degradation of the device performance in this regard are atomic scale defects which can be introduced during manufacturing or created during operation of the devices. These defects alter the function of the devices by capturing and emitting electric charges. The defects can be located in the insulating oxide of the transistor, at the interface between the oxide and the semiconductor, or in the semiconductor bulk. Most relevant for the performance of a MOSFET are the defects in the oxide and at the interface, which can exchange charges with either the conducting channel or the gate contact.

A number of models have been developed describing the charge trapping behavior of such defects, among them the Shockley–Read–Hall (SRH) model, the non-radiative multi-phonon (NMP) model and the recently proposed hydrogen release model. Detailed ab-initio studies of the atomistic nature of the defects can be performed using density functional theory (DFT) calculations and allow linking their properties to model parameters, while the influence of the defects on device performance can be simulated using TCAD tools employing the charge trapping models. To explain the statistical impact of the defects on devices, the defect centric model can be used. All of these methods, however, require characterization studies on sample MOSFETs to calibrate or verify them. Characterization of the defects is possible using a variety of methods, either probing whole defect distributions or single defects. Popular measurement schemes include methods investigating the defects' influence on the channel current, e.g. RTN and extended MSM (eMSM), methods measuring the charging or discharging current of the defects, e.g. capacitance-voltage (CV) and charge pumping (CP) measurements, and physical characterization methods, e.g. the secondary ion mass spectroscopy (SIMS) and electron paramagnetic resonance (EPR) characterization. From the data obtained from applying these methods, defect parameters can then be extracted. Especially in the case of single defect measurements, the defect parameter extraction from the often noisy measurements can be quite demanding, and in particular for RTN measurements a number of different methods are available to do so. Single defect measurements are helpful in the development of physical defect models and allow to verify if they properly explain the defects' behavior. Distributions of single defects can then be calculated to explain also the behavior of large area devices. To obtain the shape of these distributions one may either try to infer them from single measurements on large area devices containing the cumulative response of many defects or from a sufficient number of separate single defect measurements on small area devices.

In Chapter 6, two studies using different approaches of defect characterization to investigate statistical parameters of defects in two different commercial-grade silicon technologies were shown. In the first study, RTN measurements were used together with an extraction methodology based on the Canny algorithm to find the distribution of the defects' impact on the threshold voltage shift, the trap level, and the vertical depth of the defects. This revealed a bimodal distribution of step heights in these devices, whose origin requires further work. In the second study on silicon devices, an array chip designed especially for characterization purposes was used to conduct mixed BTI and RTN measurements. A methodology based on the defect centric model was then used to find the impact of various bias conditions on the degradation in these devices. Aside from the studies on mature silicon technology, a study exploring single defects in novel devices manufactured using MoS_2 as a channel material was presented. For this study, RTN measurements were performed in a large temperature range and evaluated using an approach based on HMMs to obtain first insights into single defects in experimental MoS_2 devices.

Outlook

Single Defect Characterization

Specialized test structures as the one studied in Section 6.2 significantly decrease the measurement effort required to study statistically significant numbers of single defects. Improved structures of this kind may allow the characterization of multiple devices in parallel and thus also reduce the time required for such studies. Further improvement of the methodologies used to extract the defects' properties from the recorded data may enable such studies on devices with larger gate areas or higher defect densities.

Capacitance-Voltage Measurements

CV measurements have been used for the characterization of interface defects, for oxide defects, however, faster methods of characterization are generally preferred. In particular when studying the creation or passivation of oxide defects, e.g. due to mechanisms involving hydrogen movement, measurement speed is not as critical a

TU Bibliothek WIEN Your knowledge hub
Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

parameter. For this, CV measurements may prove useful as they provide the energetic distribution of the defects.

Influence on the Gate Current

The charge trapping behavior of defects is currently mainly studied via their influence on the drain-source current. However, as outlined in Section 4.1.1, they also influence the gate current either by directly exchanging charge with the gate or by influencing tunneling currents. Studying these effects in conjunction with measurements of the channel current may provide additional insights into the defects' behavior.

REFERENCES

[1] J. E. Lilienfeld. “Method and Apparatus for Controlling Electric Currents”. US1745175. Filed 1926. Jan. 28, 1930.

[2] O. Heil. “Improvements in or Relating to Electrical Amplifiers and other Control Arrangements and Devices”. GB439457. Filed 1935. Dec. 6, 1935.

[3] W. Shockley and G. L. Pearson. “Modulation of Conductance of Thin Films of Semi-Conductors by Surface Charges”. In: *Physical Review* 74.2 (1948), p. 232. DOI: `10.1103/PhysRev.74.232`.

[4] E. H. Nicollian and J. R. Brews. *MOS (Metal Oxide Semiconductor) Physics and Technology*. Wiley New York, 1982. ISBN: 978-0-471-08500-3.

[5] W. L. Brown. “n-Type Surface Conductivity on p-Type Germanium”. In: *Physical Review* 91.3 (Aug. 1953), pp. 518–527. ISSN: 0031-899X. DOI: `10.1103/physrev.91.518`.

[6] I. M. Ross. “Semiconductive Translating Device”. US2791760. Filed 1955. May 7, 1957.

[7] M. M. Atalla, E. Tannenbaum, and E. J. Scheibner. “Stabilization of Silicon Surfaces by Thermally Grown Oxides”. In: *Bell System Technical Journal* 38.3 (May 1959), pp. 749–783. DOI: `10.1002/j.1538-7305.1959.tb03907.x`.

[8] D. Kahng. “Electric Field Controlled Semiconductor Device”. US3102230. Filed 1960. Aug. 27, 1963.

[9] M. M. Atalla. “Semiconductor Devices Having Dielectric Coatings”. US3206670. Filed 1960. Sept. 14, 1965.

[10] B. E. Deal. “Method of Making Stable Semiconductor Devices”. US3426422. Filed 1965. Feb. 11, 1969.

[11] W. H. Miller and B. Fred. “Semiconductor Devices and Passivation Thereof”. US3343049. Filed 1964. Sept. 19, 1967.

[12] D. R. Kerr, J. Hopewell, and D. R. Young. “Methods of Improving Electrical Characteristics of Semiconductor Devices and Products so Produced”. US3303059. Filed 1964. Feb. 7, 1967.

[13] F. M. Wanlass. “Low Stand-by Power Complementary Field Effect Circuitry”. US3356858. Filed 1963. Dec. 5, 1967.

[14] J. C. Sarace, R. E. Kerwin, D. L. Klein, and R. Edwards. “Metal-Nitride-Oxide-Silicon Field-Effect Transistors, with Self-Aligned Gates”. In: *Solid-State Electronics* 11.7 (1968), pp. 653–660. DOI: 10.1016/0038-1101(68)90067-1.

[15] D. Kahng and S. M. Sze. “A Floating Gate and its Application to Memory Devices”. In: *The Bell System Technical Journal* 46.6 (1967), pp. 1288–1295. DOI: 10.1002/j.1538-7305.1967.tb01738.x.

[16] R. H. Dennard. “Field-Effect Transistor Memory”. US3387286. Filed 1967. June 4, 1968.

[17] R. W. Bower, H. G. Dill, K. G. Aubuchon, and S. A. Thompson. “MOS Field Effect Transistors Formed by Gate Masked Ion Implantation”. In: *IEEE Transactions on Electron Devices* 15.10 (1968), pp. 757–761. DOI: 10.1109/T-ED.1968.16511.

[18] G. E. Smith. “Nobel Lecture: The invention and early history of the CCD”. In: *Reviews of Modern Physics* 82.3 (2010), p. 2307. DOI: 10.1103/RevModPhys.82.2307.

[19] M. Denais, V. Huard, C. Parthasarathy, G. Ribes, F. Perrier, N. Revil, and A. Bravaix. “Interface Trap Generation and Hole Trapping under NBTI and PBTI in Advanced CMOS Technology with a 2-nm Gate Oxide”. In: *IEEE Transactions on Devices and Materials Reliability* 4.4 (2004), pp. 715–722. DOI: 10.1109/TDMR.2004.840856.

[20] T. Grasser, B. Kaczer, W. Gös, H. Reisinger, T. Aichinger, P. Hehenberger, P.-J. Wagner, 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 (Nov. 2011), pp. 3652–3666. DOI: 10.1109/ted.2011.2164543.

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

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

[23] A. Ghetti, C. M. Compagnoni, A. S. Spinelli, and A. Visconti. “Comprehensive Analysis of Random Telegraph Noise Instability and its Scaling in Deca–Nanometer Flash Memories”. In: *IEEE Transactions on Electron Devices* 56.8 (2009), pp. 1746–1752. DOI: 10.1109/TED.2009.2024031.

[24] M. Luo, R. Wang, S. Guo, J. Wang, J. Zou, and R. Huang. “Impacts of Random Telegraph Noise (RTN) on Digital Circuits”. In: *IEEE Transactions on Electron Devices* 62.6 (June 2015), pp. 1725–1732. ISSN: 0018-9383. DOI: 10.1109/ted.2014.2368191.

[25] E. Rosenbaum and L. F. Register. "Mechanism of Stress-Induced Leakage Current in MOS Capacitors". In: *IEEE Transactions on Electron Devices* 44.2 (1997), pp. 317–323. DOI: 10.1109/16.557724.

[26] W. Goes, M. Waltl, Y. Wimmer, G. Rzepa, and T. Grasser. "Advanced Modeling of Charge Trapping: RTN, $1/f$ Noise, SILC, and BTI". In: *2014 International Conference on Simulation of Semiconductor Processes and Devices (SISPAD)*. IEEE, Sept. 2014, pp. 77–80. DOI: 10.1109/sispad.2014.6931567.

[27] G. Rzepa, M. Waltl, W. Goes, B. Kaczer, and T. Grasser. "Microscopic Oxide Defects Causing BTI, RTN, and SILC on High-k FinFETs". In: *2015 International Conference on Simulation of Semiconductor Processes and Devices (SISPAD)*. IEEE, Sept. 2015, pp. 144–147. DOI: 10.1109/sispad.2015.7292279.

[28] F. Jiménez-Molinos, A. Palma, F. Gámiz, J. Banqueri, and J. A. López-Villanueva. "Physical Model for Trap-assisted Inelastic Tunneling in Metal-Oxide-Semiconductor Structures". In: *Journal of Applied Physics* 90.7 (Oct. 2001), pp. 3396–3404. ISSN: 0021-8979. DOI: 10.1063/1.1398603.

[29] S. A. Abbas and R. C. Dockerty. "Hot-Carrier Instability in IGFET's". In: *Applied Physics Letters* 27.3 (1975), pp. 147–148. DOI: 10.1063/1.88387.

[30] T. H. Ning, P. W. Cook, R. H. Dennard, C. M. Osburn, S. E. Schuster, and H. Yu. "1 μm MOSFET VLSI Technology: Part IV—Hot-Electron Design Constraints". In: *IEEE Transactions on Electron Devices* 26.4 (1979), pp. 346–353. DOI: 10.1109/T-ED.1979.19433.

[31] M. L. Reed and J. D. Plummer. "Chemistry of Si-SiO$_2$ Interface Trap Annealing". In: *Journal of Applied Physics* 63.12 (June 1988), pp. 5776–5793. ISSN: 0021-8979. DOI: 10.1063/1.340317.

[32] E. H. Poindexter. "MOS Interface States: Overview and Physicochemical Perspective". In: *Semiconductor Science and Technology* 4.12 (Dec. 1989), pp. 961–969. DOI: 10.1088/0268-1242/4/12/001.

[33] 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.

[34] G. J. Gerardi, E. H. Poindexter, P. J. Caplan, and N. M. Johnson. "Interface Traps and Pb Centers in Oxidized (100) Silicon Wafers". In: *Applied Physics Letters* 49 (1986), pp. 348–350. ISSN: 0003-6951. DOI: 10.1063/1.97611.

[35] N. H. Thoan, K. Keunen, V. V. Afanas'ev, and A. Stesmans. "Interface State Energy Distribution and Pb Defects at Si(110)/SiO$_2$ Interfaces: Comparison to (111) and (100) Silicon Orientations". In: 109 (2011), p. 013710. ISSN: 0021-8979. DOI: 10.1063/1.3527909.

[36] J. P. Campbell and P. M. Lenahan. "Density of States of Pb1 Si/SiO$_2$ Interface Trap Centers". In: *Applied Physics Letters* 80.11 (2002), pp. 1945–1947. DOI: 10.1063/1.1461053.

[37] D. M. Fleetwood. "'Border Traps' in MOS Devices". In: *IEEE Transactions on Nuclear Science* 39.2 (1992), pp. 269–271. DOI: 10.1109/23.277495.

[38] M. Bina, K. Rupp, S. Tyaginov, O. Triebl, and T. Grasser. "Modeling of Hot Carrier Degradation using a Spherical Harmonics Expansion of the Bipolar Boltzmann Transport Equation". In: *2012 International Electron Devices Meeting*. IEEE, Dec. 2012. DOI: 10.1109/iedm.2012.6479138.

[39] 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.

[40] 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: *Physical Review B* 89.12 (Mar. 2014), p. 125201. DOI: 10.1103/physrevb.89.125201.

[41] A.-M. El-Sayed, Y. Wimmer, W. Goes, T. Grasser, V. V. Afanas'ev, and A. L. Shluger. "Theoretical Models of Hydrogen-Induced Defects in Amorphous Silicon Dioxide". In: *Physical Review B* 92.1 (July 2015), p. 014107. DOI: 10.1103/physrevb.92.014107.

[42] A.-M. El-Sayed, M. B. Watkins, T. Grasser, V. V. Afanas'ev, and A. L. Shluger. "Hydrogen-Induced Rupture of Strained Si-O Bonds in Amorphous Silicon Dioxide". In: *Physical Review Letters* 114.11 (Mar. 2015), p. 115503. DOI: 10.1103/physrevlett.114.115503.

[43] A. L. Shluger and K. P. McKenna. "Models of Oxygen Vacancy Defects Involved in Degradation of Gate Dielectrics". In: *2013 IEEE International Reliability Physics Symposium (IRPS)*. Apr. 2013, 5A.1.1–5A.1.9. DOI: 10.1109/IRPS.2013.6532018.

[44] S. T. Pantelides, Z.-Y. Lu, C. Nicklaw, T. Bakos, S. N. Rashkeev, D. M. Fleetwood, and R. D. Schrimpf. "The E$'$ Center and Oxygen Vacancies in SiO$_2$". In: *Journal of Non-Crystalline Solids* 354.2-9 (2008), pp. 217–223. DOI: 10.1016/j.jnoncrysol.2007.08.080.

[45] Y. Wimmer, A.-M. El-Sayed, W. Goes, 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*. Vol. 472. 2190. The Royal Society. 2016, p. 20160009. DOI: 10.1098/rspa.2016.0009.

[46] D. T. Gillespie. *Markov Processes: An Introduction for Physical Scientists*. Elsevier Science, 1991. ISBN: 9780080918372.

TU Bibliothek WIEN Your knowledge hub
Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

[47] N. G. Van Kampen. *Stochastic Processes in Physics and Chemistry*. Vol. 1. Elsevier Science, 1992. ISBN: 9780080571386.

[48] W. Shockley and W. T. Read. "Statistics of the Recombinations of Holes and Electrons". In: *Physical review* 87.5 (1952), p. 835. DOI: 10.1103/PhysRev.87.835.

[49] 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: 10.1016/j.microrel.2018.04.002.

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

[51] M. J. Kirton and M. J. 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.

[52] M. Schulz and N. M. Johnson. "Evidence for Multiphonon Emission from Interface States in MOS Structures". In: *Solid State Communications* 25.7 (1978), pp. 481–484. DOI: 10.1016/0038-1098(78)90162-X.

[53] K. Huang, A. Rhys, and N. F. Mott. "Theory of Light Absorption and Non-Radiative Transitions in F-Centres". In: *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*. Vol. 204. 1078. 1950, pp. 406–423. DOI: 10.1098/rspa.1950.0184.

[54] C. Henry and D. Lang. "Nonradiative Capture and Recombination by Multiphonon Emission in GaAs and GaP". In: *Physical Review B* 15.2 (1977), p. 989. DOI: 10.1103/PhysRevB.15.989.

[55] W. Goes, 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. ISSN: 0026-2714. DOI: 10.1016/j.microrel.2017.12.021.

[56] M. Born and R. Oppenheimer. "Zur Quantentheorie der Molekeln". In: *Annalen der Physik* 389.20 (1927), pp. 457–484. DOI: 10.1002/andp.19273892002.

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

[58] E. U. Condon. “Nuclear Motions Associated with Electron Transitions in Diatomic Molecules”. In: *Physical Review* 32 (6 Dec. 1928), pp. 858–872. DOI: 10.1103/PhysRev.32.858.

[59] E. Fermi. *Nuclear Physics: A Course Given by Enrico Fermi at the University of Chicago*. University of Chicago Press, 1950. ISBN: 978-0-226-24365-8.

[60] W. Goes, F. Schanovsky, and T. Grasser. “Advanced Modeling of Oxide Defects”. In: *Bias Temperature Instability for Devices and Circuits*. Springer, 2014, pp. 409–446. DOI: 10.1007/978-1-4614-7909-3_16.

[61] Y.-Y. Liu, F. Liu, R. Wang, J.-W. Luo, X. Jiang, R. Huang, S.-S. Li, and L.-W. Wang. “Characterizing the Charge Trapping across Crystalline and Amorphous Si/SiO_2 /HfO_2 Stacks from First-Principle Calculations”. In: *Physical Review Applied* 12.6 (Dec. 2019), p. 064012. DOI: 10.1103/physrevapplied.12.064012.

[62] Y.-Y. Liu, F. Zheng, X. Jiang, J.-W. Luo, S.-S. Li, and L.-W. Wang. “Ab Initio Investigation of Charge Trapping Across the Crystalline-Si–Amorphous-SiO_2 Interface”. In: *Physical Review Applied* 11.4 (2019), p. 044058. DOI: 10.1103/PhysRevApplied.11.044058.

[63] K. V. Mikkelsen and M. A. Ratner. “Electron Tunneling in Solid-State Electron-Transfer Reactions”. In: *Chemical Reviews* 87.1 (1987), pp. 113–153. DOI: 10.1021/cr00077a007.

[64] F. Schanovsky, O. Baumgartner, and T. Grasser. “Multi Scale Modeling of Multi Phonon Hole Capture in the Context of NBTI”. In: *2011 International Conference on Simulation of Semiconductor Processes and Devices*. IEEE. IEEE, Sept. 2011, pp. 15–18. DOI: 10.1109/sispad.2011.6035038.

[65] T. Grasser, M. Waltl, W. Goes, Y. Wimmer, A.-M. El-Sayed, A. Shluger, and B. Kaczer. “On the Volatility of Oxide Defects: Activation, Deactivation and Transformation”. In: *IEEE Int. Reliab. Phys. Symp. Proc.* IEEE, Apr. 2015, 5A.3.1–5A.3.8. DOI: 10.1109/irps.2015.7112739.

[66] T. Grasser, H. Reisinger, W. Goes, T. Aichinger, P. Hehenberger, P.-J. Wagner, M. Nelhiebel, J. Franco, and B. Kaczer. “Switching Oxide Traps as the Missing Link Between Negative Bias Temperature Instability and Random Telegraph Noise”. In: *IEEE Int. Electron Devices Meet.* IEEE, Dec. 2009, pp. 1–4. DOI: 10.1109/iedm.2009.5424235.

[67] E. Bury, R. Degraeve, M. J. Cho, B. Kaczer, W. Goes, T. Grasser, N. Horiguchi, and G. Groeseneken. “Study of (correlated) Trap Sites in SILC, BTI and RTN in SiON and HKMG Devices”. In: Marina Bay Sands (June 30, 2014). Marina Bay Sands: IEEE, June 30, 2014, pp. 250–253. ISBN: 978-1-4799-3909-1. DOI: 10.1109/IPFA.2014.6898196.

[68] M. Waltl, W. Goes, K. Rott, H. Reisinger, and T. Grasser. “A Single-trap Study of PBTI in SiON nMOS Transistors: Similarities and Differences to the NBTI/pMOS Case”. In: IEEE, June 1, 2014, XT.18.1–XT.18.5. DOI: 10.1109/IRPS.2014.6861195.

[69] M. Waltl, G. Rzepa, A. Grill, W. Goes, J. Franco, B. Kaczer, L. Witters, J. Mitard, N. Horiguchi, and T. Grasser. “Superior NBTI in High-k SiGe Transistors - Part I: Experimental”. In: *IEEE Transactions on Electron Devices* 64.5 (May 2017), pp. 2092–2098. ISSN: 0018-9383. DOI: 10.1109/TED.2017.2686086.

[70] T. Grasser, M. Waltl, Y. Wimmer, W. Goes, R. Kosik, G. Rzepa, H. Reisinger, G. Pobegen, A. El-Sayed, A. Shluger, et al. “Gate-Sided Hydrogen Release as the Origin of ”Permanent” NBTI Degradation: From Single Defects to Lifetimes”. In: *2015 IEEE International Electron Devices Meeting (IEDM)*. IEEE. IEEE, Dec. 2015, pp. 20–1. DOI: 10.1109/iedm.2015.7409739.

[71] E. Cartier, J. H. Stathis, and D. A. Buchanan. “Passivation and Depassivation of Silicon Dangling Bonds at the Si/SiO_2 Interface by Atomic Hydrogen”. In: *Applied Physics Letters* 63.11 (1993), pp. 1510–1512. DOI: 10.1063/1.110758.

[72] E. Cartier, D. A. Buchanan, J. H. Stathis, and D. J. DiMaria. “Atomic Hydrogen-Induced Degradation of Thin SiO_2 Gate Oxides”. In: *Journal of Non Crystalline Solids* 187 (1995), pp. 244–247. DOI: 10.1016/0022-3093(95)00143-3.

[73] A. Yokozawa and Y. Miyamoto. “First-Principles Calculations for Charged States of Hydrogen Atoms in SiO_2”. In: *Physical Review B* 55.20 (1997), p. 13783. DOI: 10.1103/PhysRevB.55.13783.

[74] T. Grasser, K. Rott, H. Reisinger, M. Waltl, P. Wagner, F. Schanovsky, W. Goes, G. Pobegen, and B. Kaczer. “Hydrogen-Related Volatile Defects as the Possible Cause for the Recoverable Component of NBTI”. In: *2013 IEEE International Electron Devices Meeting*. IEEE, Dec. 2013, pp. 15–5. DOI: 10.1109/iedm.2013.6724637.

[75] D. L. Griscom. “Diffusion of Radiolytic Molecular Hydrogen As a Mechanism for the Post-irradiation Buildup of Interface States in SiO_2-on-Si Structures”. In: 58 (1985), pp. 2524–2533. ISSN: 0021-8979. DOI: 10.1063/1.335931.

[76] T. Grasser, M. Waltl, G. Rzepa, W. Goes, Y. Wimmer, A. El-Sayed, A. Shluger, H. Reisinger, and B. Kaczer. “The ”Permanent” Component of NBTI Revisited: Saturation, Degradation-Reversal, and Annealing”. In: *2016 IEEE International Reliability Physics Symposium (IRPS)*. IEEE, Apr. 2016, 5A–2. DOI: 10.1109/irps.2016.7574504.

[77] D. Waldhoer, A.-M. B. El-Sayed, Y. Wimmer, M. Waltl, and T. Grasser. “Atomistic Modeling of Oxide Defects”. In: *Noise in Nanoscale Semiconductor Devices*. Springer International Publishing, 2020, pp. 609–648. DOI: 10.1007/978-3-030-37500-3_18.

[78] K. Binder, D. M. Ceperley, J.-P. Hansen, M. Kalos, D. Landau, D. Levesque, H. Mueller-Krumbhaar, D. Stauffer, and J.-J. Weis. *Monte Carlo Methods in Statistical Physics*. Vol. 7. Springer Science & Business Media, 2012. ISBN: 978-3-642-03162-5. DOI: 10.1007/978-3-642-03163-2.

[79] B. Kaczer, P. Roussel, T. Grasser, and G. Groeseneken. "Statistics of Multiple Trapped Charges in the Gate Oxide of Deeply Scaled MOSFET Devices-Application to NBTI". In: *IEEE Electron Device Letters* 31.5 (May 2010), pp. 411–413. DOI: 10.1109/led.2010.2044014.

[80] L. W. Nagel and D. O. Pederson. *SPICE (Simulation Program with Integrated Circuit Emphasis)*. Tech. rep. EECS Department, University of California, Berkeley, 1973. URL: http://www2.eecs.berkeley.edu/Pubs/TechRpts/1973/22871.html.

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

[82] P. Weckx, B. Kaczer, J. Franco, P. J. Roussel, E. Bury, A. Subirats, G. Groeseneken, F. Catthoor, D. Linten, P. Raghavan, and A. Thean. "Defect-centric Perspective of Combined BTI and RTN Time-dependent Variability". In: *2015 IEEE International Integrated Reliability Workshop (IIRW)*. Oct. 2015, pp. 21–28. DOI: 10.1109/IIRW.2015.7437060.

[83] P. Weckx, B. Kaczer, C. Chen, J. Franco, E. Bury, K. Chanda, J. Watt, P. J. Roussel, F. Catthoor, and G. Groeseneken. "Characterization of Time-dependent Variability using 32k Transistor Arrays in an Advanced HK/MG Technology". In: *2015 IEEE International Reliability Physics Symposium*. Apr. 2015, 3B.1.1–3B.1.6. DOI: 10.1109/IRPS.2015.7112702.

[84] B. Kaczer, J. Franco, P. Weckx, P. J. Roussel, V. Putcha, E. Bury, M. Simicic, A. Chasin, D. Linten, B. Parvais, et al. "A Brief Overview of Gate Oxide Defect Properties and their Relation to MOSFET Instabilities and Device and Circuit Time-Dependent Variability". In: *Microelectronics Reliability* 81 (Feb. 2018), pp. 186–194. DOI: 10.1016/j.microrel.2017.11.022.

[85] B. Ullmann, K. Puschkarsky, M. Waltl, H. Reisinger, and T. Grasser. "Evaluation of Advanced MOSFET Threshold Voltage Drift Measurement Techniques". In: *IEEE Transactions on Device and Materials Reliability* 19.2 (2019), pp. 358–362. DOI: 10.1109/TDMR.2019.2909993.

[86] A. L. McWhorter. "1/f Noise and Related Surface Effects in Germanium". PhD thesis. Massachusetts Institute of Technology, 1955.

[87] H. Mikoshiba. "1/f Noise in n-Channel Silicon-gate MOS Transistors". In: *IEEE Transactions on Electron Devices* 29.6 (June 1982), pp. 965–970. ISSN: 0018-9383. DOI: 10.1109/T-ED.1982.20815.

[88] K. S. Ralls, W. J. Skocpol, L. D. Jackel, R. E. Howard, L. A. Fetter, R. W. Epworth, and D. M. Tennant. "Discrete Resistance Switching in Submicrometer Silicon Inversion Layers: Individual Interface Traps and Low-Frequency ($\frac{1}{f}$) Noise". In: *Physical Review Letters* 52 (3 Jan. 1984), pp. 228–231. DOI: 10.1103/PhysRevLett.52.228.

[89] T. Grasser, K. Rott, H. Reisinger, M. Waltl, J. Franco, and B. Kaczer. "A Unified Perspective of RTN and BTI". In: *2014 IEEE International Reliability Physics Symposium*. June 2014, 4A.5.1–4A.5.7. DOI: 10.1109/IRPS.2014.6860643.

[90] S. Machlup. "Noise in Semiconductors: Spectrum of a Two-Parameter Random Signal". In: *Journal of Applied Physics* 25.3 (1954), pp. 341–343. DOI: 10.1063/1.1721637.

[91] H. Reisinger. "The Time-Dependent Defect Spectroscopy". In: *Bias Temperature Instability for Devices and Circuits*. Springer, 2014, pp. 75–109. DOI: 10.1007/978-1-4614-7909-3_4.

[92] G. Kapila and V. Reddy. "Impact of Sampling Rate on RTN Time Constant Extraction and its Implications on Bias Dependence and Trap Spectroscopy". In: *IEEE Transactions on Device and Materials Reliability* 14.2 (2014), pp. 616–622. DOI: 10.1109/TDMR.2014.2305972.

[93] C.-Y. Chen, Q. Ran, H.-J. Cho, A. Kerber, Y. Liu, M.-R. Lin, and R. W. Dutton. "Correlation of Id- and Ig-Random Telegraph Noise to Positive Bias Temperature Instability in Scaled High-κ/Metal Gate n-Type MOSFETs". In: *2011 International Reliability Physics Symposium*. IEEE. 2011, 3A–2. DOI: 10.1109/IRPS.2011.5784475.

[94] M. Toledano-Luque, B. Kaczer, E. Simoen, R. Degraeve, J. Franco, P. J. Roussel, T. Grasser, and G. Groeseneken. "Correlation of Single Trapping and Detrapping Effects in Drain and Gate Currents of Nanoscaled nFETs and pFETs". In: *2012 IEEE International Reliability Physics Symposium (IRPS)*. Apr. 2012, XT.5.1–XT.5.6. DOI: 10.1109/IRPS.2012.6241935.

[95] F. Crupi, G. Giusi, G. Iannaccone, P. Magnone, C. Pace, E. Simoen, and C. Claeys. "Analytical Model for the 1/f Noise in the Tunneling Current through Metal-Oxide-Semiconductor Structures". In: *Journal of Applied Physics* 106.7 (2009), p. 073710. DOI: 10.1063/1.3236637.

[96] B. Kaczer, T. Grasser, J. Roussel, J. Martin-Martinez, R. O'Connor, B. O'sullivan, and G. Groeseneken. "Ubiquitous Relaxation in BTI Stressing–New Evaluation and Insights". In: *2008 IEEE International Reliability Physics Symposium*. IEEE. 2008, pp. 20–27. DOI: 10.1109/RELPHY.2008.4558858.

TU Bibliothek WIEN Your knowledge hub
Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

[97] M. Denais, C. Parthasarathy, G. Ribes, Y. Rey-Tauriac, N. Revil, A. Bravaix, V. Huard, and F. Perrier. “On-the-fly Characterization of NBTI in Ultra-thin Gate Oxide PMOSFET’s”. In: *IEDM Technical Digest. IEEE International Electron Devices Meeting, 2004.* IEEE. 2004, pp. 109–112. DOI: `10.1109/IEDM.2004.1419080`.

[98] T. Grasser, P.-J. Wagner, P. Hehenberger, W. Goes, and B. Kaczer. “A Rigorous Study of Measurement Techniques for Negative Bias Temperature Instability”. In: *IEEE Transactions on Device and Materials Reliability* 8.3 (2008), pp. 526–535. DOI: `10.1109/tdmr.2008.2002353`.

[99] G. Rescher, G. Pobegen, T. Aichinger, and T. Grasser. “On the Subthreshold Drain Current Sweep Hysteresis of 4H-SiC nMOSFETs”. In: *2016 IEEE International Electron Devices Meeting (IEDM)*. IEEE, Dec. 2016, pp. 10–8. DOI: `10.1109/iedm.2016.7838392`.

[100] R. Castagne and A. Vapaille. “Description of the SiO_2-Si Interface Properties by Means of Very Low Frequency MOS Capacitance Measurements”. In: *Surface Science* 28.1 (1971), pp. 157–193. DOI: `10.1016/0039-6028(71)90092-6`.

[101] L. M. Terman. “An Investigation Of Surface States at a Silicon/Silicon Oxide Interface employing Metal-Oxide-Silicon Diodes”. In: *Solid-State Electronics* 5.5 (1962), pp. 285–299. ISSN: 0038-1101. DOI: `https://doi.org/10.1016/0038-1101(62)90111-9`.

[102] C. N. Berglund. “Surface States at Steam-Grown Silicon-Silicon Dioxide Interfaces”. In: *IEEE Transactions on Electron Devices* 10 (1966), pp. 701–705. DOI: `10.1109/T-ED.1966.15827`.

[103] J. S. Brugler and P. G. A. Jespers. “Charge Pumping in MOS Devices”. In: *IEEE Transactions on Electron Devices* 16.3 (Mar. 1969), pp. 297–302. ISSN: 0018-9383. DOI: `10.1109/T-ED.1969.16744`.

[104] G. Groeseneken, H. E. Maes, N. Beltran, and R. F. De Keersmaecker. “A Reliable Approach to Charge-Pumping Measurements in MOS Transistors”. In: *IEEE Transactions on Electron Devices* 31.1 (Jan. 1984), pp. 42–53. ISSN: 0018-9383. DOI: `10.1109/T-ED.1984.21472`.

[105] D. K. Schroder. *Semiconductor Material and Device Characterization*. New York: John Wiley & Sons, 2015. ISBN: 978-0-471-73906-7.

[106] D. V. Lang. “Deep-level Transient Spectroscopy: A New Method to Characterize Traps in Semiconductors”. In: *Journal of Applied Physics* 45.7 (1974), pp. 3023–3032. DOI: `10.1063/1.1663719`.

[107] K. L. Wang and A. O. Evwaraye. “Determination of Interface and Bulk-trap States of IGFET’s using Deep-level Transient Spectroscopy”. In: *Journal of Applied Physics* 47.10 (1976), pp. 4574–4577. DOI: `10.1063/1.322381`.

[108] A. Neugroschel, C.-T. Sah, K. M. Han, M. S. Carroll, T. Nishida, J. T. Kavalieros, and Y. Lu. "Direct-Current Measurements of Oxide and Interface Traps on Oxidized Silicon". In: *IEEE Transactions on Electron Devices* 42.9 (1995), pp. 1657–1662. DOI: 10.1109/16.405281.

[109] J. G. Simmons and H. A. Mar. "Thermal Bulk Emission and Generation Statistics and Associated Phenomena in Metal-Insulator-Semiconductor Devices under Non-Steady-State Conditions". In: *Physical Review B* 8.8 (1973), p. 3865. DOI: 10.1103/PhysRevB.8.3865.

[110] P. J. Caplan, E. H. Poindexter, B. E. Deal, and R. R. Razouk. "ESR Centers, Interface States, and Oxide Fixed Charge in Thermally Oxidized Silicon Wafers". In: *Journal of Applied Physics* 50.9 (1979), pp. 5847–5854. DOI: 10.1063/1.326732.

[111] F. J. Grunthaner, P. J. Grunthaner, R. P. Vasquez, B. F. Lewis, J. Maserjian, and A. Madhukar. "High-Resolution X-Ray Photoelectron Spectroscopy as a Probe of Local Atomic Structure: Application to Amorphous SiO_2 and the Si-SiO_2 Interface". In: *Physical Review Letters* 43.22 (1979), p. 1683. DOI: 10.1103/PhysRevLett.43.1683.

[112] B. V. Crist. *Example of a "Wide Scan Survey Spectrum" using XPS. Used to Determine what Elements Are and Are Not Present.* [Online; accessed 10-June-2020], https://creativecommons.org/licenses/by-sa/2.5/legalcode. 2006. URL: https://commons.wikimedia.org/wiki/File:Wide.jpg.

[113] 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*. IEEE. IEEE, 2010, pp. 1063–1068. DOI: 10.1109/irps.2010.5488672.

[114] J. J. Thomson. "LXXXIII. Rays of Positive Electricity". In: *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science* 20.118 (1910), pp. 752–767. DOI: 10.1080/14786441008636962.

[115] R. F. K. Herzog and F. P. Viehböck. "Ion Source for Mass Spectrography". In: *Physical Review* 76.6 (1949), pp. 855–856. DOI: 10.1103/PhysRev.76.855.

[116] R. R. Greenberg, P. Bode, and E. A. D. N. Fernandes. "Neutron Activation Analysis: A Primary Method of Measurement". In: *Spectrochimica Acta Part B: Atomic Spectroscopy* 66.3 (2011), pp. 193–241. ISSN: 0584-8547. DOI: 10.1016/j.sab.2010.12.011.

[117] Y. Yuzhelevski, M. Yuzhelevski, and G. Jung. "Random Telegraph Noise Analysis in Time Domain". In: *Review of Scientific Instruments* 71.4 (2000), pp. 1681–1688. DOI: 10.1063/1.1150519.

[118] T. Nagumo, K. Takeuchi, S. Yokogawa, K. Imai, and Y. Hayashi. “New Analysis Methods for Comprehensive Understanding of Random Telegraph Noise”. In: *Electron Devices Meeting (IEDM), 2009 IEEE International*. IEEE. IEEE, Dec. 2009, pp. 1–4. DOI: 10.1109/iedm.2009.5424230.

[119] J. Martin-Martinez, J. Diaz, R. Rodriguez, M. Nafria, and X. Aymerich. “New Weighted Time Lag Method for the Analysis of Random Telegraph Signals”. In: *IEEE Electron Device Letters* 35.4 (Apr. 2014), pp. 479–481. ISSN: 0741-3106. DOI: 10.1109/LED.2014.2304673.

[120] W. A. Taylor. *Change-point Analysis: A Powerful New Tool for Detecting Changes*. 2000.

[121] M. Waltl, P.-J. Wagner, H. Reisinger, K. Rott, and T. Grasser. “Advanced Data Analysis Algorithms for the Time-Dependent Defect Spectroscopy of NBTI”. In: *2012 IEEE International Integrated Reliability Workshop Final Report*. IEEE, Oct. 2012, pp. 74–79. DOI: 10.1109/iirw.2012.6468924.

[122] J. Canny. “A Computational Approach to Edge Detection”. In: *Readings in Computer Vision*. Elsevier, 1987, pp. 184–203. DOI: 10.1016/b978-0-08-051581-6.50024-6.

[123] L. E. Baum, T. Petrie, G. Soules, and N. Weiss. “A Maximization Technique Occurring in the Statistical Analysis of Probabilistic Functions of Markov Chains”. In: *The Annals of Mathematical Statistics* 41.1 (1970), pp. 164–171. DOI: 10.1214/aoms/1177697196.

[124] L. E. Baum et al. *An Inequality and Associated Maximization Technique in Statistical Estimation for Probabilistic Functions of Markov Processes*. 1972.

[125] D. J. Frank and H. Miki. “Analysis of Oxide Traps in Nanoscale MOSFETs Using Random Telegraph Noise”. In: *Bias Temperature Instability for Devices and Circuits*. Springer, 2014, pp. 111–134. DOI: 10.1007/978-1-4614-7909-3_5.

[126] Z. Ghahramani and M. I. Jordan. “Factorial Hidden Markov Models”. In: *Advances in Neural Information Processing Systems*. 1996, pp. 472–478.

[127] 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 (2014), pp. 24–29.

[128] R. Weiss, S. Du, J. Grobler, S. Lebedev, and G. Varoquaux. *hmmlearn 0.2.2*. 2017.

[129] J. Schreiber. “Pomegranate: Fast and Flexible Probabilistic Modeling in Python”. In: *The Journal of Machine Learning Research* 18.1 (2017), pp. 5992–5997.

[130] P. H. C. Eilers. “A Perfect Smoother”. In: *Analytical chemistry* 75.14 (2003), pp. 3631–3636. DOI: 10.1021/ac034173t.

[131] W. S. Cleveland. “Robust Locally Weighted Regression and Smoothing Scatterplots”. In: *Journal of the American statistical association* 74.368 (1979), pp. 829–836. DOI: 10.1080/01621459.1979.10481038.

[132] T. Nagumo, K. Takeuchi, T. Hase, and Y. Hayashi. “Statistical Characterization of Trap Position, Energy, Amplitude and Time Constants by RTN Measurement of Multiple Individual Traps”. In: *Electron Devices Meeting (IEDM), 2010 IEEE International*. IEEE. IEEE, Dec. 2010, pp. 28–3. DOI: 10.1109/iedm.2010.5703437.

[133] K. K. Hung, P. K. Ko, C. Hu, and Y. C. 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.

[134] S. Villa, G. De Geronimo, A. Pacelli, A. L. Lacaita, and A. Longoni. “Application of 1/f Noise Measurements to the Characterization of Near-interface Oxide Traps in ULSI n-MOSFETs”. In: *Microelectronics Reliability* 38.12 (1998), pp. 1919–1923. DOI: 10.1016/S0026-2714(98)00069-9.

[135] T. Grasser, H. Reisinger, P.-J. Wagner, W. Goes, F. Schanovsky, and B. Kaczer. “The Time Dependent Defect Spectroscopy (TDDS) for the Characterization of the Bias Temperature Instability”. In: *IEEE Int. Reliab. Phys. Symp. Proc.* IEEE, 2010, pp. 16–25. DOI: 10.1109/irps.2010.5488859.

[136] P. V. Gray and D. M. Brown. “Density of SiO_2–Si Interface States”. In: *Applied Physics Letters* 8.2 (1966), pp. 31–33. DOI: 10.1063/1.1754468.

[137] G. T. Sasse, F. G. Kuper, and J. Schmitz. “MOSFET Degradation Under RF Stress”. In: *IEEE Transactions on Electron Devices* 55.11 (2008), pp. 3167–3174. DOI: 10.1109/TED.2008.2004650.

[138] E. H. Nicollian and A. Goetzberger. “MOS Conductance Technique for Measuring Surface State Parameters”. In: *Applied Physics Letters* 7.8 (1965), pp. 216–219. DOI: 10.1063/1.1754385.

[139] E. H. Nicollian and A. Goetzberger. “The Si-SiO, Interface–Electrical Properties as Determined by the Metal-Insulator-Silicon Conductance Technique”. In: *The Bell System Technical Journal* 46.6 (1967), pp. 1055–1033. DOI: 10.1002/j.1538-7305.1967.tb01727.x.

[140] A. Oshima, T. Komawaki, K. Kobayashi, R. Kishida, P. Weckx, B. Kaczer, T. Matsumoto, and H. Onodera. “Physical-based RTN Modeling of Ring Oscillators in 40-nm SiON and 28-nm HKMG by Bimodal Defect-Centric Behaviors”. In: *2016 International Conference on Simulation of Semiconductor Processes and Devices (SISPAD)*. IEEE, Sept. 2016, pp. 327–330. DOI: 10.1109/sispad.2016.7605213.

[141] 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)*. IEEE. IEEE, Oct. 2015, pp. 73–76. DOI: `10.1109/iirw.2015.7437071`.

[142] A. Kerber, K. Maitra, A. Majumdar, M. Hargrove, R. J. Carter, and E. A. Cartier. "Characterization of Fast Relaxation During BTI Stress in Conventional and Advanced CMOS Devices With HfO_2/TiN Gate Stacks". In: *IEEE Transactions on Electron Devices* 55.11 (2008), pp. 3175–3183. DOI: `10.1109/TED.2008.2004853`.

[143] A. R. Klots, A. K. M. Newaz, B. Wang, D. Prasai, H. Krzyzanowska, J. Lin, D. Caudel, N. J. Ghimire, J. Yan, B. L. Ivanov, K. A. Velizhanin, A. Burger, D. G. Mandrus, N. H. Tolk, S. T. Pantelides, and K. I. Bolotin. "Probing Excitonic States in Suspended Two-Dimensional Semiconductors by Photocurrent Spectroscopy". In: *Scientific Reports* 4.1 (Oct. 2014), p. 6608. DOI: `10.1038/srep06608`.

[144] F. A. Rasmussen and K. S. Thygesen. "Computational 2D Materials Database: Electronic Structure of Transition-Metal Dichalcogenides and Oxides". In: *Journal of Physical Chemistry C* 119.23 (June 2015), pp. 13169–13183. DOI: `10.1021/acs.jpcc.5b02950`.

[145] Y. Y. Illarionov, G. Rzepa, M. Waltl, T. Knobloch, A. Grill, M. M. Furchi, T. Mueller, and T. Grasser. "The Role of Charge Trapping and MoS_2/SiO_2 and MoS_2/hBN Field-Effect Transistors". In: *2D Materials* 3 (2016), p. 035004. DOI: `10.1088/2053-1583/3/3/035004`.

[146] Y. Y. Illarionov, A. D. Smith, S. Vaziri, M. Ostling, T. Mueller, M. C. Lemme, and T. Grasser. "Bias-Temperature Instability in Single-Layer Graphene Field-Effect Transistors". In: *Applied Physics Letters* 105.14 (Oct. 2014), p. 143507. DOI: `10.1063/1.4897344`.

[147] Y. Guo, X. Wei, J. Shu, B. Liu, J. Yin, C. Guan, Y. Han, S. Gao, and Q. Chen. "Charge Trapping at the MoS_2-SiO_2 Interface and Its Effects on the Characteristics of MoS_2 Metal-Oxide-Semiconductor Field Effect Transistors". In: *Applied Physics Letters* 106.10 (Mar. 2015), p. 103109. DOI: `10.1063/1.4914968`.

[148] Y. Park, H. W. Baac, J. Heo, and G. Yoo. "Thermally Activated Trap Charges Responsible for Hysteresis in Multilayer MoS_2 Field-Effect Transistors". In: *Applied Physics Letters* 108.8 (Feb. 2016), p. 083102. DOI: `10.1063/1.4942406`.

[149] D. J. Late, B. Liu, H. S. S. R. Matte, V. P. Dravid, and C. N. R. Rao. "Hysteresis in Single-Layer MoS_2 Field Effect Transistors". In: *ACS Nano* 6.6 (2012), pp. 5635–5641. DOI: `10.1021/nn301572c`.

[150] S. Das, H. Y. Chen, A. V. Penumatcha, and J. Appenzeller. “High Performance Multilayer MoS_2 Transistors with Scandium Contacts”. In: *Nano Letters* 13 (2013), pp. 100–105. DOI: `10.1021/nl303583v`.

[151] J. Appenzeller, F. Zhang, S. Das, and J. Knoch. “Transition Metal Dichalcogenide Schottky Barrier Transistors”. In: *2D Mater. Nanoelectron.* 17 (2016), p. 207. DOI: `10.1201/b19623-11`.

LIST OF PUBLICATIONS

Scientific Journals, Book Contribution

[BSJ1] **B. Stampfer**, A. Grill, and M. Waltl. “Advanced Electrical Characterization of Single Oxide Defects Utilizing Noise Signals”. In: *Noise in Nanoscale Semiconductor Devices*. Springer International Publishing, 2020, pp. 229–257. DOI: 10.1007/978-3-030-37500-3_7.

[BSJ2] **B. Stampfer**, F. Schanovsky, T. Grasser, and M. Waltl. “Semi-Automated Extraction of the Distribution of Single Defects for nMOS Transistors”. In: *Micromachines* 11.4 (Apr. 2020), p. 446. ISSN: 2072-666X. DOI: 10.3390/mi11040446.

[BSJ3] **B. Stampfer**, M. Simicic, P. Weckx, A. Abbasi, B. Kaczer, T. Grasser, and M. Waltl. “Extraction of Statistical Gate Oxide Parameters from Large MOSFET Arrays”. In: *IEEE Transactions on Device and Materials Reliability* (2020), pp. 1–1. ISSN: 1530-4388. DOI: 10.1109/tdmr.2020.2985109.

[BSJ4] 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* (2020), p. 113746. ISSN: 0026-2714. DOI: 10.1016/j.microrel.2020.113746.

[BSJ5] A. Grill, **B. Stampfer**, K.-S. Im, J.-H. Lee, C. Ostermaier, H. Ceric, M. Waltl, and T. Grasser. “Electrostatic Coupling and Identification of Single-Defects in GaN/AlGaN Fin-MIS-HEMTs”. In: *Solid-State Electronics* 156 (June 2019), pp. 41–47. ISSN: 0038-1101. DOI: 10.1016/j.sse.2019.02.004.

[BSJ7] **B. Stampfer**, F. Zhang, Y. Y. Illarionov, T. Knobloch, P. Wu, M. Waltl, A. Grill, J. Appenzeller, and T. Grasser. “Characterization of Single Defects in Ultrascaled MoS_2 Field-Effect Transistors”. In: *ACS Nano* 12.6 (June 2018), pp. 5368–5375. ISSN: 1936-0851. DOI: 10.1021/acsnano.8b00268.

[BSJ6] T. Knobloch, G. Rzepa, Y. Y. Illarionov, M. Waltl, F. Schanovsky, **B. Stampfer**, M. M. Furchi, T. Mueller, and T. Grasser. “A Physical Model for the Hysteresis in MoS_2 Transistors”. In: *IEEE Journal of the Electron Devices Society* 6 (2018), pp. 972–978. ISSN: 2168-6734. DOI: 10.1109/jeds.2018.2829933.

Conference Proceedings

[BSC1] T. Grasser, B. Kaczer, B. O'Sullivan, G. Rzepa, **B. Stampfer**, and M. Waltl. "The Mysterious Bipolar Bias Temperature Stress from the Perspective of Gate-Sided Hydrogen Release". In: *2020 IEEE International Reliability Physics Symposium (IRPS)*. 2020, pp. 1–6. DOI: `10.1109/IRPS45951.2020.9129198`.

[BSC2] **B. Stampfer**, M. Simicic, P. Weckx, A. Abbasi, B. Kaczer, T. Grasser, and M. Waltl. "Statistical Characterization of BTI and RTN using Integrated pMOS Arrays". In: *2019 IEEE International Integrated Reliability Workshop (IIRW)*. IEEE, Oct. 2019. DOI: `10.1109/iirw47491.2019.8989904`.

[BSC3] T. Grasser, **B. Stampfer**, M. Waltl, G. Rzepa, K. Rupp, F. Schanovsky, G. Pobegen, K. Puschkarsky, H. Reisinger, B. 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: *2018 IEEE International Reliability Physics Symposium (IRPS)*. IEEE, Mar. 2018. DOI: `10.1109/irps.2018.8353540`.

[BSC4] T. Grasser, M. Waltl, K. Puschkarsky, **B. Stampfer**, G. Rzepa, G. Pobegen, H. Reisinger, H. Arimura, and B. Kaczer. "Implications of Gate-sided Hydrogen Release for Post-stress Degradation Build-up After BTI Stress". In: *2017 IEEE International Reliability Physics Symposium (IRPS)*. IEEE, Apr. 2017. DOI: `10.1109/irps.2017.7936334`.

[BSC5] A. Grill, **B. Stampfer**, M. Waltl, K.-S. Im, J.-H. Lee, C. Ostermaier, H. Ceric, and T. Grasser. "Characterization and Modeling of Single Defects in GaN/AlGaN Fin-MIS-HEMTs". In: *2017 IEEE International Reliability Physics Symposium (IRPS)*. IEEE, Apr. 2017. DOI: `10.1109/irps.2017.7936285`.

[BSC6] T. Knobloch, G. Rzepa, Y. Y. Illarionov, M. Waltl, F. Schanovsky, M. Jech, **B. Stampfer**, M. M. Furchi, T. Muller, and T. Grasser. "Physical Modeling of the Hysteresis in MoS_2 Transistors". In: *2017 47th European Solid-State Device Research Conference (ESSDERC)*. IEEE, Sept. 2017. DOI: `10.1109/essderc.2017.8066647`.

[BSC7] G. Rzepa, J. Franco, A. Subirats, M. Jech, A. Chasin, A. Grill, M. Waltl, T. Knobloch, **B. Stampfer**, T. Chiarella, N. Horiguchi, L. A. Ragnarsson, D. Linten, B. Kaczer, and T. Grasser. "Efficient Physical Defect Model Applied to PBTI in High-κ Stacks". In: *2017 IEEE International Reliability Physics Symposium (IRPS)*. IEEE, Apr. 2017. DOI: `10.1109/irps.2017.7936425`.

Master Thesis

[BST1] **B. Stampfer**. "Trap Assisted Tunneling and Band Interaction using the Non-Radiative Multi Phonon Model". MA thesis. Technische Universität Wien, 2016.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN