Technische Universität Wien
Fakultät für Elektrotechnik und Informationstechnik

Negative bias temperature instability and hot-carrier degradation of 130 nm technology transistors including recovery effects

DISSERTATION

ausgeführt zum Zwecke der Erlangung des akademischen Grades eines
Doktors der technischen Wissenschaften

eingereicht an der Technischen Universität Wien
Fakultät für Elektrotechnik und Informationstechnik
von

Gunnar Andreas Rott

Stockmannstraße 34
81477 München, Deutschland
Matrikelnummer: 1129988
geboren am 13. November 1985 in Flensburg

April 2018

1. Gutachter: Prof. Dr.-techn. Klaus-Tibor Grasser
2. Gutachter: Prof. Dr.-techn. ...
3. Gutachter: Prof. Dr.-techn. ...

Betreuer: Prof. Dr.-techn. Klaus-Tibor Grasser

Contents

To my wife Karina, my family and my cats.

Kurzfassung

Metall-Oxid-Halbleiter-Feldeffekttransistoren (englisch metal-oxide-semiconductor field-effect transistor, MOSFET) werden heutzutage in zahlreichen elektronischen Anwendungen verwendet. Verglichen mit bipolaren Transistoren wird ein sehr niedriger Eingangsstrom an der Steuerelektrode (Gate) benötigt, um den Stromfluss von Drain nach Source zu beeinflussen. Aus diesem Grund haben Schaltungen, welche aus MOSFETs bestehen, einen sehr niedrigen Stromverbrauch und werden als Grundlage für verschiedenste analoge und digitale Anwendungen eingesetzt. Zudem können MOSFETs auf wenige Nanometer herunterskaliert werden und ermöglichen so einen sehr hohen Integrationsgrad. Der Nachteil dieses sehr hohen Integrationsgrads (VLSI) liegt nicht nur in der Streuung der elektrischen Parameter der gefertigten Transistoren, sondern auch in deren unterschiedlichem Alterungsverhalten.

Zwei Haupteffekte der Transistoralterung, “negative” und “positive bias temperature instability” (N/PBTI) sowie “hot-carrier degradation” (HCD), führen basierend auf dem Schalt- und Lastprofil jedes einzelnen Transistors zu einer Veränderung der ursprünglichen elektrischen Parameter der Transistoren in einer elektronischen Schaltung. Um diese Veränderung vorhersagen zu können, werden Messdaten zur Zuverlässigkeit der Parameter aufgenommen und physikalische Modelle entwickelt. Die Vorhersagbarkeit der Parameteränderungen ist entscheidend, um das Verhalten der Schaltung über ihre Lebensdauer zu simulieren und verstehen zu können. Hierbei liegt die Lebensdauer zwischen wenigen Jahren für Unterhaltungselektronik und einigen Jahrzehnten für Automobil- sowie militärischen Anwendungen.

Diese Arbeit befasst sich mit der Messung und Analyse von HCD, NBTI, inklusive dem “recovery”-Effekt, sowie deren Zusammenspiel. Für die Aufnahme von HCD-Daten können standardisierte Laborgeräte verwendet werden, wohingegen NBTI nur mit spezieller Messtechnik untersucht werden sollte, um das “recovery”-Verhalten richtig abbilden zu können. Daher wurde ein spezielles Messinstrument für diese Arbeit aufgebaut und für sämtliche Messungen verwendet, bei welchen der “recovery”-Effekt zu erwarten ist. Das Besondere an diesem Messinstrument ist die sehr kurze Umschaltzeit ($\approx 1\,\mu\mathrm{s}$) zwischen der Stress- und Messphase, wodurch die “recovery”-Information über schnell relaxierende Störstellen erhalten bleibt, welche an der $\mathrm{Si}-\mathrm{SiO_2}$-Grenzfläche und im Gate-Oxid liegen. Aus diesem Grund wurden auch diejenigen Messungen, bei welchen das Zusammenspiel

von HCD und NBTI untersucht wurde, mit diesem besonderen Messinstrument durchgeführt.

“Hot-Carrier”-Messdaten einer 130 nm Transistortechnologie mit einer Nominalspannung von 3.3 V, welche in Automobil- und analogen Anwendungen eingesetzt wird, zeigen zwei Hauptalterungseffekte. Der erste Effekt ($D_{\mathrm{Isub,max}}$) liegt in der Nähe der traditionellen “worst-case” HCD Bedingung und ein weiterer ($D_{\psi,\mathrm{max}}$) bei $V_{\mathrm{ds}} = V_{\mathrm{gs}}$. Beide Alterungseffekte kompensieren die Änderung der Einsatzspannung nach einer gewissen Zeit, wobei dieser Vorgang sehr gut anhand der Änderung des Substratstroms während der Stresszeit beobachtet werden kann. In der Literatur werden mehrere Mechanismen für “Hot-Carrier”-Alterung berichtet. Der erste Effekt beruht auf der Dissoziation von Bindungen, welche durch einen einzelnen hochenergetischen Ladungsträger verursacht werden kann, wohingegen der zweite Effekt durch das Anregen der Bindung durch die Schwingungswechselwirkung zahlreicher niederenergetischer Ladungsträger begründet wird. Die gezeigten Ergebnisse betonen die Bedeutsamkeit dieser Methode und ermöglichen einen Benchmark für die Simulation der Transistoralterung, da die beiden beobachteten Alterungseffekte gut voneinander unterschieden und gezielt durch eine geeignete Wahl der Gate-Source-Spannung angeregt werden können. Langzeitaufnahmen der Transistoralterung zeigen ferner, dass der Alterungseffekt $D_{\mathrm{Isub,max}}$ selbst für ein niedriges V_{gs} durch $D_{\psi,\mathrm{max}}$ kompensiert wird.

Außerdem werden HCD- und NBTI-Messergebnisse desselben Technologieknotens, jedoch eines Transistors mit einer Nominalspannung von 1.5 V, in Form von Diagrammen der Einsatzspannungsverschiebung während der Stress- und “recovery”-Phase sowie mit Hilfe von “continuous capture emission time (CET) maps” vorgestellt. Transistoren mit $l = 100\,\mathrm{nm}$, $w = 10\,\mu\mathrm{m}$ wurden basierend auf einer zweidimensionalen Matrix der Gate- und Drain-Source-Stressspannung bei 398 K untersucht. Die gewählten Stressbedingungen beinhalten sowohl homogenen ($V_{\mathrm{ds}} = 0$) und inhomogenen ($V_{\mathrm{ds}} \ll V_{\mathrm{nom}}$) NBTI, reinen HCD ($V_{\mathrm{gs}} < V_{\mathrm{nom}}$) als auch die Mischung von NBTI und HCD. Die Ergebnisse zeigen deutlich, dass die NBTI-”recovery” für eine ansteigende Drain-Source-Spannung ($V_{\mathrm{ds}} > V_{\mathrm{nom}}$) verhältnismäßig gering ausfällt und überwiegend die permanente Alterung durch HCD nach dem Ende des Stresses bestehen bleibt. Des Weiteren kann ein Minimum der NBTI-Alterung für ein bestimmtes V_{ds} beobachtet werden. Mit Hilfe von “CET Maps” wird ersichtlich, dass die Wahrscheinlichkeit für eine Relaxation von Störstellen bei langen Stresszeiten sehr gering wird, wohingegen für kürzere Stresszeiten eine relaxierende Komponente deutlich erkennbar ist.

Da die meisten Zuverlässigkeitsmodelle das statistisch gemittelte Alterungsverhalten eines Transistors beschreiben, wurden auch Informationen über die Streuung der Einsatzspannung vor und nach dem NBTI-Stress aufgezeichnet und analysiert. Darüber hinaus werden der permanente und relaxierbare Anteil der Alterung voneinander getrennt betrachtet. Dies ermöglicht eine genauere Vorhersage des Verhaltens einer Schaltung über deren Lebensdauer. Da die Streuung der elektrischen Parameter eines Transistors antiproportional zur Gateoxid-Fläche zunimmt, sind die erlangten Ergebnisse unter anderem von großer Bedeutung für die Simulation der Alterung von Transistoren in SRAM-Zellen und CPU-Kernen.

Entgegen anderer Ansätze wurde in dieser Arbeit eine neue Methodik für das Fitting von “CET Map”-Daten entwickelt, bei welcher die “Recovery”-Messkurven in Abhängigkeit ihrer Stresszeit gemeinsam gefittet werden. Das Ergebnis ist ein Algorithmus, welcher nicht nur sehr effizient arbeitet, sondern auch die Messdaten in einer hohen Güte reproduziert. Aufgrund des dreidimensionalen Fitting-Verfahrens (Stresszeit, “Recovery”-Zeit, Einsatzspannungsdrift) sind die erzeugten “CET Maps” mathematisch stetig und erlauben eine Extrapolation zu realistischen Zeitskalen. Das am Ende dieser Arbeit gezeigte “CET Map”-Modell nutzt die Stress- und “Recovery”-Zeit, die Gate- und Drain-Source-Spannung sowie die Stresstemperatur als Eingangsparameter, um die Einsatzspannungsverschiebung zu berechnen, und kann aufgrund seiner analytischen Formulierung in Schaltkreissimulatoren verwendet werden.

In einer gemeinsamen Studie wurde die NBTI-Alterung unter analogen Stressspannungssignalen von p-MOSFETs experimentell und theoretisch untersucht. Als Ergebnis dieser Studie wird ein Kompaktmodell für die Alterung von Transistoren vorgestellt, welches auf einer umfangreichen TCAD-Analyse und Messdaten beruht, welche das “Recovery”-Verhalten 1 µs nach dem Stress wiedergeben. Da das Kompaktmodell numerisch sehr effizient ist, kann es unmittelbar in Schaltkreissimulatoren verwendet werden, um die NBTI-Alterung realitätsnah zu bestimmen. Neben der Berechnung der statistisch gemittelten Alterungsinformation eines einzelnen Transistors kann mit dem Kompaktmodell auch die Streuung elektrischer Transistorparameter sowie der Schaltkreisperformanz ermittelt werden.

Abstract

Metal-oxide-semiconductor field-effect-transistors (MOSFETs) are used in various electronic applications everyday. Compared to bipolar transistors, a very low input current is needed to control the load current. That is why, circuits based on MOSFETs have very low power consumption and it is the main reason why they are used as the basis for a variety of analog and digital applications. On the other hand, MOSFETs can be scaled down to few nanometers and enable a high density of integration. The drawback of this very-large-scale integration (VLSI) is not only the spread of electrical parameters of the fabricated transistors but also their different aging behavior.

Two main aging effects, negative and positive bias temperature instability (N/PBTI) as well as hot-carrier degradation (HCD), change the initial electrical parameters of each single transistor in the circuit based on its control and load profile. To predict the drift of each parameter, reliability measurement data is recorded and physics-based models are developed. The prediction of the parameter drift is crucial to simulate and understand the behavior of the circuit over lifetime. At this, the end of life of an electronic circuit ranges between few years for consumer electronics up to several decades for automotive, aerospace and military applications.

This work focuses on the measurement and analysis of HCD, NBTI, including the recovery effect, and their interplay. In order to record HCD reliability data a standard lab setup has been used. During NBTI measure-stress-measure (MSM) sequences recovery effects are expected. Therefore, a special measurement device with an ultra-short delay about 1 µs between the stress and measurement phase was built-up and has been employed to preserve the information of fast recovering traps at the $\mathrm{Si} - \mathrm{SiO}_2$ interface and in the oxide. That is why, all tests covering stress signals for mixed HCD and NBTI stress were also performed with the ultra-fast measurement equipment.

Hot-carrier measurement results on a 130 nm dual gate oxide MOS transistor technology node, which is used for automotive and analog applications with a nominal voltage of 3.3 V, show two main degradation effects with one drift type ($D_{\mathrm{Isub,max}}$) close to the traditional hot-carrier degradation worst-case condition and another ($D_{\psi,\mathrm{max}}$) for $V_{\mathrm{ds}} = V_{\mathrm{gs}}$. Both effects compensate the drift after a specific stress time. The drifts and their compensation are clearly observable by analyzing the change of the substrate current characteristics over stress time. In the literature several mechanisms for hot-carrier

degradation have been reported. The first effect is related to the bond dissociation caused by a single high energetic carrier while the second one is due to the multiple vibrational excitation of the bond by several "colder" carriers. The results underline the importance of that approach and provide a benchmark for device degradation simulations due to the good separability of the observed effects. Long term stress data show that even for low V_{gs} the drift type $D_{\mathrm{Isub,max}}$ will be compensated by $D_{\psi,\mathrm{max}}$.

In addition to that, measurement results of the same technology, but on transistors with a nominal voltage of 1.5 V, in the form of threshold voltage drift plots, recovery traces and continuous capture emission time (CET) maps including NBTI and HCD are presented. Devices of $l = 100\,\mathrm{nm}$, $w = 10\,\mu\mathrm{m}$ have been stressed using a 2-dimensional parameter space of gate and drain voltage combinations at 398 K. The chosen stress conditions include the homogeneous ($V_{\mathrm{ds}} = 0$) and inhomogeneous ($V_{\mathrm{ds}} \ll V_{\mathrm{nom}}$) NBTI case, the pure HCD ($V_{\mathrm{gs}} < V_{\mathrm{nom}}$) case as well as the mixture of NBTI and HCD. The results clearly show that for increasing $V_{\mathrm{ds}} > V_{\mathrm{nom}}$ NBTI recovery becomes less severe and mainly the permanent degradation due to HCD endures after the end of stress. Furthermore, there is a drift minimum of NBTI observable for a specific V_{ds}. Using CET maps it is quite evident that for high stress times the probability density of emission becomes very small whereas for shorter stress times there is a recoverable component notable.

Because most reliability models describe the average degradation of a transistor, also information about the variability before and after NBTI stress was recorded and analyzed. Additionally, the permanent and recoverable parts of the degradation were separated. This enables a more precise prediction of the drift over lifetime of the circuit. As the variability of a transistor increases inversely proportional to its gate area, the obtained information is very important for aging simulations of transistors found in SRAM cells or CPU cores.

Contrary to other methods, in this work a new approach of CET map data fitting was developed. Here, the recovery traces and their stress time dependence are fitted in a single step. As a result, not only the fitting algorithm is very efficient but also the accuracy of the fit is very high. Due to the three dimensional fitting approach (stress time, recovery time, threshold voltage drift) the final CET maps are mathematically continuous and allow an extrapolation to reasonable time scales. The finally presented CET map model allows the stress and recovery time, the gate- and drain-source voltages as well as the stress temperature as input parameters for the calculation of the threshold voltage drift and can be used in circuit simulators due to its analytic formulation.

In a cooperative study the NBTI degradation of p-MOSFETs was experimentally and theoretically investigated due to analog stress voltages, which goes beyond existing NBTI studies for digital stress. As a result, a physics-based compact model for analog stress

NBTI was proposed which builds upon the extensive TCAD analysis of the ultra-short-delay experimental data. The numerical efficiency of the compact model allows its direct coupling to electric circuit simulators and permits to accurately account for NBTI degradation already during circuit design. This compact model enables the calculation of the time-dependent NBTI variability of single device parameters and of circuit performance parameters.

1. Degradation mechanisms

Nowadays several degradation mechanisms in CMOS technologies are known. The most critical ones for transistor reliability are hot-carrier degradation (HCD), negative and positive bias temperature instability (NBTI, PBTI) and time-dependent dielectric breakdown (TDDB). Because HCD and BTI are in the focus of this work the following sections will give a brief overview of these two degradation mechanisms, consider the history of their models and explain the physics behind those mechanisms as well as their effect on the electrical behavior of MOSFETs.

1.1. Hot-carrier degradation

Hot-carrier degradation (HCD) describes all mechanisms which lead to a change of the transistor parameters (e.g. threshold voltage, drain-source current) caused by the build-up of interface states at or near the Si/SiO$_2$ interface. “Hot” carriers which have gained sufficient high energy through the electric field dependent acceleration between the source and drain side can change the microscopic structure of a MOSFET (i.e. Si-H bond breakage). “Sufficient” in this case depends on a concrete HCD model. It may refer to the energy to overcome the potential barrier at the Si/SiO$_2$ interface or to the energy to trigger a the Si-H bond rupture. Dependent on their energy single carriers can trigger different mechanisms. For rather low energies about 1.12 eV light emission can be observed [1–3] at the drain side of the p-n junction of an n-MOSFET. Carriers with an energy above the ionization threshold (1.3 eV) can produce electron-hole pairs through the interaction with orbital electrons of Si lattice atoms, a process called impact ionization. If the energy exceeds the SiO$_2$ barrier height an injection of carriers into the oxide takes place. For electrons this barrier energy is lower (3.1 eV) than for holes (4.8 eV). Prima facie, this suggests the assumption that the injection of electrons is more probable than that of holes. Nevertheless, once the carriers are injected into the oxide the amount of trapped holes is higher compared to trapped electrons because of their lower mobility [4]. The created interface traps with density N_{it} can capture electrons and holes and become charged. N_{it} is a distributed quantity and varies with the coordinate along the Si/SiO$_2$ interface as well as in energy. As a result, the electrostatics of the

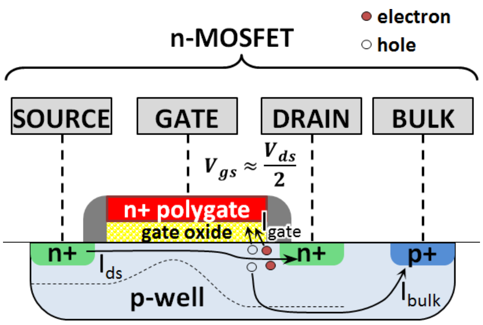


Figure 1.1: Drain avalanche hot-carrier (DAHC): A high energy incident carrier creates an electron-hole pair through impact ionization. Following, the hole and electron are either accelerated by the gate-drain field towards the oxide or the hole contributes to the substrate current. If the carriers directed to the oxide have a sufficiently high energy (electron: 3.1 eV, hole: 4.8 eV) they can overcome the SiO_2 barrier. Because the mobility of holes is lower compared to electrons their probability of becoming trapped in the oxide is higher. The trapped carriers build strongly localized volume charges inside the oxide and interface thereby changing the device characteristics. The dashed line symbolizes the depletion layer.

transistor are changed by these localized charges.

Several hot-carrier degradation mechanisms are described in the literature: Channel hot-carrier (CHC), drain avalanche hot-carrier (DAHC), secondary generated hot-carrier (SGHC), substrate hot-carrier (SHC), Fowler-Nordheim and direct tunneling injection. CHC is linked to the acceleration of the carriers by the electric field in the channel of the transistor. For DAHC and SGHC a solitary hot carrier causes a cascade of impact ionization events. This process leads to the avalanche generation of electron-hole pairs. SHC describes the injection of carriers from the channel-substrate p-n junction into the Si/SiO_2 interface. Fowler-Nordheim tunneling is present if a carrier is injected into the SiO_2 conduction band via the triangular potential barrier caused by a steep potential. In contrast to that, direct tunneling describes the process when carriers are injected through a trapezoidal barrier which is caused by a shallow potential. Historically the description of these mechanisms is based on single carrier processes.

DAHC [5] is most pronounced at $V_{\mathrm{gs}} \approx V_{\mathrm{ds}}/2$. Here, the charge carriers are accelerated from the source towards the drain side and collide there with Si lattice atoms (Fig. 1.1). On the one hand, a highly energetic electron can enter the conduction band of Si and transfer its energy to a valence band electron. On the other hand, highly energetic holes can enter the valence band and transfer its energy to a valence band electron. Because of the excessive energy the valence band electron is excited to the conduction band and as a result of this an electron-hole pair is generated. These additional free carriers at the drain side either (i) can be accelerated by the drain-source field, collide with another Si lattice atom and produce a further electron-hole pair (avalanche multiplication), (ii) are accelerated by the gate-drain field, have sufficient energy to overcome the SiO_2 barrier and (a) get trapped in the oxide or interface or (b) create a gate current or (iii) create a substrate current. The strongly localized volume charge (ii a) leads to a change of the

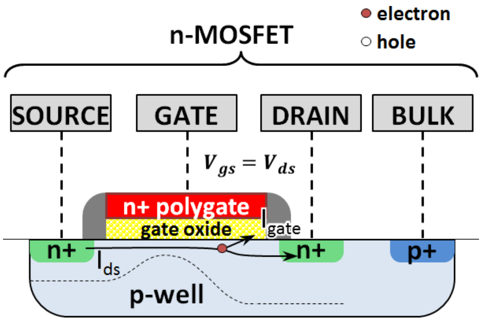


Figure 1.2: Channel hot-carrier (CHC): The charge carriers being accelerated from the source to the drain can be deflected towards the gate oxide because of the high gate-source voltage. If their energy is high enough they directly get injected into the gate oxide.

device characteristics because of their local contribution to the electric field at the drain side.

Channel hot-carrier [6, 7] has its worst-case at $V_{\text{gs}} \approx V_{\text{ds}}$. Due to the high gate voltage the carriers, which are accelerated from the source to the drain, are deflected towards the gate oxide (Fig. 1.2). If their energy is high enough they can overcome the SiO_2 barrier and get trapped while changing the device characteristics.

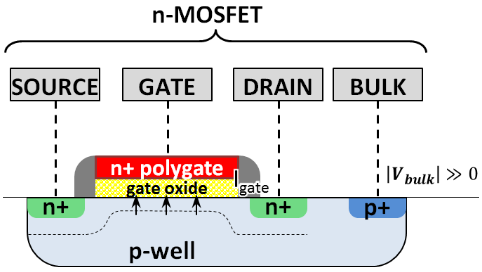


Figure 1.3: Substrate hot-carrier (SHC): For high absolute substrate voltages the gate-substrate field accelerates carriers towards the oxide where they are injected into the gate oxide if their energy exceeds the SiO_2 barrier.

Substrate hot-carrier [8] occurs if the absolute substrate bias is strongly elevated (Fig. 1.3). The high gate-substrate field enables a direct injection of electrons or holes into the gate oxide and their subsequent trapping.

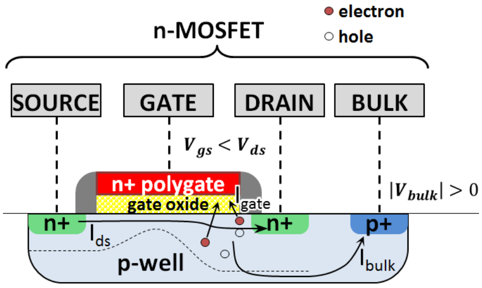


Figure 1.4: Secondary generated hot-carrier (SGHC): Free charge carriers which were generated by impact ionization are accelerated by the gate-substrate field towards the oxide. If their energy is higher than the SiO_2 barrier they are injected into the oxide and can get trapped.

Secondary generated hot-carrier [9] is possible if $V_{\text{ds}} > V_{\text{gs}}$ and $|V_{\text{bulk}}| > 0$ and can be interpreted as a mixture of DAHC and SHC. Once an electron-hole pair was generated by impact ionization the back bias dependent gate-substrate field accelerates one charge carrier towards the oxide (path of lower electron in Fig. 1.4). The carriers which are trapped in the oxide ($\approx 10^{-10}$ of all carriers) cause a drift of the device characteristics.

Additionally, the temperature has a huge impact on the magnitude of hot-carrier degradation. Often, transistors with a channel length above 0.25 µm show an increased hot-carrier degradation at low temperatures while devices with a channel length below 65 nm are more sensitive to high temperature stress. But also the degradation of short channel transistors can be pronounced at low temperatures [10]. Only the consideration of the interplay between various mechanisms can explain the complex phenomenon of hot-carrier degradation [10–14].

Since more than three decades hot-carrier degradation models employed the electric field dependences inside the transistor. Unfortunately, most of these empirical models fail if they are applied to state of the art transistors with channel lengths below a few hundred nanometer, special doping profiles (e.g. lightly or medium doped drain, pocket or halo implants) or nitrided oxides. Nowadays, due to the scaling of MOSFETs, single carriers may not reach the required energy to directly dissociate the Si-H bond. Therefore, the interaction of multiple carriers needs to be taken into account to explain the observed device degradation. Modern hot-carrier degradation models are energy driven, consider multi vibrational excitation and self heating effects. One of the most recent approaches incorporates the evaluation of the carrier energy distribution function (as described in Section 1.1.6). This method shows a very good agreement between measured and simulated data for older as well as for newer technologies as it takes the microscopic structure of the MOSFET into account.

The following sections will present the previously mentioned hot-carrier degradation effects and the history of their corresponding models in detail.

1.1.1. Hu's Lucky Electron Model

In 1985 a noticeable type of MOSFET degradation was related to the presence of hot-electrons in the transistor channel [15]. Former studies related channel hot-electron emission [6] to be the cause of degradation through electron trapping in the oxide using the gate current I_g as a monitor. In contrast to that, it has been shown that the generation of interface traps, which are surface states, is the main reason of device degradation [16–18]. Using a model which "is at best a hypothesis" and "at worst [...] a working phenomenological model" [15] it was found out that "hot" electrons ("not hot holes" [15]) with energies larger than 3.7 eV evidently lead to a MOSFET degradation caused by interface-states generation ("rather than electron trapping" [15]) as shown in Fig. 1.5. At the time when the Hu model was proposed the authors expected the bond energy (the average bond dissociation energy of the system) of $\equiv \mathrm{Si_sH}$ to be about 0.3 eV following [19]. Nowadays, this is known to be wrong. The dissociation of the Si-H bond occurs either at 1.5 eV and $\hbar\omega = 0.075\,\mathrm{eV}$ (bending vibrational mode) or 2.5 eV and $\hbar\omega = 0.25\,\mathrm{eV}$ (stretching vibrational mode) [20, 21]. Experimental investigations suggest that the desorption of H from the Si/SiO$_2$ interface occurs with a bond-breakage energy of $E_a = 2.56\,\mathrm{eV}$ [22, 23]. Adding the barrier energy (3.2 eV) to the wrongly assumed bond energy of 0.3 eV one finds a value close to $\phi_\mathrm{it} = 3.7\,\mathrm{eV}$. The generated interface-states are located very closely ($< 100\,\mathrm{nm}$) to the drain and their number is proportional to the resulting shift in the linear threshold voltage ΔV_th [24], the reduction in the transconductance in the linear or saturation region Δg_m [17] and the shift in the sub-threshold current swing ΔS [16]. In addition, a time dependent power-law formulation of hot-carrier degradation, $\Delta V_\mathrm{th} \propto t^\mathrm{n}$, was proposed.

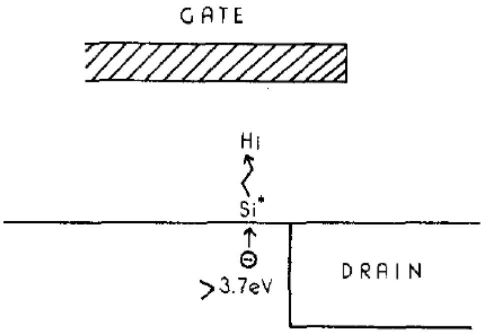


Figure 1.5.: The concept of the lucky electron model as proposed by the group of Hu in 1985 with $V_\mathrm{gs} < V_\mathrm{ds}$. [15]

Hot-electron effects were explained by the group of Hu using the following model equations. The substrate current is known to follow the relation

$$I_\mathrm{sub} = C_1 I_\mathrm{d} \mathrm{e}^{-\beta_\mathrm{i}/E_\mathrm{m}} = C_1 I_\mathrm{d} \mathrm{e}^{-\varphi_\mathrm{i}/q\lambda E_\mathrm{m}} \tag{1.1}$$

where $C_1 \approx 2$ is weakly dependent on E_m, the maximum electric field of the channel, processing parameters, the geometry and $\beta_\mathrm{i} = \varphi_\mathrm{i}/\,(q\lambda) = 1.7 \times 10^6\,\mathrm{V\,cm^{-1}}$. $\lambda\,[\mathrm{cm}]$ is the mean-free path of the hot electron and the energy which is needed by a hot electron to trigger an impact ionization event is given by φ_i. To gain the energy φ_i an electron must travel the distance $-\varphi_\mathrm{i}/qE_\mathrm{m}$ in the electric field E_m. Hence, $\mathrm{e}^{-\varphi_\mathrm{i}/qE_\mathrm{m}}$ is the probability of one electron to gain energy φ_i without suffering a collision. This concept was already demonstrated by W. Shockley for bulk phenomena [25]. The rate of hot electrons is described by Hu et al. to be directly dependent on the current of cold electrons given by I_d. Because of that, the gate current should be written as

$$I_\mathrm{g} = C_2 I_\mathrm{d} \mathrm{e}^{-\varphi_\mathrm{b}/q\lambda E_\mathrm{m}} \tag{1.2}$$

with φ_b as the barrier energy at the Si/SiO$_2$ interface and $C_2 = 2 \times 10^{-3}$ for $V_\mathrm{g} > V_\mathrm{d}$ [7]. If the drain voltage exceeds the gate voltage the polarity of the electric field of the oxide repulses electrons from the gate. In general, to express the number of interface traps generated by a known drain current I_d and device width W one could write

$$\Delta N_\mathrm{it} = C_4 \left(t \frac{I_\mathrm{d}}{W} \mathrm{e}^{-\varphi_\mathrm{it}/q\lambda E_\mathrm{m}} \right)^n , \tag{1.3}$$

with t and n describing the time behavior of the degradation. For lifetime predictions (1.3) can be rewritten to calculate the time at which ΔN_it reaches a defined limit

$$\tau = C_5 \frac{W}{I_\mathrm{d}} \mathrm{e}^{\varphi_\mathrm{it}/q\lambda E_\mathrm{m}} \tag{1.4}$$

where C_5 contains C_4, n and the limit ΔN_it. Using the relation $I_\mathrm{sub} \propto \mathrm{e}^{1.3/q\lambda E_\mathrm{m}}$ and fits of experimental data of a $W/L = 100\,\mu\mathrm{m}/2\,\mu\mathrm{m}$ n-MOSFET with a 35.8 nm thick oxide it was shown for the threshold voltage shift ΔV_th that $\tau \propto I_\mathrm{sub}^{-2.9} I_\mathrm{d}^{1.9} \Delta V_\mathrm{th}^{1.5} W$.

The used proportionality factor C_5 varies for different technologies by over a factor of 100 (Fig. 1.6). Generally, this means for short or wide channel transistors with e.g. a small L, a large V_ds or a small V_gs, resulting in a large I_sub, the lifetime will be short. The experimental data nicely showed that the gate current increases linearly with the substrate current and independently of the chosen device parameters and bias. With the slope being $\varphi_\mathrm{b}\left(E_\mathrm{ox}\right)/\varphi_\mathrm{i}$ the critical electron energy for the impact ionization was determined to $\varphi_\mathrm{i} = 1.3\,\mathrm{eV}$[15]. Consequently the lifetime was identified to be inversely proportional to the substrate current which is dependent on the applied gate-source, drain-source voltage or the chosen channel length.

As a model for the electric field of the channel is very important to understand the mechanisms causing the degradation of the device a simplified version [26] is presented. The potential of the channel can be expressed as an exponential function of the lateral position y [27–29]

$$V\left(y\right) \approx V_\mathrm{ds,sat} + U\mathrm{e}^{y/L} \tag{1.5}$$

$$E\left(y\right)=\left|\frac{\mathrm{d}V}{\mathrm{d}y}\right|=\frac{U}{L}\mathrm{e}^{y/L}=\frac{V\left(y\right)-V_{\mathrm{ds,sat}}}{L} \tag{1.6}$$

with the potential $V_{\mathrm{ds,sat}}$ at the “pinchoff” point ($y=0$), $U \ll V_{\mathrm{ds,sat}}$ and the channel length L. For the drain side where $E=E_{\mathrm{m}}$ and $V=V_{\mathrm{ds}}$ Eq. (1.6) becomes

$$E_{\mathrm{m}}=\frac{V_{\mathrm{ds}}-V_{\mathrm{ds,sat}}}{L}=\frac{V_{\mathrm{ds}}-V_{\mathrm{ds,sat}}}{\sqrt{3X_{\mathrm{ox}}X_{\mathrm{j}}'}} \tag{1.7}$$

where X_{j}' represents a rough estimator of the junction depth. Hence, the *lucky electron model* assumes “short-channel” technologies to show smaller degradation by a reduction of the supply voltage only because $V_{\mathrm{ds,sat}}$, which is proportional to V_{gs}, is a function of the effective channel length [15].

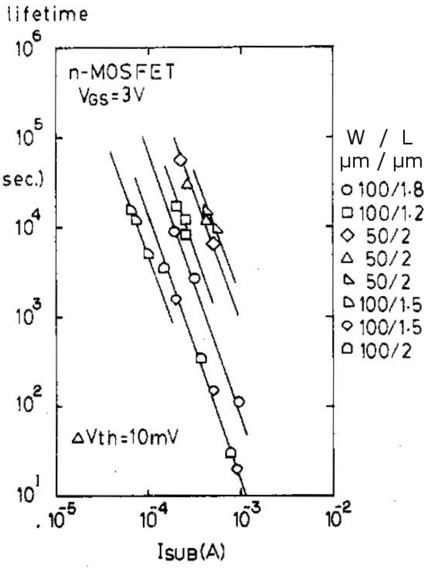


Figure 1.6.: The proportionality constant of the device lifetime is strongly dependent on the characterized technology. Modified from [15].

As pointed out in the overview about hot-carrier degradation (Section 1.1) this empirical model is only valid for a limited range of stress conditions including the stress voltages, temperature and device geometry because the microscopic structure of the transistor is not taken into account and only the contribution of single carriers to the degradation is described. It can be used to understand the results of the single carrier related degradation regime of the analog p-MOSFET characterized in this work with a rather long channel $l = 370\,\mathrm{nm}$ presented in Section 4.2, Fig. 4.13. Here, as typically for p-MOSFETs, the maximum effect was obtained for V_{gs} set to a value where the maximum gate current can be observed, which is about $-0.75\,\mathrm{V}$ and independent of the chosen V_{ds} (Fig. 4.2). But it fails to explain any hot-carrier degradation mechanisms which are related to multiple carrier processes. Hu et al. stated that, due to the increase of the substrate current by a high drain-source voltage or a small channel length, the lifetime will be shortened. Following this single-carrier model no degradation will occur at drain-source voltages which do not sufficiently accelerate the electron above the 3.7 eV barrier energy to break the Si-H bond. Therefore, this model cannot explain the degradation of modern short channel transistors with very low operating voltages at which the 3.7 eV can never reached by a solitary carrier. The first model which follows a different approach and takes into account the information about the energy distribution of defect generation as well as multiple particle processes will be presented in the next section.

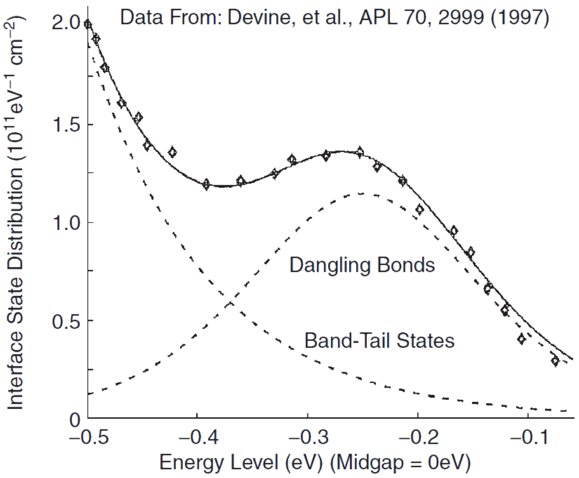


Figure 1.7.: The distribution half-width of the electronic energy levels at the interface is about 0.1 eV. Therefore, the variations in the defect activation energies are expected to be similar or larger. [11]

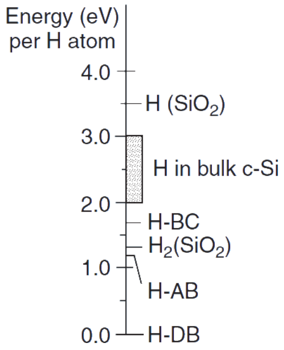


Figure 1.8.: Energy needed to release hydrogen to different locations. Dependent on the position of the hydrogen release the mean energy for hydrogen desorption changes. Therefore one has to consider a bimodal distribution of the activation energy. [11]

1.1.2. Hess Model

A physics-based statistical model for hot-carrier induced degradation (HCD) was presented in 2001 by Hess et al. [11]. The model allows to quantify the time-dependence of several transistor degradation mechanisms like HCD, negative bias temperature instability (NBTI) and time-dependent dielectric breakdown (TDDB). In addition, it has been shown that the statistical fluctuations in the degradation of sub-0.1 µm devices require new standards to determine appropriate design parameters (e.g. operating voltage). Compared to other hot-carrier models, here, the presence of two competing mechanisms, single- and multiple-electron related degradation was introduced for the first time.

The interface states which are created during HCD change the electrostatics of the transistor in two ways. On the one hand, the trapping of negative charges at the drain side leads to an increase of the local resistance, and consequently, a decrease of the drain-source current. On the other hand, the scattering by the interface charge causes a reduction of the local carrier mobility close to the interface [11]. Significant variations of the electronic energy levels of the defects, which are located at hydrogen passivated $\mathrm{Si}-\mathrm{SiO}_2$ interface, have been observed (see. Fig. 1.7) and are supported by theoretical models [30]. This is why also similar variations ($\approx 1\,\mathrm{eV}$) in the defect activation energies are expected.

Furthermore, disorder-induced variations in the $\mathrm{Si}-\mathrm{H}$ activation energies are expected. Comparable to measurements of the thermally induced decay of fiber Bragg gratings, the best fit for the distribution of activation energies can be obtained by a

Fermi-derivative function [31]. Because low-energy bonds are activated more quickly than high-energy bonds, a sublinear time-dependence of device degradation is supposed by Hess et al.. This theory of defect activation enables the derivation of an analytical formula for a time-dependent hot-carrier degradation of MOSFETs [11],

$$\mathcal{F}_{\mathrm{HCI}}\left(t\right)=\frac{\Delta D_{\mathrm{it}}}{D_{\mathrm{it0}}}\approx\frac{p}{1+\left(t/\tau_{1,\mathrm{IT}}\right)^{-\alpha_{1,\mathrm{HCI}}}}+\frac{1-p}{1+\left(t/\tau_{2,\mathrm{IT}}\right)^{-\alpha_{2,\mathrm{HCI}}}} \tag{1.8}$$

with the interface defect density D_{it0}, the hot-carrier activated defect-density ΔD_{it}, $\tau_{\mathrm{i,IT}}$ as lifetime constants, $\log\tau_{\mathrm{i,IT}}=ME_{\mathrm{A,IT}}^{(\mathrm{mean},i)}/\left(qV_{\mathrm{DD}}\right)+C$ (Arrhenius relation), power-law exponents $\alpha_{\mathrm{i,HCI}}\approx qV_{\mathrm{DD}}/\sigma_{\mathrm{i,IT}}M\ln\left(10\right)$ and a technology-dependent constant M.

Two main reasons explain the distribution in defect activation energies, the variation of bond energies due to $\mathrm{SiO_2}$ and $\mathrm{Si/SiO_2}$ disorder as well as the availability of multiple pathways to activation. On the one hand, it was shown via electron spin resonance experiments (ESR) that thermal activation of hydrogen from a passivated defect confirms a distribution in the bond energies of 0.08 eV [32]. On the other hand, different activation pathways, which can be simulated by Density Functional Theory [33], show a wide range of possible activation energies to desorb the hydrogen into the $\mathrm{SiO_2}$ (3.5 eV) and for a final hydrogen state close to the bulk ($< 3\,\mathrm{eV}$).

A single electron can induce transistor degradation if it has a very high energy. This is the case where the maximum substrate current is observed (i.e. $V_{\mathrm{gs}}\approx V_{\mathrm{ds}}/2$, Fig. 1.1). On the contrary also multiple electron collisions can cause degradation and are expected to take place for the maximum drain current (i.e. $V_{\mathrm{gs}}=V_{\mathrm{ds}}$).

Hess et al. chose a bimodal formulation of the activation energy in their model. This is due to the fact that the mean energy for the hydrogen desorption depends on whether the hydrogen is released closer to the silicon bulk or to the dielectric (see. Fig. 1.8) [11]. Then, the failure probability of a transistor is given by the finite binomial probability

$$\begin{aligned}F_{\mathrm{HCI}}\left(E_{\mathrm{A,IT}}\left(t\right)\right)&=\sum_{n=N_{\mathrm{f}}}^{n=N_{\mathrm{H}}}\begin{pmatrix}N_{\mathrm{H}}\\n\end{pmatrix}\left[\mathcal{F}_{\mathrm{HCI}}\left(E_{\mathrm{A,IT}}\right)\right]^{n}\left[1-\mathcal{F}_{\mathrm{HCI}}\left(E_{\mathrm{A,IT}}\right)\right]^{N_{\mathrm{H}}-n}\\&\approx\sum_{n=N_{\mathrm{f}}}^{n=N_{\mathrm{H}}}\frac{\left[N_{\mathrm{H}}\mathcal{F}_{\mathrm{HCI}}\left(E_{\mathrm{A,IT}}\right)\right]^{n}\mathrm{e}^{-N_{\mathrm{H}}\mathcal{F}_{\mathrm{HCI}}\left(E_{\mathrm{A,IT}}\right)}}{n!}\end{aligned} \tag{1.9}$$

with the initial number of $\mathrm{Si-H}$ bonds at the interface $N_{\mathrm{H}}=W\int\rho_{\mathrm{H}}\left(x\right)\mathrm{d}x_{0}^{L_{\mathrm{eff}}}$, the number needed for the device to fail $N_{\mathrm{f}}=W\int\rho_{\mathrm{it}}\left(x\right)\mathrm{d}x_{0}^{L_{\mathrm{eff}}}$, ρ_{H} and ρ_{it} the density of $\mathrm{Si-H}$ bonds and interface traps until failure. The shape of the failure probability is a Weibull distribution as a function of $E_{\mathrm{A,IT}}\left(t\right)$ [34].

The slope of the Weibull distribution is affected by the width of the activation energy distribution [35]. If the damage caused by hot-carriers is located at the $\mathrm{Si}-\mathrm{SiO}_2$ interface and independent of other defect generation events, the failure function for a given number of bonds is

$$F\left(E_{\mathrm{A}}\left(t\right)\right)=\sum_{n=N_{\mathrm{f}}}^{n=N_{\mathrm{H}}}\left(\begin{array}{c}N_{\mathrm{H}}\\ n\end{array}\right)\left[\int^{E_{\mathrm{f}}}\rho\left(E_{\mathrm{A}}\right)\mathrm{d}E_{\mathrm{A}}\right]^{n}\left[1-\int^{E_{\mathrm{f}}}\rho\left(E_{\mathrm{A}}\right)\mathrm{d}E_{\mathrm{A}}\right]^{N_{\mathrm{H}}-n} \tag{1.10}$$

The smaller a transistor is designed the fewer defects are needed to cause failure. If one keeps the defect percentage constant, small devices show a much shorter lifetime compared to larger devices. In addition, the exponent of the time power law has a huge impact on the device lifetime [35].

The strongest degradation is observed in materials without long range order (amorphous, e.g. SIO_2) [35]. In disordered solids the barriers for ionic motion are lower for ions at a surface or an interface. Therefore, amorphous materials have a larger range of barrier energies for ionic motion. Furthermore, localized electron states, which are rather present at surfaces and in amorphous solids, can transfer energy from electronic degrees of freedom better to localized ionic degrees of freedom than delocalized electronic states of crystals [35].

The large disparity in electron mass and ionic mass makes it difficult for electrons to transfer their energies to the Si-H bonds. Therefore, the energy needed for a single electron to cause bond-breakage can only be achieved if the electron is localized near to the bond (resonance scattering) or if it transfers its energy to a valence electron which passes the energy to the bond (electronic excitation) [35]. As a consequence of both processes the potential energy surface of the bond is changed. On the one hand the bonding state can be delocalized and the valence excitation would be an impact ionization process. On the other hand the electron can be localized and the excitation would be able to transfer a significant amount of energy to the bond [35]. Scanning tunneling microscopy induced desorption of hydrogen from a silicon (100) surface shows that the cross section initially is an exponential function of the electron energy and subsequently saturates [36]. This direct way to transfer a large energy amount to a bond was chosen in [35] to model the degradation process.

If one assumes independent defect creation events and **single electrons** to be responsible for bond-breaking events, the rate equation for this process is

$$\frac{\mathrm{d}N}{\mathrm{d}t}=\frac{\left(N_0-N\right)}{\tau} \tag{1.11}$$

$$N\left(t\right)=1-\mathrm{e}^{t/\tau} \tag{1.12}$$

with lifetime $\tau=\left(\int I\left(E_{\mathrm{e}}\right)\sigma\left(E_{\mathrm{e}}\right)\mathrm{d}E_{\mathrm{e}}\right)^{-1}$, where $I\left(E_{\mathrm{e}}\right)$ is the density of electrons (cm^{-2}),

σ is the cross section of the single defect creation, N the number of defects and N_0 the number of defect precursors.

The lucky electron model, which is based on a single carrier degradation mechanism, fails to adequately explain the degradation of scaled devices because it cannot explain the damage of devices with voltages below an activation energy of 3.7 eV. In addition, the activation energy has a large isotope effect ($\approx 1\,\mathrm{eV}$) in the model of Hu et al.[35]. In the lucky electron model hot-carrier degradation is only related to solitary "hot"-carriers which have sufficient energy to induce impact ionization. Thus, impact ionization and interface trap generation were used synonymously.

Nowadays, device scaling results in lower voltages and may increase the current density in the transistor channel [37]. Hence, there is a larger number of low energy electrons available which can transfer their energy to the lattice. Single high energy electrons become rather rare and cannot solely explain the device degradation. Delocalized excitations will merely result in diffuse lattice heating. Hence, the *localized excitation* of vibrational modes of a bond has to be the dominant degradation process where traps are created by **multiple electrons** of low energy above the phonon mode energy and with a typical current density above $100\,\mathrm{nA\,nm^{-2}}$. The lifetime of a multiple carrier model can be written like [38]

$$\frac{1}{\tau} = \sum_{\mathrm{i}}^{n} \int I\left(E_\mathrm{e}\right) \sigma_\mathrm{d}\left(E_\mathrm{e}\right) \mathrm{d}E_\mathrm{e} A^i \left(\frac{\int I\left(E_\mathrm{e}\right) \sigma_\mathrm{in}\left(E_\mathrm{e}\right) \mathrm{d}E_\mathrm{e} + \frac{\mathrm{e}^{\frac{\hbar\omega}{k_\mathrm{B}T}}}{\tau_\mathrm{p}}}{\int I\left(E_\mathrm{e}\right) \sigma_\mathrm{in}\left(E_\mathrm{e}\right) \mathrm{d}E_\mathrm{e} + \frac{1}{\tau_\mathrm{p}}} \right)^i \tag{1.13}$$

where $\int I\left(E_\mathrm{e}\right) \sigma_\mathrm{d}\left(E_\mathrm{e}\right)$ is the rate of electronic excitation, A the probability enhancement for higher phonon modes, $\int I\left(E_\mathrm{e}\right) \sigma_\mathrm{in}\left(E_\mathrm{e}\right)$ the rate of phonon excitation, T the lattice temperature, $\hbar\omega$ the phonon mode energy and τ_p the phonon mode relaxation time. The phonon excitation P_u and decay P_d rates can be obtained from the formalism presented in [39]

$$P_\mathrm{u} \approx \int_{E_\mathrm{th}}^{\infty} I\left(E\right) \sigma_\mathrm{emi}\left(E\right) \left[1 - f_\mathrm{ph}\left(E + \hbar\omega\right)\right] \mathrm{d}E \tag{1.14}$$

$$P_\mathrm{d} \approx \int_{E_\mathrm{th}}^{\infty} I\left(E\right) \sigma_\mathrm{ab}\left(E\right) \left[1 - f_\mathrm{ph}\left(E - \hbar\omega\right)\right] \mathrm{d}E \tag{1.15}$$

with σ_emi and σ_ab as the phonon emission and absorption cross sections and the phonon occupation numbers $f_\mathrm{ph}\left(E\right)$. These two phonon process related rates lead to the bond-breakage rate for the multiple carrier process

$$R_{\mathrm{MP}} = \left(\frac{E_{\mathrm{B}}}{\hbar\omega} + 1\right)\left[P_{\mathrm{d}} + \mathrm{e}^{\frac{-\hbar\omega}{k_{\mathrm{B}}T_{\mathrm{L}}}}\right]\left[\frac{P_{\mathrm{u}} + \omega_{\mathrm{e}}}{P_{\mathrm{d}} + \mathrm{e}^{\frac{-\hbar\omega}{k_{\mathrm{B}}T_{\mathrm{L}}}}}\right]^{-\frac{E_{\mathrm{B}}}{\hbar\omega}} \tag{1.16}$$

where E_{B} represents the energy of the last bonded level in the quantum well (Fig. 1.9), ω_{e} the phonon reciprocal lifetime and T_{L} the lattice temperature.

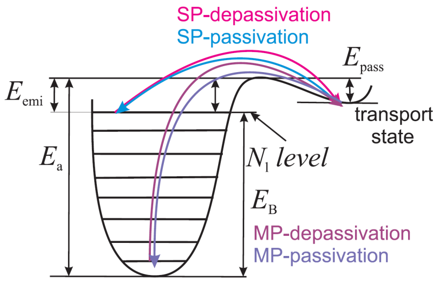


Figure 1.9.: The interaction of single- and multiple-carrier mechanisms with the Si-H bond represented by a harmonic oscillator. [40]

Using a multiple carrier degradation model the observed [38] low threshold damage and isotope effect can be simply explained because the excited phonon modes reduce the barrier height and the phonon modes of deuterium relax much faster than hydrogen [35]. In addition, the physical interpretation for the higher damage of n-MOSFET compared to p-MOSFET transistors can be explained easily by the accessibility of resonant states for carriers which excite phonon modes of the $\mathrm{Si}-\mathrm{H}$ bond. For holes the accessible state is deeper located in the valence band than for electrons. Therefore, it is more difficult for them to cause the excitation of the $\mathrm{Si}-\mathrm{H}$ bond, resulting in a lower damage of the p-MOSFET. To describe the hot-carrier degradation of a specific transistor properly one needs to find the correct carrier energy distribution function because it is implicitly presented by the particle flux differential $I\left(E\right)$. Once the carrier energy distribution function is determined (as presented in Section 1.1.6) also the results of this work, which are related to multiple carrier degradation mechanisms, can be explained on a physical basis. Of special interest are the two degradation regimes of the analog p-MOSFET, which can be selected by the gate-source stress voltage (see Fig. 4.13 and Fig. 4.14).

1.1.3. Penzin Model

The time-dependent trap formation based degradation model presented by Penzin et al. [41] in 2003 adapted the statistical degradation model from the group of Hess for TCAD device simulations with the help of phenomenological approximations. The model is based on a simple first order kinetic equation and a $\mathrm{Si}-\mathrm{H}$ concentration dependent formulation of the activation energy. In addition, the model includes the electrical field dependence of different degradation mechanisms. Although the diffusion of hydrogen away from the $\mathrm{Si/SiO_2}$ interface is neglected, it can be considered by a higher order term in the kinetic equation. In order to predict the device lifetime, a simulation of the trap generation is needed, which has to accurately reflect the physics of the interface trap formation process. Because 10^{12} silicon dangling bonds per cm^2 are expected [42] at the interface and hydrogen or deuterium are used to compensate them during fabrication of high quality $\mathrm{Si} - \mathrm{SiO}_2$ interfaces, the electro-chemical trap formation process needs to be understood. Measurements of the interface trap concentration N_{it} show that $\log\left(N_{\mathrm{it}}\right)$ is almost proportional to $\lg\left(t\right)$ and many models are available to explain and fit this relation. The sub-linearity can be related to the disorder-induced variations of the bond activation energies [43] and the corresponding time-dependences have been explained by hydrogen diffusion and its possible recovery [15, 44] or a mobility reduction caused by the generated interface traps [45, 46]. Anyhow, the Penzin Model explains this effect by considering the dependence of the $\mathrm{Si} - \mathrm{H}$ activation energy on its density and uses a field dependence similar to [47]. This setup is implemented into a general purpose device simulator to solve the enhanced trap formation kinetics equation. When Penzin et al. published their model in 2003, it was unknown how many trap levels correspond to each stress-generated dangling bond, and also the energetic properties of the generated traps had not been investigated in detail before [48, 49]. Therefore, the following simplification was used,

$$N_{\mathrm{it}} = N - n, \tag{1.17}$$

where N_{it} is the concentration of interface traps. The concentration of the silicon dangling bonds consists of N (total concentration of Si bonds appearing as dangling ones if hydrogen leaves the bond) and n (concentration of $\mathrm{Si} - \mathrm{H}$ bonds at the interface). The time dependence of trap generation was experimentally confirmed [50] to be

$$N_{\mathrm{it}} - N_{\mathrm{it}}^0 = \frac{n_0}{1 + (kt)^{-\alpha}} \tag{1.18}$$

$$\Leftrightarrow n = n_0 \left(1 + (kt)^{\alpha}\right)^{-1}, \tag{1.19}$$

where the initial concentrations of $\mathrm{Si} - \mathrm{H}$ bonds and interface traps are given by n_0 and N_{it}^0 and k and α are fitting parameters.

A variation of the hydrogen concentration from n_0 to n, taking the potential barrier change $\epsilon_0 - \epsilon_n$ into account, leads to the activation energy which is needed to release hydrogen from the interface. Because the bond can be stretched or pressed by the electric field (after [47] this is valid for $\mathrm{Si} - \mathrm{O}$ and according to [41] this can also be applied to $\mathrm{Si} - \mathrm{H}$) a correctional term was introduced leading to the final expression of the activation energy

$$\epsilon_{\mathrm{A}} = \epsilon_{\mathrm{A}}^0 + \delta\left|F\right|^{\rho} + \beta k_{\mathrm{B}} T \ln\left(\frac{N - n}{N - n_0}\right). \tag{1.20}$$

Here, the energy needed to break a Si − H bond is given by ϵ_A^0. The last summand with prefactor β is the potential energy needed to overcome the potential barrier of the 2-D potential system, $|F|$ is the modulus of the electric field, $\beta = 1 + \beta_\perp F_\perp$, δ and ρ are two fitting parameters. The fitting parameter $\beta_\perp = 0.5\epsilon_{\text{ins}} S_{\text{eff}}/q$ is dependent on the insulator permittivity ϵ_{ins} (the system is considered to be a capacitor) and the effective area of one hydrogen atom S_{eff}. The polarity dependent perpendicular component of the electric field to the interface is given by $F_\perp$.

The trap distributions for negative and positive gate bias stress as well as for two types of hot carrier stress are plotted along with a comparison to experimental data in Fig. 1.10. The Penzin Model well adapts different slopes for the time dependence of the trap formation process while the parameter $\beta_\perp$ controls the time-dependence of the degradation slope, δ and ρ shift the degradation curve.

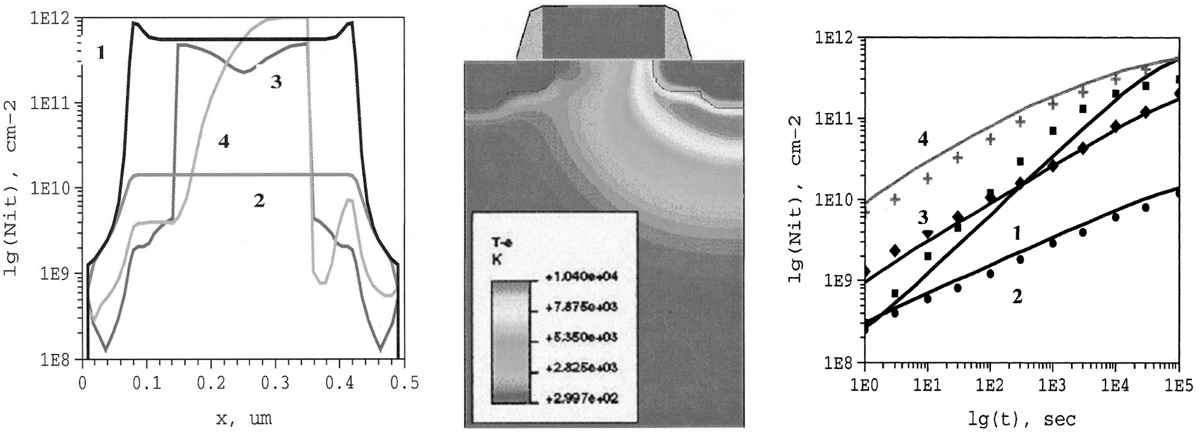


Figure 1.10.: TCAD simulation results of the Penzin model.
Left: The trap distribution along the interface of an n-MOSFET ($l = 350\,\text{nm}$, $t_{\text{ox}} = 6.5\,\text{nm}$) for four different stress types (1: $V_g = -9\,\text{V}$, $V_d = V_s = V_b = 0\,\text{V}$; 2: $V_g = 12\,\text{V}$, $V_b = 0\,\text{V}$, source and drain floating; 3: $V_g = 1\,\text{V}$, $V_b = -11\,\text{V}$, $V_d = V_s = 0\,\text{V}$; 4 (typ. HCD): $V_d = 5.5\,\text{V}$, $V_g = 2.5\,\text{V}$ and $V_s = V_b = 0\,\text{V}$).
Middle: The electron temperature for stress type 4 ranges from $3 \times 10^2\,\text{K}$ to $10^4\,\text{K}$. The extracted parameters are (from field stresses 1&2) $\beta_\perp = 10^{-7}\,(\text{V/cm})^{-1}$, $\rho = 0.33$, $\delta = 1.95 \times 10^{-3}\,(\text{V/cm})^{-\rho}$ and (from hot carrier stresses 3&4) $\delta_{\text{HC}}|_{\rho_{\text{HC}}=1} = 6\times10^5\,\left(\text{A/cm}^2\right)$. A nonuniform spatial dependence of the degradation is observable with the maximum at the drain end as expected for hot carrier stress.
Right: Comparison between the experimental (symbols) and simulation data (lines) of an n-MOSFET stressed with the conditions 1 to 4. The highest degradation occurs for the typical worst-case hot-carrier stress condition 4 ($V_{\text{gs}} \approx V_{\text{ds}}/2$). [41]

1.1.4. Reaction-Diffusion Framework

The Reaction-Diffusion (R-D) Framework presented by Kufluoglu and Alam in 2004 [51, 52] distinguishes between the degradation time-exponents of NBTI ($t^{1/4}$) and HCD ($t^{1/2}$) (Fig. 1.11) caused by the breakage of Si – H bonds using a geometrical approach. They use the classical R-D model [15, 53] as presented in Fig. 1.12 and speculate that the hydrogen diffusion mechanism of HCD has a 2-D nature.

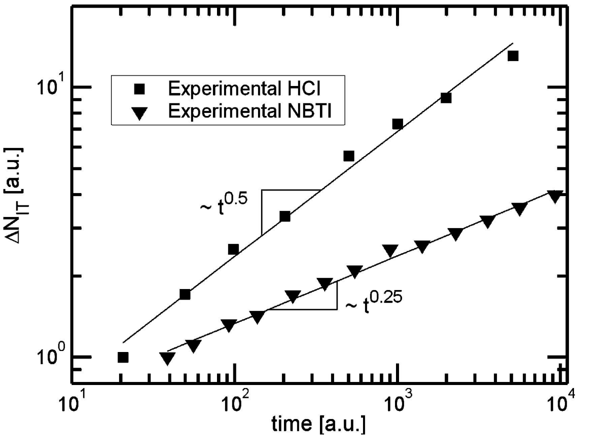


Figure 1.11.: The experimental interface trap density has NBTI and HCD like power-law exponents. For NBTI it is $t^{1/4}$ whereas for HCD one can extract $t^{1/2}$. The wrong time power law exponent of NBTI originates from a too long delay between stress and measurement during which the degradation partially recovered. Data taken from [15, 53]. [51]

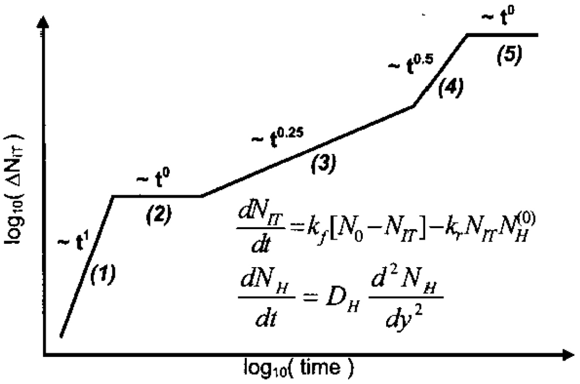


Figure 1.12.: The R-D model consists of five regions with different time behavior following the inset equations (D_H: diffusion constant, k_f: forward dissociation rate, k_r: reverse annealing rate). Please note that only region (3) can be seen experimentally during degradation measurements followed by a saturation of the drift without the intermediate region (4). [52]

The classical R-D model [53] consists of five different regimes of interface trap generation [52],

1. N_IT increase due to Si – H bond breaking
2. diffusion of hydrogen begins to take over
3. trap creation is limited by hydrogen diffusion
4. hydrogen diffuses in the gate with infinite diffusion velocity
5. maximum trap density, N_0, is reached, and N_IT no longer increases with time

Here NBTI and HCD are related to region 3 and 4 because of their $\sim t^{1/4}$ and $\sim t^{1/2}$ degradation behavior. Although it has not been proven experimentally, according to this model, NBTI should show a $\sim t^{1/2}$ dependence for long and HCD a $\sim t^{1/4}$ dependence for short time scales. From an opposing point of view NBTI measurements indicate that the $t^{1/2}$ region does not exist [54] due to the limited speed of hydrogen diffusion in the oxide. Therefore, the classical R-D model will yield wrong results. In order to fix this problem, the authors proposed that a geometrical correction for the degradation mechanisms has to be included in the model. For NBTI the interface trap distribution over the channel is homogeneous. But for the 1-dimensional problem of HCD, where the defect generation is located at the drain side, the detached hydrogen can diffuse 2-dimensionally into the oxide. Contrary to the assumptions of the reaction diffusion framework all the NBTI results which are presented in this work (Section 4.3) show that the $\sim t^{1/4}$ of NBTI has to be related to a too long measurement delay (see also Fig. 4.34). Hence, the reaction diffusion framework fails to explain the degradation of the transistor.

For

$$N_{\mathrm{it}}\left(t\right) = \left(\frac{1}{A_{\mathrm{d}}}\right) \int N_{\mathrm{H}}\left(r,t\right) \mathrm{d}^3 r \tag{1.21}$$

with the hydrogen diffusivity D_{H}, the degradation area A_{d} and the hydrogen density at the interface $N_{\mathrm{H}}^{(0)}$ the interface trap densities for NBTI and HCD are given by

$$N_{\mathrm{it,NBTI}}\left(t\right) = \left(\frac{1}{A_{\mathrm{d}}}\right) \int_0^{\sqrt{D_{\mathrm{h}} t}} N_{\mathrm{H}}^{(0)} \left[1 - \frac{r}{\sqrt{D_{\mathrm{H}} t}}\right] A_{\mathrm{d}} \mathrm{d}r = \left(\frac{1}{2}\right) N_{\mathrm{H}}^{(0)} \sqrt{D_{\mathrm{H}} t} \tag{1.22}$$

$$N_{\mathrm{it,HCD}}\left(t\right) = \left(\frac{\pi}{2A_{\mathrm{d}}}\right) \int_0^{\sqrt{D_{\mathrm{h}} t}} N_{\mathrm{H}}^{(0)} \left[r - \frac{r^2}{\sqrt{D_{\mathrm{H}} t}}\right] \mathrm{d}r = \left(\frac{\pi}{12 A_{\mathrm{d}}}\right) N_{\mathrm{H}}^{(0)} \left(D_{\mathrm{H}} t\right) \tag{1.23}$$

If the density of interface traps and hydrogen at the interface is kept constant ($N_{\mathrm{it}} N_{\mathrm{H}}^{(0)} =$ const.) this leads with the net rate of the interface trap density increase,

$$\frac{\mathrm{d}N_{\mathrm{it}}}{\mathrm{d}t} = k_{\mathrm{f}} \left(N_0 - N_{\mathrm{it}}\right) - k_{\mathrm{r}} N_{\mathrm{it}} N_{\mathrm{H}}^{(0)} \tag{1.24}$$

$$\frac{\mathrm{d}N_{\mathrm{H}}}{\mathrm{d}t} = D_{\mathrm{H}} \frac{\mathrm{d}^2 N_{\mathrm{H}}}{\mathrm{d}y^2}, \tag{1.25}$$

where k_{f} is the forward dissociation and k_{r} the reverse annealing rate to

$$N_{\mathrm{it,NBTI}} \sim \sqrt{\frac{k_{\mathrm{f,NBTI}} N_0}{k_{\mathrm{r}}}} \left(D_{\mathrm{H}} t\right)^{1/4}, \tag{1.26}$$

$$N_{\mathrm{it,HCD}} \sim \sqrt{\frac{k_{\mathrm{f,HCD}} N_0}{k_{\mathrm{r}}}} \left(D_{\mathrm{H}} t\right)^{1/2}. \tag{1.27}$$

In addition to the wrong power law constant of NBTI, the R-D framework also predicts a full recovery of HCD within observable time scales. Anyhow, all measurements in this work show that at operational temperatures no recovery of HCD occurs and other technologies show only very slow HCD recovery. Therefore, HCD has to be considered as a reaction-limited process [40]. Although the calculations are correct, the physical basis and some of the chosen parameter values of the R-D framework are wrong and far away from reality. That is why, the R-D framework is not applicable for a correct determination of the degradation.

1.1.5. Bravaix Model

A model based on a combination of the ideas from Hess, Rauch and La Rosa et al. (the “energy-driven paradigm”, [55]) was presented by the group of Bravaix in 2009 [20, 56]. Here, both the interplay of single and multiple particle processes, which was introduced by Hess, and the substitution of the electron energy distribution function by empirical stress-dependent factors, following the approximation of Rauch and La Rosa, are included in a “general framework about defect creation at the Si/SiO2 interface”. This framework reflects “the reliability of metal oxide semiconductor field effect transistors, and particularly negative bias temperature instability permanent part, and channel hot carrier to cold carrier damage”[20].

Degradation takes place in a MOSFET if incident carriers (i.e. channel carriers (n-MOSFET: electrons, p-MOSFET: holes), tunneling gate carriers and substrate carriers) exceed the potential barrier of the oxide, a critical energy. The anharmonic coupling, an adsorbate-induced resonant state, between the adsorbate and the incident carrier can be modeled by a Lorentzian density of states (DOS) [57–59]. For electrons this is the $\mathrm{Si}-\mathrm{H}$ $6\sigma*$ [58, 59] and for holes the 5σ [57] state. The shift of the Fermi levels between the incident carriers and the substrate is shown in Fig. 1.13. First, the carrier energy is transmitted to the bond via inelastic scattering and after that dissipated by adsorbate vibrational modes. For $\mathrm{Si}-\mathrm{H}$ these can either be stretching or bending modes [60] which have a bond breaking energy E_{B}, a vibrational mode energy $\hbar\omega$ and a relaxation time τ_{e}.

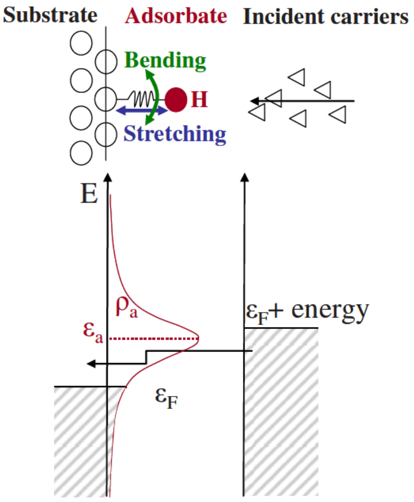


Figure 1.13.: The adsorbate (H) has two different vibrational modes, bending and stretching, each with a bond breaking energy E_{B}, vibrational mode energy $\hbar\omega$ and relaxation time τ_{e}. The incident carriers at energy ϵ_{a} and DOS ρ_{a} near the Fermi level of the substrate and the adsorbate are anharmonically coupled via an adsorbate-induced resonance state. This allows an energy transfer from the incident carriers to the bending or stretching bond vibrational modes. [20]

The bond breaking energy can be taken from tabular data provided by several groups [60–63] and is about 2.5 eV (with $\hbar\omega_{\mathrm{s}} = 0.25\,\mathrm{eV}$) for the stretching mode and 1.5 eV (with $\hbar\omega_{\mathrm{b}} = 0.075\,\mathrm{eV}$) for the bending mode. Also, the bond vibrational life time can be derived from infrared absorption spectroscopy (transient bleaching technique). The derived values at 5 K are $w_{\mathrm{e,stretching}} = 1/295\,\mathrm{ps}^{-1}$ and $w_{\mathrm{e,bending}} = 1/10\,\mathrm{ps}^{-1}$. Note that Bravaix et al. chose a constant value for the vibrational life time, although it is temperature dependent as discussed in [64]. In addition, it is assumed that the bending mode is related only to the

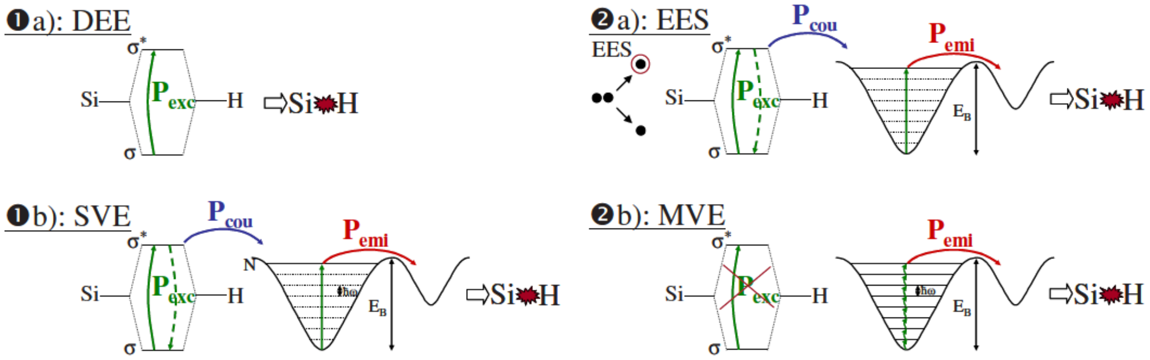


Figure 1.14.: Four carrier-induced excitation related Si–H bond breaking modes are possible: DEE (direct electron excitation), SVE (single vibrational excitation) (with or without EES (electron-electron scattering)) and MVE (multiple vibrational excitation). The coupling and thermal emission from the Nth level lead to probabilities of resonance excitation of the Nth level occupation, P_{exc}, P_{cou} and P_{emi}. [20]

sum of transversal optical ($460\,\mathrm{cm}^{-1}$) and transversal acoustic ($150\,\mathrm{cm}^{-1}$) phonon modes (following the observation that the decay time increases exponentially with increasing number of phonons needed for the decay N_i [65]).

The physical framework of Bravaix et al. considers four different ways of Si – H bond breaking for hot-carrier degradation (see. Fig. 1.14). Two of them are related to single high-energetic incident carriers which have enough energy to excite the adsorbate resonance state, direct electron excitation (DEE) and single vibrational excitation (SVE). If a fraction of the Si – H bonds stays long enough in the excited resonance state it can induce a direct bond breakage and is called DEE. The resulting rate of bond breaking is defined as the product of the carrier energy distribution function with the excitation probability P_{exc} of the carriers. The integral of the carrier energy distribution function over the energy is proportional to the particle current,

$$R_{\mathrm{b,DEE}} = IP_{\mathrm{exc}}, \mathrm{HCD}: I = I_{\mathrm{ds}} \tag{1.28}$$

The energy of the Si-H bond experiencing a resonance decay will be transferred to low-energy modes via an anharmonic coupling. A competing process at that stadium is the fast relaxation caused by the emission of bulk phonons. The coupling probability P_{cou} between the resonance state and the localized vibrational modes is related to the occupation probability of the last level N (which is induced by the resonance state decay if only this level contributes to the hydrogen desorption). This special case, where in a single jump (i.e. a single carrier) a multi-step vibrational excitation takes place is called SVE. Here, for the Nth level the variation in the occupation density n_N with stress time t is defined as

$$\frac{\mathrm{d}n_N}{\mathrm{d}t} = IP_{\mathrm{exc}}P_{\mathrm{cou}}n_0 - w_{\mathrm{e}}\,[\mathrm{Si}^*]\,[\mathrm{H}^*] \tag{1.29}$$

with the ground state density n_0 and the total relaxation time $w_\mathrm{e} = \tau_\mathrm{e}^{-1}$. The ensemble rapidly becomes stationary because the interface trap generation is orders of magnitude slower compared to the bond relaxation time (s vs. ps) leading to

$$\frac{\mathrm{d}n_N}{\mathrm{d}t} = 0 \Rightarrow [\mathrm{Si}^*] = \Delta N_\mathrm{it} = \frac{I P_\mathrm{exc} P_\mathrm{cou} n_0}{w_\mathrm{e}\,[\mathrm{H}^*]}. \tag{1.30}$$

The hydrogen atom in the Nth level can be dissociated but also a reformation can take place by excitation decay. A dissociation of the hydrogen atom from silicon is only possible if it is transmitted to transport states. The probability of the emission of the hydrogen is P_emi. For an attempt frequency ν hydrogen can be emitted by thermal activation over the weak barrier E_emi from the excited state with a probability λ per unit time, i.e.

$$\lambda = \nu \mathrm{e}^{\frac{-E_\mathrm{emi}}{k_\mathrm{B}T}} \Rightarrow P_\mathrm{emi} = 1 - \mathrm{e}^{-\lambda t}. \tag{1.31}$$

Experimental data [20] suggest that the emission time λ^{-1} is much longer than the stress time leading to the approximation

$$P_\mathrm{emi} \approx \lambda t = t\nu \mathrm{e}^{\frac{-E_\mathrm{emi}}{k_\mathrm{B}T}}. \tag{1.32}$$

Because the creation of interface traps is related to P_emi, implying $\Delta N_\mathrm{it} = [H^*]\,\lambda t$, the total number of interface traps is given by

$$\Delta N_\mathrm{it} = \left(\frac{n_0 \lambda I P_\mathrm{exc} P_\mathrm{cou}}{w_\mathrm{e}} t\right)^{1/2}. \tag{1.33}$$

For long stress times the time exponent ($1/2$) will decrease because of the saturation of degradation. For that case the number of generated interface traps is no longer negligible compared to the total number of Si-H bonds n_0. The lifetime τ until a critical amount of interface traps is generated can be related to SVE by

$$\Delta N_\mathrm{it,crit} = (R_\mathrm{bSVE}\tau) \tag{1.34}$$

with the bond breaking rate $R_\mathrm{bSVE} \propto (n_0 \lambda I P_\mathrm{exc} P_\mathrm{cou})/w_\mathrm{e}$. Also electron-electron scattering (EES), which causes a second knee in the carrier energy distribution function, can be taken into account for degradation. Because the carrier energy density function is no longer proportional to I but to I^2 for EES, one has to be aware that $R_\mathrm{bEES} \propto (n_0 \lambda I^2 P_\mathrm{exc} P_\mathrm{cou})/w_\mathrm{e}$.

If the excitation cannot be induced by a single carrier (including EES), the energy barrier E_B can still be overcome by multiple low-energy incident carriers. Each of these carriers contributes to a multiple-step vibrational ladder with $\hbar\omega$. To accomplish this multiple vibrational excitation (MVE) the number of electrons, which “interact” with the bond per second, has to be high enough to produce excited states. Also, the lifetime of the excited states has to be long enough to let them contribute to the multiple-step

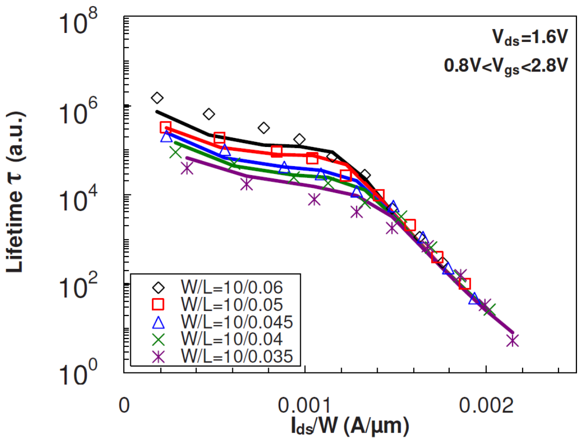


Figure 1.15.: The comparison between experimental data (symbols) and the model (lines) shows that in contrast to electron-electron scattering (EES) for multiple vibrational excitation (MVE) the gate-length dependence becomes less severe and only the I_{ds} magnitude is important for the lifetime prediction. [20]

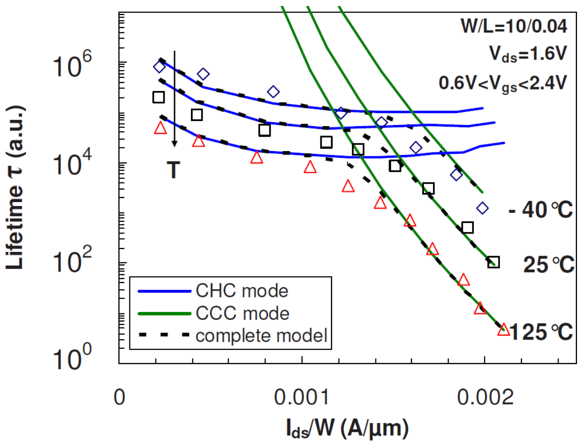


Figure 1.16.: The model (lines) consisting of channel hot carrier (CHC) and channel cold carrier (CCC) modes is in good agreement with the data of $L = 40\,\mathrm{nm}$ n-MOSFETs with $T_{\mathrm{ox}} = 1.7\,\mathrm{nm}$ stressed at three different temperatures (symbols). [20]

vibrational ladder. To describe this competing process the empirical factor S_{MVE}, the "inelastic tunneling probability" [21], is used to describe the "up" and "down" rates of the phonon excitation (compare with Eq. 1.14 and 1.15) can be expressed by

$$P_{\mathrm{u}} = S_{\mathrm{MVE}}\,(I_{\mathrm{ds}}/q) + w_{\mathrm{e}}\mathrm{e}^{\frac{-\hbar\omega}{k_{\mathrm{B}}T}} \tag{1.35}$$

$$P_{\mathrm{d}} = S_{\mathrm{MVE}}\,(I_{\mathrm{ds}}/q) + w_{\mathrm{e}} \tag{1.36}$$

leading to the final breaking rate

$$R_{\mathrm{bMVE}} = n_0\nu \left[\frac{S_{\mathrm{MVE}}\left(\frac{I_{\mathrm{ds}}}{q}\right) + w_{\mathrm{e}}\mathrm{e}^{\frac{-\hbar\omega}{k_{\mathrm{B}}T}}}{S_{\mathrm{MVE}}\left(\frac{I_{\mathrm{ds}}}{q}\right) + w_{\mathrm{e}}}\right]^{\frac{E_{\mathrm{B}}}{\hbar\omega}} \mathrm{e}^{\frac{-E_{\mathrm{emi}}}{k_{\mathrm{B}}T}}. \tag{1.37}$$

The proposed model for both SVE (including EES (Fig. 1.15)) and MVE is in good agreement with experimental data (Fig. 1.16). Devices with ultra-thin (1.7 nm) and thin (3 nm) oxides (to prevent electron trapping known for medium and thick oxides at high gate biases [66, 67]) and a channel length of 40 nm were stressed at various drain-source and gate-source voltages. To neglect the contribution of fast interface traps the delay before readout was set to 60 s. At low drain-source currents ($V_{\mathrm{gs}} = 0.8\,\mathrm{V}$) SVE and EES are causing the degradation by high-energy carriers and the influence of the chosen stress drain-source voltage is high on the resulting lifetime. High drain-source currents ($V_{\mathrm{gs}} = 2.8\,\mathrm{V}$) show only degradation caused by MVE (low-energy carriers) and therefore, the influence of the drain-source voltage is much smaller. Following, the two regimes are called channel hot carrier (CHC) and channel cold carrier (CCC) mode by the authors. In addition, the data depict that for MVE the lifetime is reduced with increasing temperature and the energy impact is still present at high drain-source currents because shorter channels show a higher impact-ionization ratio, $R_{\mathrm{ii}} = \frac{I_{\mathrm{bs}}}{I_{\mathrm{ds}}}$.

Finally an “Age” function is proposed by the group of Bravaix considering all three competing degradation mechanisms (SVE, EES, MVE):

$$\begin{aligned}\mathrm{Age} = \frac{t}{\tau} = t \Bigg[& K_{\mathrm{SVE}} \left(\frac{I_{\mathrm{ds}}}{W}\right)^{a_1} \left(\frac{I_{\mathrm{bs}}}{I_{\mathrm{ds}}}\right)^{m} + K_{\mathrm{EES}} \left(\frac{I_{\mathrm{ds}}}{W}\right)^{a_2} \left(\frac{I_{\mathrm{bs}}}{I_{\mathrm{ds}}}\right)^{m} \\ & + K_{\mathrm{MVE}} V_{\mathrm{ds}}^{a_3/2} \left(\frac{I_{\mathrm{ds}}}{W}\right)^{a_3} \mathrm{e}^{\frac{-E_{\mathrm{emi}}}{k_{\mathrm{B}}T}} \Bigg] \\ = t & \left[\frac{K_{\mathrm{SVE}}}{\tau_{\mathrm{SVE}}} + \frac{K_{\mathrm{EES}}}{\tau_{\mathrm{EES}}} + \frac{K_{\mathrm{MVE}}}{\tau_{\mathrm{MVE}}}\right]\end{aligned} \tag{1.38}$$

with fitting constants a_1, a_2 and a_3, degradation mechanism probabilities K_{SVE}, K_{EES} and K_{MVE} and the body effect coefficient m. Using the Age function the time power law exponent is found to be $1/2$ for all stress conditions.

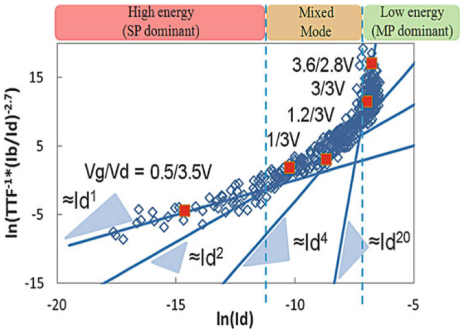


Figure 1.17.: The increase of $V_{\mathrm{gs}}/V_{\mathrm{ds}}$ shows the change of the power law exponent of I_{ds} from CHC to CCC via mixed mode stress. The two and four carrier degradation modes dominate the mixed mode. The data (full symbols) is taken from a $L_{\mathrm{G}} =$ 65 nm n-MOSFET while the open symbols are used to extract $S_{\mathrm{it}}(E) \propto \left(\sqrt[m]{TTF} I_{\mathrm{ds}}\right)^{-1}$. Here m is, roughly speaking, understood as the number of carriers involved in the bond-breakage process. [13]

Recent publications from Alain Bravaix et al., e.g. [13], describe an intermediate degradation mode between SVE and MVE called mixed mode (MM) (see Fig. 1.17). Here, the hot-carrier damage is related to the anharmonic coupling between the stretching and bending mode. If the potential well of the bending mode is divided into several energy levels the time-dependent defect density can be written as ([68], [13])

$$N_{\mathrm{it}} = N_0 \left[1 - \mathrm{e}^{-cR_{\mathrm{MM}}t/2}\right] \tag{1.39}$$

with the multi mode bond-breaking rate R_{MM} as well as the constants N_0 and c being derived from measurement data.

Unfortunately, the model from the group of Bravaix substitutes the carrier energy distribution function with empirical stress-dependent factors which leads to non-universal but technology-dependent constants. Thus, the whole HCD picture with its physical mechanisms cannot completely be described by that model.

1.1.6. Carrier distribution function based hot-carrier model

A physics-based hot-carrier degradation model was proposed and verified for a 5 V n-MOSFET with a channel length of about 0.5 µm in 2010 by Tyaginov and Starkov [69–72]. To describe the interplay between high- and low-energetic electrons the carrier energy distribution function was entirely evaluated employing a full-band Monte Carlo device simulator. Several stress conditions for the degradation of the linear drain current were investigated and show a good agreement between theory and experiment. The simulations exactly reflect the strong localization of hot-carrier degradation with high-energy tails of the carrier energy distribution function at the drain side of the gate where the maximum of the carrier acceleration integral is located. In contrast to previous hot-carrier aging models (e.g. [20, 56]) all links between the microscopic and device level are covered. For the computation of the degradation only few empirical factors (e.g. capture cross section, attempt rate, threshold energy) need to be used.

A three step setup is used in order to describe the mechanisms of defect generation for a specific transistor architecture and stress regime, to calculate the corresponding carrier energy distribution function, time-dependent interface state density profiles $N_{\text{it}}(x)$ and trap density-of-states (DOS) as well as the aged device characteristics (Fig. 1.18).

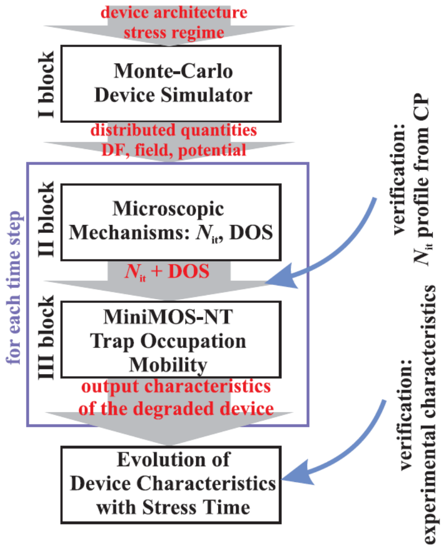


Figure 1.18.: The flowchart of the carrier distribution function based hot-carrier model employing the stochastic Boltzmann transport equation solver consists of three main modules: carrier transport, microscopic mechanisms for defect creation and the simulation of degraded device characteristics. [40]

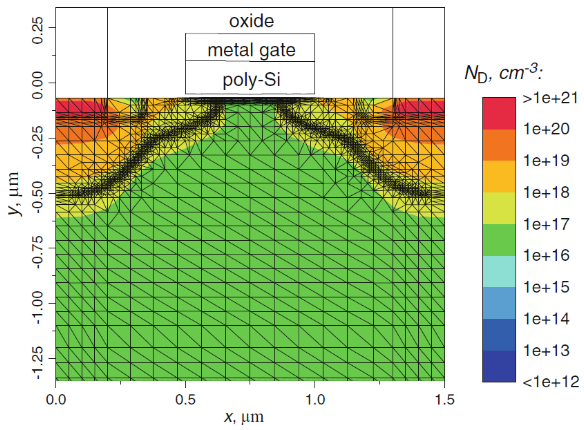


Figure 1.19.: To validate the model the architecture of a 5 V n-MOSFET with $L = 0.5$ µm was used. The color map represents the phosphorous doping profile. [14]

During the first step the carrier transport is identified via a full-band Monte Carlo device simulator (i.e. MONJU [73]) which calculates the solution of the stochastic Boltzmann transport equation. The device simulator considers energy exchange mechanisms as impact ionization, surface scattering, electron-phonon scattering and scattering at ionic impurities. Transistor degradation with emphasized electron-electron scattering cannot be taken into account because this mechanism is not implemented in MONJU.

The resulting carrier energy distribution function is then fed into the device simulator MiniMOS-NT which characterizes the bond-breakage kinetics as interface state densities $N_{\mathrm{it}}\left(x,t\right)$ with the lateral coordinate x and the aging time t.

The most crucial quantity of the presented model is the acceleration integral (AI) which controls the simulated rates of the hot-carrier degradation. It is based on the carrier energy distribution function derived from the full-band Monte Carlo device simulator for single (SP) and multiple particle (MP) mechanisms as well as for electrons (e) and holes (h):

$$I_{\mathrm{SP/MP}}^{(e/h)} = \int_{E_{\mathrm{th}}}^{\infty} f^{(e/h)}\left(E\right) g^{(e/h)}\left(E\right) \sigma_{\mathrm{SP/MP}}^{(e/h)}\left(E\right) \nu\left(E\right) \mathrm{d}E \tag{1.40}$$

with the carrier energy distribution function $f^{(e/h)}\left(E\right)$, the DOS $g^{(e/h)}\left(E\right)$, the Keldysh-like reaction cross section $\sigma_{\mathrm{SP/MP}}^{(e/h)}\left(E\right)$ and the group velocity $\nu\left(E\right)$.

$$\sigma_{\mathrm{SP/MP}}^{(e/h)}\left(E\right) = \sigma_{0,\mathrm{SP/MP}}^{(e/h)}\left(E - E_{\mathrm{th,SP/MP}}\right)^{p_{\mathrm{it}}} \tag{1.41}$$

where $\sigma_{0,\mathrm{SP/MP}}$ is the attempt rate ($p_{\mathrm{it}} = 11$) and the threshold energy for SP and MP processes is set to $E_{\mathrm{th}} = 1.5\,\mathrm{eV}$.

The proposed model considers the SP and MP phenomenons independently which simplifies the calculation effort but neglects their interplay. With the energy barriers E_{emi} for hydrogen emission and E_{pass} for hydrogen passivation and considering the Arrhenius relation the dissociation and passivation rates are defined as

$$R_{\mathrm{MP}} = \nu_{\mathrm{MP,act}} \mathrm{e}^{\frac{-E_{\mathrm{emi}}}{k_{\mathrm{B}}T_{\mathrm{L}}}} \tag{1.42}$$

$$P_{\mathrm{MP}} = \nu_{\mathrm{MP,pass}} \mathrm{e}^{\frac{-E_{\mathrm{pass}}}{k_{\mathrm{B}}T_{\mathrm{L}}}} \tag{1.43}$$

with the attempt rates $\nu_{\mathrm{MP,act}}$ and $\nu_{\mathrm{MP,pass}}$. Omitting the rather slow bond-breakage and -passivation process for N_{l} (last bonded state level) the system is solved recursively to find the occupancy of all levels. As a result of this for prevailing ground state bonds ($N_0 = \sum_i n_i \approx n_0$) the interrelations $\frac{n_i}{n_0} = \left(\frac{P_{\mathrm{u}}}{P_{\mathrm{d}}}\right)^i$ can be written. Reincorporating the bond-breakage and -passivation rates and assuming constant occupation numbers n_i with respect to the slow nature of these processes the solution obtained for the multiple particle mechanism interface state density (for intially fresh bonds) is

$$N_{\mathrm{MP}} = N_0 \left[\frac{R_{\mathrm{MP}}}{P_{\mathrm{MP}}} \left(\frac{P_{\mathrm{u}}}{P_{\mathrm{d}}} \right)_{\mathrm{l}}^{N} \left(1 - \mathrm{e}^{R_{\mathrm{MP}} t} \right) \right]^{\frac{1}{2}} . \tag{1.44}$$

with

$$P_{\mathrm{u}} = \nu_{\mathrm{MP}}^{(e)} I_{\mathrm{MP}}^{(e)} + \nu_{\mathrm{MP}}^{(h)} I_{\mathrm{MP}}^{(h)} + \omega_{\mathrm{e}} \mathrm{e}^{\frac{-\hbar\omega}{k_{\mathrm{B}} T_{\mathrm{L}}}} \tag{1.45}$$

$$P_{\mathrm{d}} = \nu_{\mathrm{MP}}^{(e)} I_{\mathrm{MP}}^{(e)} + \nu_{\mathrm{MP}}^{(h)} I_{\mathrm{MP}}^{(h)} + \omega_{\mathrm{e}}. \tag{1.46}$$

Considering the competing SP and MP modes their contributions can be weighted with particular probabilities:

$$N_{\mathrm{it}} = p_{\mathrm{SP}} N_{\mathrm{SP}} + p_{\mathrm{MP}} N_{\mathrm{MP}} \tag{1.47}$$

The model was validated against the linear drain current degradation of a 5 V n-MOSFET with channel lengths of 0.5 µm, 1.2 µm and 2.0 µm (see Fig. 1.19). Electron energy distribution functions for various hot-carrier stress conditions, the corresponding AI as well as the interface state density profiles and relative HCD contribution of the MP mechanism are plotted in Fig. 1.20.

To correctly describe the degradation for long channel devices also the secondary holes generated by impact ionization and accelerated towards the source side of the device need to be taken into account, because they can either contribute to the SP or MP mechanism. To also include electron-electron scattering (EES) [74] and the interplay between SP and MP mechanisms, the current version of the carrier energy distribution function based hot-carrier degradation model employs a deterministic BTE solver (ViennaSHE [75–78]) in contrast to the previous stochastic method. Here, the distribution function is represented by a spherical harmonical expansion (SHE) of an arbitrary order [74, 76, 79–81], which consumes high amounts of RAM instead of CPU time. That is why, it became practicable during the last few years even on average workstations [76].

The process, which is related to a single hot carrier exciting one of the bonding electrons to anti-bonding state, is called the AB mechanism. Bond dissociation caused by multi-vibrational excitation is called the MVE mechanism. To describe the interplay of the AB and the MVE mechanism correctly, the bond-breakage and -passivation rates

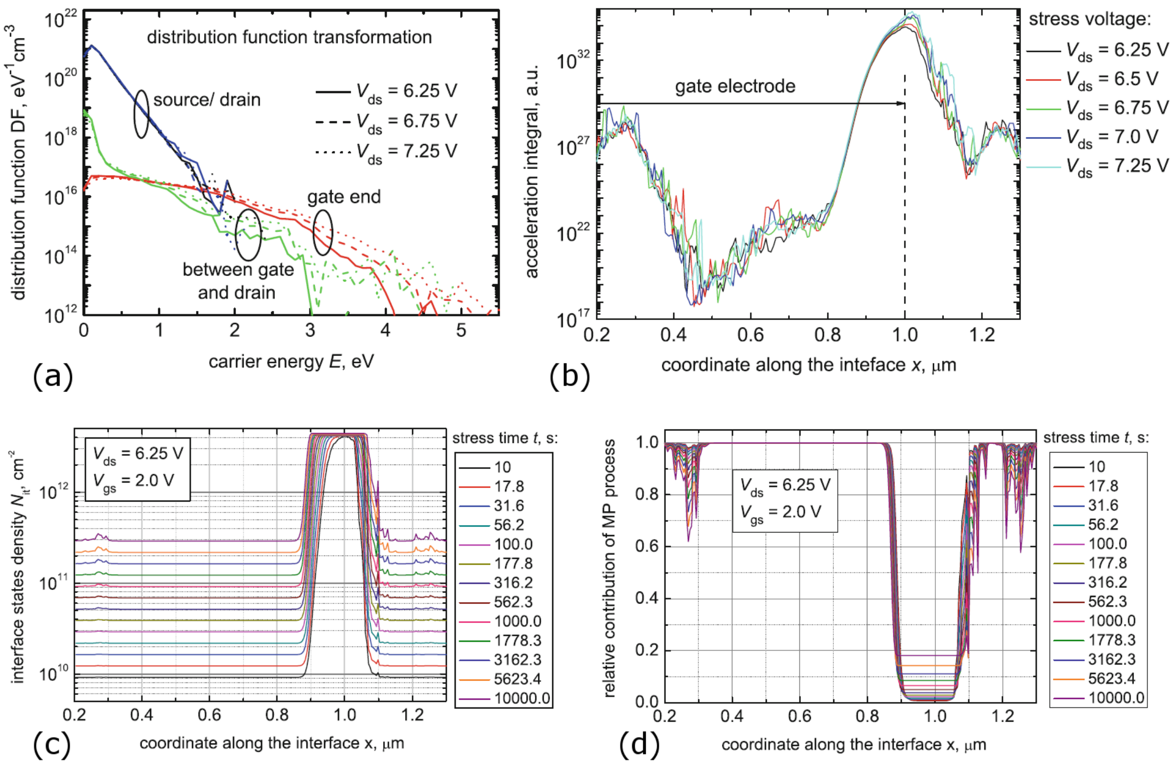


Figure 1.20.: a) For $V_{\mathrm{gs}} = 2.0\,\mathrm{V}$ and $V_{\mathrm{ds}} = 6.25\,\mathrm{V}$, $6.75\,\mathrm{V}$ and $7.25\,\mathrm{V}$ the carrier distribution function has been calculated for the $5\,\mathrm{V}$ n-MOSFET with $L = 0.5\,\mu\mathrm{m}$. The channel of the device is located between $0.5\,\mu\mathrm{m}$ and $1.0\,\mu\mathrm{m}$. Shown are distribution functions at the source and drain end which are close to the Maxwellian distribution, and the two non-equilibrium located in the middle of the channel and beyond the drain-sided gate end.
b) For the same n-MOSFET the carrier AI has been calculated for $V_{\mathrm{gs}} = 2.0\,\mathrm{V}$ and $V_{\mathrm{ds}} = 6.25\,\mathrm{V}$, $6.5\,\mathrm{V}$, $6.75\,\mathrm{V}$, $7.0\,\mathrm{V}$ and $7.25\,\mathrm{V}$. That HCD is a strongly localized mechanism can be seen by the peak of the AI at the drain side.
c) It is obvious that the peak of the interface state density profiles $N_{\mathrm{it}}(x)$ for the stress condition $V_{\mathrm{gs}} = 2.0\,\mathrm{V}$, $V_{\mathrm{ds}} = 6.25\,\mathrm{V}$ is dominated by the SP process because the relative contribution of the MP process (d) shows its saturation in the center of the device and close to the drain. The saturation is caused by the high stress voltages which lead to a coordinate independent interface state density. [14]

need to be considered for each level,

$$\frac{\mathrm{d}n_0}{\mathrm{d}t} = P_{\mathrm{d}} n_1 - P_{\mathrm{u}} n_0 - R_0 n_0 + P_0 N_{\mathrm{it}}^2 \tag{1.48}$$

$$\frac{\mathrm{d}n_i}{\mathrm{d}t} = P_{\mathrm{d}}\left(n_{i+1} - n_i\right) - P_{\mathrm{u}}\left(n_i - n_{i-1}\right) - R_i n_i + P_i N_{\mathrm{it}}^2 \tag{1.49}$$

$$\frac{\mathrm{d}n_{N_{\mathrm{l}}}}{\mathrm{d}t} = P_{\mathrm{u}} n_{N_{\mathrm{l}}-1} - P_{\mathrm{d}} n_{N_{\mathrm{l}}} - R_{N_{\mathrm{l}}} n_{N_{\mathrm{l}}} + P_{N_{\mathrm{l}}} N_{\mathrm{it}}^2. \tag{1.50}$$

The bond-rupture rates R_i of the ith level are defined as

$$R_{\mathrm{a,n_i}} = w_{\mathrm{th}} \mathrm{e}^{\frac{-(E_{\mathrm{a}} - E_i)}{k_{\mathrm{B}} T}} + \nu_{\mathrm{AB}} I_{\mathrm{AB,i}} \tag{1.51}$$

with

$$I_{\mathrm{AB,i}} = \int f(E)\, g(E)\, \sigma (E - E_{\mathrm{a}} + E_i)^p \, \nu (E) \, \mathrm{d}E. \tag{1.52}$$

Following [14] the system (Eq. 1.50) can be reduced to

$$\frac{\mathrm{d}N_{\mathrm{it}}}{\mathrm{d}t} = (N_0 - N_{\mathrm{it}})\, \mathfrak{R} - N_{\mathrm{it}}^2 \mathfrak{P} \tag{1.53}$$

with the cumulative bond-breakage and total passivation rates

$$\mathfrak{R} = \frac{1}{k} \sum_i R_i \left(\frac{P_{\mathrm{u}}}{P_{\mathrm{d}}} \right)^i \tag{1.54}$$

$$\mathfrak{P} = \nu_p \mathrm{e}^{\frac{-E_{\mathrm{pass}}}{k_{\mathrm{B}} T_{\mathrm{L}}}} \tag{1.55}$$

and the normalization factor $k = \sum_i \left(\frac{P_{\mathrm{u}}}{P_{\mathrm{d}}} \right)^i$. This leads to the solution of the system

$$N_{\mathrm{it}}(t) = \frac{(\mathfrak{R}^2/4 + N_0 \mathfrak{R}\mathfrak{P})^{\frac{1}{2}}}{\mathfrak{P}} \frac{1 - f(t)}{1 + f(t)} - \frac{\mathfrak{R}}{2\mathfrak{P}} \tag{1.56}$$

with

$$f(t) = \frac{(\mathfrak{R}^2/4 + N_0 \mathfrak{R}\mathfrak{P})^{\frac{1}{2}} - \mathfrak{R}/2}{(\mathfrak{R}^2/4 + N_0 \mathfrak{R}\mathfrak{P})^{\frac{1}{2}} + \mathfrak{R}/2} \mathrm{e}^{-2t\left(\mathfrak{R}^2/4 + N_0 \mathfrak{R}\mathfrak{P}\right)^{\frac{1}{2}}}. \tag{1.57}$$

For comparison EES was included and ignored for the via ViennaSHE calculated carrier energy distribution functions. The results are plotted in Fig. 1.21 with the corresponding AIs. The comparison to experimental data including curves showing the neglect of several degradation mechanisms is shown in Fig. 1.22. The plots clearly reveal that only the model considering the interplay between SP and MP mechanisms is in good agreement with the experimental data, which emphasizes the importance of the exact solution of the BTE, taking into account all noticeable degradation effects.

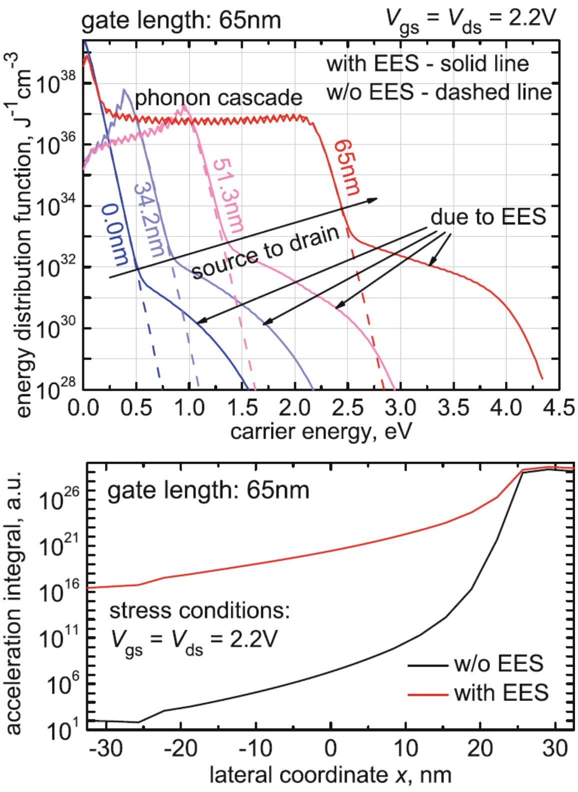


Figure 1.21.: Top: Calculations of the distribution functions for a 65 nm device which was stressed at $V_{gs} = V_{ds} = 2.2\,\mathrm{V}$. It is evident that EES needs to be taken into account to correctly reflect the high-energy tails of this transistor.
Bottom: The contribution of EES leads to a severe increase of the acceleration integral which is pronounced in the areas of rather cold carriers (source and channel center). [14]

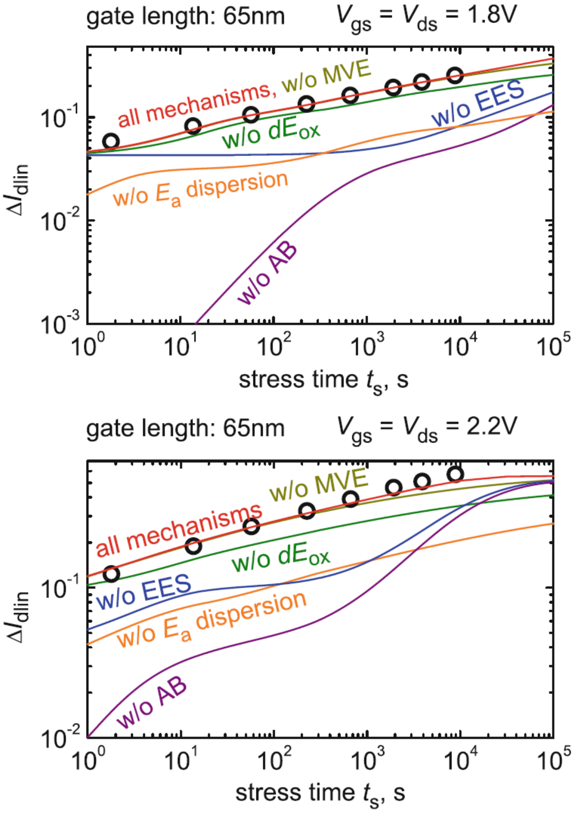


Figure 1.22.: Comparison between the experimental data (symbols) and the simulation (lines) of the linear drain-source current degradation at worst-case conditions ($V_{gs} = V_{ds} = 1.8\,\mathrm{V}$ **(top)** and $2.2\,\mathrm{V}$ **(bottom)**) for a channel length of 65 nm. Only the consideration of all modeled mechanisms (SP, MP, EES, AB, E_a dispersion, activation energy reduction dE_{ox} caused by the electrical field at the interface) leads to a good agreement between experiment and simulation. [14]

1.2. Bias Temperature Instability (BTI)

After its experimental evidence in the 1960s the bias temperature instability (BTI) remained obscure until Nitrogen was introduced for oxide processing around 2000. Through this NBTI for p-MOSFETs became more pronounced. Due to the emerging high-k gate stack technology (excl. Nitrogen) also PBTI in n-MOSFETs became a prominent degradation effect. 50 years later BTI is still a very controversial topic due to contradicting experimental observations, their explanations and physical models with an exponentially growing number of publications.

One decade ago it was realized that the recovery effect has a dramatic impact on the characterization and lead to different physical models. In addition to that, different characterization methods like measure-stress-measure (used for this work), on the fly, fast voltage ramps or poly heater setups [82, 83] allow a quantization of the recovery and facilitate advanced aging models. In the 1980s the monitoring of discrete charge capture and emission events was performed. While at that time large transistors with non-nitrided oxides were analyzed, nowadays, scaled transistors with nitrided oxides or high-k gate stacks are in the focus of research. Like for random telegraph noise (RTN) discrete changes in the threshold voltage are observable. These are caused by the discharge of single defects having a very wide time distribution [84]. The time dependent dielectric spectroscopy (TDDS, see Section 1.2.6) (Reisinger, Infineon) as analysis method yielded numerous results, e.g. that the same traps are responsible for RTN and BTI [85].

The distribution properties, scaling issues and the impact on digital (SRAM) and analog circuits have been discussed by Rauch (IBM) in detail including the extrinsic as well as the intrinsic variability [86]. Comprehensive research has been carried out by Kaczer, Toledano-Luque and Weckx (IMEC) with the so-called defect-centric view which focuses on the single defect contribution to the threshold voltage, its aging and recovery related shift and describes a connection to the 0h-variability [87]. A correct identification of the defects which are subject to the degradation is very important. For this purpose electron spin resonance experiments (ESR) are the state of the art resource [88]. Here, Pb, E' and K centers respond to NBTI stress in SiO2, SiON as well as in high-k gate stacks. Furthermore hydrogen related defects have been identified for a wide range of technologies [89]. Also electrical measurements are able to confirm possible defects in SiO2, SiON, high-k dielectrics as shown by Zhang [90]. Hole trapping and trap transformation during cyclic BTI stress were studied by Ang [91] and show an inconsistency to conventional reaction-diffusion (RD) theory.

To model BTI the theoretical properties of the most frequent observed oxide defects (e.g. E' centers) need to be known. The charging of those defects causes them to interact with the discrete dopants in the channel, causing strong changes of the device characteristics. The very first model from 1977 [19] is based on the RD theory and is being refined until today. The state of the art RD model is being developed by Mahapatra (IIT Bombay) (see Section 1.2.1). However, stochastic chemical kinetics show limitations and problems of the RD theory as discussed by Schanovsky and Grasser (TU Vienna) (see Section 1.2.2). Charge trapping in non-radiative multiphonon (NMP) theory and additional metastable states allow an explanation of the experimental TDDS data and in addition to that, an approximate description of NBTI can be given with a collection of independent first order reactions via capture emission time (CET) maps.

Not only the usage of hydrogen during the process has an influence on BTI but also the associated process steps including nitrogen, flouride and deuterium in SiON and high-k dielectrics [92, 93]. Ultrascaled technologies employ high-k materials to suppress the leakage currents although this causes PBTI in n-MOSFETs. Experimental results of nanoscaled high-k devices were obtained by Toledano-Luque and Kaczer (IMEC) focusing on the stochastic charge trapping and statistical lifetime prediction [94]. Alternative channel materials and device topologies have been suggested to overcome scaling issues. Franco and Kaczer (IMEC) proved that e.g. SiGe channel devices show an improved reliability compared to Si [95].

The impact of BTI on circuits should not be underestimated. A cooperative study between Intel and the University of Minnesota utilized beat frequency detection odometers for the direct on-chip assessment of BTI on circuits for a better measurement and timing control [96]. But also bottom-to-top approaches of reliability analysis from device level to system level aging including statistical modeling of charge trapping in the context of RTN and BTI are in huge interest of the reliability community and one main aspect of the European research project MoRV (**Mo**delling **R**eliability under **V**ariability). With these findings it is possible to translate the calculated NBTI degradation of the electrical properties into circuit performance and reliability with the help of combined SPICE and Monte Carlo simulations [97].

1.2.1. NBTI Reaction Diffusion Framework

In 2013 Mahapatra et al. described a comprehensive modeling framework for DC and AC NBTI [98] which is based on three different parts, a H/H_2 Reaction-Diffusion (RD) model for the as slow processes considered generation and recovery of interface traps [99, 100], a 2-energy level model for the relatively fast hole trapping and detrapping [101, 102] and an empirical formulation for the generation of bulk insulator traps (although their contribution has been found negligible for thinner oxides at use conditions [98, 103]). The authors allege that the modeling framework can explain features of

NBTI like the time evolution of the threshold voltage during DC stress and recovery, long term stress (i.e. high temperature operating lifetime (HTOL) data), AC stress degradation including the pulse frequency and duty cycle, the impact of the measurement delay as well as the degradation dependence on the dielectric process. Unfortunately the model is only applicable with unphysically chosen constants as discussed in section 1.2.2. Interestingly and in contrast to a previous version [51, 52] the model now states NBTI degradation to follow a reduced time power law constant of $^1/_6$. Although the RD model fits the experimental stress data for stress times above $10^4\,\mathrm{s}$ using an unphysical interstitial size of 40 Å (see Fig. 1.27) the predicted recovery behavior does not reflect the measurement results presented in this work (see Fig. 4.34 as well as Fig. 1.25). The $\mathrm{H/H_2}$ RD model is depicted in Fig. 1.23 for the generation of and recovery of interface traps caused by NBTI stress. Due to the gate bias the inversion layer hole tunnels into the interface and can be captured by a Si-H bond. If the bond breaks through thermal excitation [104, 105] the released hydrogen diffuses out and can form molecular $\mathrm{H_2}$ with another diffusing hydrogen atom (conventional $\mathrm{H/H_2}$ model) or with a Si-H bond at the $\mathrm{SiO_2/poly-Si}$ interface. During the recovery phase this $\mathrm{H_2}$ will diffuse back towards the $\mathrm{Si/SiO_2}$ interface, monomerize, react with the broken Si- and repassivate the interface. It is also possible that the $\mathrm{H_2}$ reacts with the broken Si- at the $\mathrm{SiO_2/poly-Si}$ interface. The resulting hydrogen will then also be able to passivate the $\mathrm{Si/SiO_2}$ interface alike. The bond breakage at the $\mathrm{Si/SiO_2}$ interface (s) is given by [93]

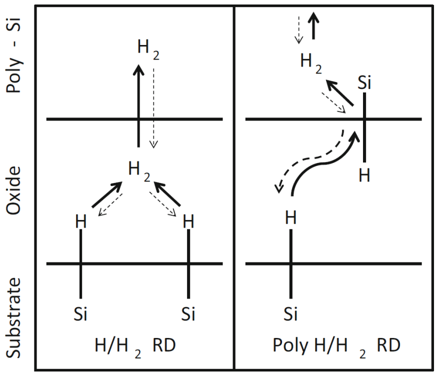


Figure 1.23.: Conventional $\mathrm{H/H_2}$ RD and poly $\mathrm{H/H_2}$ RD models. [93]

$$\frac{\mathrm{d}N_{\mathrm{IT(s)}}}{\mathrm{d}t} = k_{\mathrm{F(s)}}\left(N_{0\mathrm{(s)}} - N_{\mathrm{IT(s)}}\right) - k_{\mathrm{R(s)}}N_{\mathrm{IT(s)}}N_{\mathrm{H}}^{(s)} \tag{1.58}$$

For the conventional model the hydrogen di- and monomerization are expressed as

$$\frac{\delta}{2}\frac{\mathrm{d}N_{\mathrm{H}}^{(s)}}{\mathrm{d}t} = D_{\mathrm{H}}\frac{\mathrm{d}N_{\mathrm{H}}^{(s)}}{\mathrm{d}x} + \frac{\mathrm{d}N_{\mathrm{IT(s)}}}{\mathrm{d}t} - \delta k_{\mathrm{H}}\left[N_{\mathrm{H}}^{(s)}\right]^2 + \delta k_{\mathrm{H2}}N_{\mathrm{H2}}^{(s)} \tag{1.59}$$

$$\frac{\delta}{2}\frac{\mathrm{d}N_{\mathrm{H2}}^{(s)}}{\mathrm{d}t} = D_{\mathrm{H2}}\frac{\mathrm{d}N_{\mathrm{H2}}^{(s)}}{\mathrm{d}x} + \frac{\delta}{2}k_{\mathrm{H}}\left[N_{\mathrm{H}}^{(s)}\right]^2 - \delta k_{\mathrm{H2}}N_{\mathrm{H2}}^{(s)}, \tag{1.60}$$

the creation and dissociation of $\mathrm{H_2}$ at the $\mathrm{SiO_2/poly-Si}$ interface (p) is described by

$$\frac{\mathrm{d}N_{\mathrm{IT(p)}}}{\mathrm{d}t} = k_{\mathrm{F(p)}}\left(N_{0\mathrm{(p)}} - N_{\mathrm{IT(p)}}\right)N_{\mathrm{H}}^{(p)} - k_{\mathrm{R(p)}}N_{\mathrm{IT(p)}}N_{\mathrm{H2}}^{(p)} \tag{1.61}$$

and the diffusion of H and $\mathrm{H_2}$ is defined as

$$\frac{\mathrm{d}N_\mathrm{H}}{\mathrm{d}t} = D_\mathrm{H}\frac{\mathrm{d}^2 N_\mathrm{H}}{\mathrm{d}x^2} - k_\mathrm{H} N_\mathrm{H}^2 + k_\mathrm{H2} N_\mathrm{H2} \tag{1.62}$$

$$\frac{\mathrm{d}N_\mathrm{H2}}{\mathrm{d}t} = D_\mathrm{H2}\frac{\mathrm{d}^2 N_\mathrm{H2}}{\mathrm{d}x^2} + \frac{k_\mathrm{H} N_\mathrm{H}^2}{2} - \frac{k_\mathrm{H2} N_\mathrm{H2}}{2} \tag{1.63}$$

with

$$k_\mathrm{F(s)} = k_\mathrm{F0(s)}\left(V_\mathrm{gs} - V_\mathrm{th0}\right)^{3/2}\Gamma_\mathrm{IT}\mathrm{e}^{-\frac{E_\mathrm{AkF}}{k_\mathrm{B}T}} \qquad k_\mathrm{R(s)} = k_\mathrm{R0(s)}\mathrm{e}^{-\frac{E_\mathrm{Akr}}{k_\mathrm{B}T}} \tag{1.64}$$

$$k_\mathrm{H} = k_\mathrm{H0}\mathrm{e}^{-\frac{E_\mathrm{AkH}}{k_\mathrm{B}T}} \qquad k_\mathrm{H2} = k_\mathrm{H20}\mathrm{e}^{-\frac{E_\mathrm{AkH2}}{k_\mathrm{B}T}} \tag{1.65}$$

$$D_\mathrm{H} = D_\mathrm{H0}\mathrm{e}^{-\frac{E_\mathrm{ADH}}{k_\mathrm{B}T}} \qquad D_\mathrm{H2} = D_\mathrm{H20}\mathrm{e}^{-\frac{E_\mathrm{ADH2}}{k_\mathrm{B}T}} \tag{1.66}$$

$$k_\mathrm{F0(p)} = \left(V_\mathrm{gs} - V_\mathrm{th0}\right)^{\frac{3}{2}}\Gamma_\mathrm{IT}\mathrm{e}^{-\frac{E_\mathrm{AkF}}{k_\mathrm{B}T}} \qquad k_\mathrm{R(p)} = k_\mathrm{R0(p)}\mathrm{e}^{-\frac{E_\mathrm{AkR}}{k_\mathrm{B}T}} \tag{1.67}$$

Here the Si-H bond breaking reaction rate is given by constants $k_\mathrm{F(s)}$, $k_\mathrm{F(p)}$ as well as the bond annealing rate by $k_\mathrm{R(s)}$ and $k_\mathrm{R(p)}$. The initial Si-H bond densities for the Si/SiO$_2$ and SiO$_2$/poly − Si interface are $N_\mathrm{0(s)}$ and $N_\mathrm{0(p)}$, the interface trap densities are $N_\mathrm{IT(s)}$ and $N_\mathrm{IT(p)}$, $N_\mathrm{H}^{(s)}$ and $N_\mathrm{H}^{(p)}$ represent the atomic hydrogen density near the interface as well as $N_\mathrm{H2}^{(s)}$ and $N_\mathrm{H2}^{(p)}$ the molecular hydrogen density. The concentrations of atomic and molecular hydrogen are N_H and N_H2, the corresponding diffusivities are D_H and D_H2, k_H and k_H2 are the generation and dissociation rates of H$_2$ and the interface thickness ($\sim 1.5\,\mathrm{Å}$) is given by δ.

Besides the device dependent parameters $k_\mathrm{F0(s)}$ ($k_\mathrm{F0(p)} \approx 3/5 \times k_\mathrm{F0(s)}$) and Γ_{IT} typical parameter values are:

$$E_\mathrm{AkF} = 0.175\,\mathrm{eV} \qquad E_\mathrm{Akr} = 0.2\,\mathrm{eV} \qquad k_\mathrm{H0} = 8.56\,\mathrm{cm^2\,s^{-1}} \qquad E_\mathrm{AkH} = 0.3\,\mathrm{eV}$$

$$k_\mathrm{H20} = 507 \times 10^5\,\mathrm{s^{-1}} \qquad E_\mathrm{AkH2} = 0.3\,\mathrm{eV} \qquad E_\mathrm{ADH} = 0.2\,\mathrm{eV} \qquad E_\mathrm{ADH2} = 0.58\,\mathrm{eV}$$

For the conventional H/H$_2$ RD model the following parameter values are used,

$$k_\mathrm{R0(s)} = 9.9 \times 10^{-7} \qquad D_\mathrm{H0} = 9.56 \times 10^{-8}\,\mathrm{cm^2\,s^{-1}} \qquad D_\mathrm{H20} = 3.5 \times 10^{-5}\,\mathrm{cm^2\,s^{-1}},$$

and for the poly H/H$_2$ RD model it is

$$k_\mathrm{R0(s)} = 9.9 \times 10^{-5} \qquad k_\mathrm{R0(p)} = 8 \times 10^{-4}$$

$$D_\mathrm{H0} = 1.5 \times 10^{-5}\,\mathrm{cm^2\,s^{-1}} \qquad D_\mathrm{H20} = 9.5 \times 10^{-5}\,\mathrm{cm^2\,s^{-1}}.$$

The long-time analytic solution is derived as [105]

$$\Delta N_\mathrm{IT} = \left(\frac{k_\mathrm{F} N_0}{k_\mathrm{R}}\right)^{2/3}\left(\frac{k_\mathrm{H}}{k_\mathrm{H2}}\right)^{1/3}\left(6 \times D_\mathrm{H2} t_\mathrm{stress}\right)^{1/6}. \tag{1.68}$$

As pointed out in the introduction of this section the RD model neither adapts the degradation nor the recovery data of the device under test with physical reasonable constants. Assuming a recovery time of $10^6\,\mathrm{s}$ the conventional $\mathrm{H/H_2}$ RD model yields the following diffusion lengths,

$$\begin{aligned} d_{\mathrm{H0}} &= \sqrt{D_{\mathrm{H0}} \times 10^6\,\mathrm{s}} = \sqrt{9.56 \times 10^{-8}\,\mathrm{cm^2\,s^{-1}} \times 10^6\,\mathrm{s}} \approx 0.3\,\mathrm{cm} \\ d_{\mathrm{H20}} &= \sqrt{D_{\mathrm{H20}} \times 10^6\,\mathrm{s}} = \sqrt{3.5 \times 10^{-5}\,\mathrm{cm^2\,s^{-1}} \times 10^6\,\mathrm{s}} \approx 5.9\,\mathrm{cm}, \end{aligned}$$

which are far beyond the transistor geometries. A detailed discussion of this problem is presented in the next section. This in mind, the RD model should not or only with due care be used to interpret any measurement data.

1.2.2. The microscopic limit of the Reaction Diffusion Model

Since 1977 when Jeppson and Svensson presented the first model for NBTI [53], reaction-diffusion (RD) based degradation models became very popular because they were able to reproduce the experimental drift results. On the contrary, a microscopic formulation of the widely used reaction-diffusion model shows huge differences to the macroscopic model and the obtained experimental results [106], questioning the physical assumptions and the appropriability of the reaction-diffusion framework for systems at the investigated particle densities.

The reaction-diffusion model from Jeppson and Svensson is based on the assumption that for some silicon atoms there are no oxygen neighbors due to the $\mathrm{Si} - \mathrm{SiO_2}$ lattice mismatch. Therefore, the silicon atom has a dangling bond (a single unpaired valence electron), which leads to an increase of the states within the band-gap, and can be observed via electrical measurements [107]. These band-gap states are removed by hydrogen passivation during the wafer fabrication. Once electrical stress is applied to the gate of the transistor, holes are located at the interface and, additionally, due to the increased temperature the hydrogen atoms are released. By that, the residual silicon dangling bonds become traps for charge carriers. The model predicts a very fast establishing equilibrium between de- and repassivation [108, 109] and explains the temporal behavior of the drift via the constant hydrogen flux away from the interface. To explain the macroscopic picture, a rate equation and a Fickian diffusion equation are used to mathematically describe the interface reaction and the hydrogen motion inside the oxide. The modeled device degradation first is linear with time for short stress times and later merges into a power-law,

$$N_{\mathrm{it}}(t) = \sqrt{\frac{k_{\mathrm{f}} N_0}{2 k_{\mathrm{r}}}} (Dt)^{1/4} \tag{1.69}$$

with the depassivation (forward) rate k_{f}, the repassivation (reverse) rate k_{r}, the diffusion

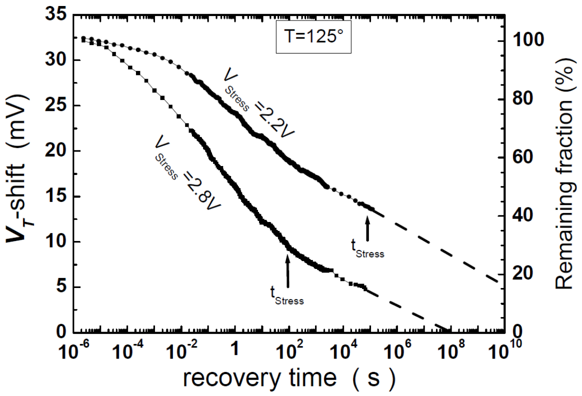


Figure 1.24.: Recovery traces for $V_{\mathrm{gs,stress}} = 2.8\,\mathrm{V}$, $t_{\mathrm{stress}} = 10^2\,\mathrm{s}$ and $V_{\mathrm{gs,stress}} = 2.2\,\mathrm{V}$, $t_{\mathrm{stress}} = 10^5\,\mathrm{s}$. It is obvious that the recovery is slower for the long stress time because of the wider diffusion distance. In contrast to the data, the RD theory predicts a recovery of 50% after $t_{\mathrm{recovery}} = t_{\mathrm{stress}}$. [112]

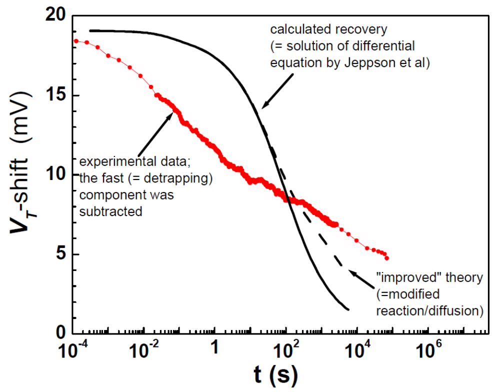


Figure 1.25.: The comparison between the experimental data (red) and the calculated recovery (black) for $V_{\mathrm{gs,stress}} = 2.8\,\mathrm{V}$, $t_{\mathrm{stress}} = 10^2\,\mathrm{s}$ shows that neither the conventional (solid) nor the "improved" RD theory (dashed) can explain the measured data. [112]

coefficient D and the initial density of interface traps N_0. However, drift results after the 1970's show different power-law exponents than $1/4$. To stay abreast with that findings, different diffusing species (e.g. H_2) were introduced for a modified RD model [110]. About four decades later NBTI came back into the focus of the scientific community, when the experimentally verified ultra-fast recovery of the degradation ($< 1\,\mu\mathrm{s}$, Fig. 1.24) [111, 112] strongly contradicted the assertions of the RD model. A comparison between experimental data and the RD model is presented in Fig. 1.25, notably showing the wrong prediction of NBTI recovery of RD models.

To verify the claims [113] of the RD framework in detail and to generally investigate the rate-equation-based MOSFET reliability description, Schanovsky et al. "derived and implemented a microscopic formulation of the RD model [114, 115], in order to study the behavior of the RD mechanism on the atomic scale." [106]. Despite the fact that the RD framework is set-up physically reasonable, it is not considered correctly that for low concentrations the partial differential equation-based description of chemical kinetics breaks down [116]. Furthermore, bimolecular reactions in RD systems require a certain reaction radius [117], which is the rate-limiting part. The RD model is normally based on a dangling bond density about $N_0 = 5 \times 10^{12}\,\mathrm{cm}^{-2}$, leading to a nearest neighbor distance of $d = \sqrt{1/N_0} \approx 4.5\,\mathrm{nm}$. Because within that distance several other atoms are present, one cannot assume any elementary reaction for such "neighbors". Hence, the bimolecular reaction has to be influenced by diffusive steps, which leads to inconsistencies between the RD framework and the physical interpretation of it. To obtain results comparable to experimental data, a competition between the diffusing hydrogen atoms for available “free” dangling bonds is needed and they have to dimerize at a specific rate. To fulfill these requirements, a diffusion coefficient of $D = 10^{-13}\,\mathrm{cm^2\,s^{-1}}$ is commonly used [118, 119], which leads depending on the published dimerization rate to a reaction radius ranging from 70 µm [118] to several kilometers [119], using the estimation from the Smoluchowski theory for irreversible bimolecular reaction [117, 120, 121],

$$\rho_\mathrm{H} = \frac{k_\mathrm{H}}{4\pi D}. \tag{1.70}$$

In contrast to that value, a reasonable reaction radius is about 4 Å [122], which indicates the limited physical validity of the parametrization of the RD framework. Therefore, an in-depth evaluation of the RD model was performed to validate its claims [113], using a grid-based atomistic RD simulator. For the calculations a cuboid is employed, which has a fixed width and length (gate dimensions) and an infinitesimal height orthogonal to the $\mathrm{Si} - \mathrm{SiO_2}$ interface plane. The uniform grid size is set to $h \approx 4\,\mathrm{Å}$ [122]. For the number of dangling bonds n_DB, the number of hydrogen atoms passivating a dangling bond n_p and the numbers n_Hi and $n_\mathrm{H_2i}$ for free atomic and molecular hydrogen inside the volume $V_i = h^3 \approx 64\,\mathrm{Å}^3$ at the interstitial i it follows

$$N_0 = \frac{n_\mathrm{DB}}{WL} \qquad N_\mathrm{it} = \frac{n_\mathrm{DB} - n_\mathrm{p}}{WL} \tag{1.71}$$

$$H(x_i) = \frac{n_{\mathrm{H}i}}{V_i} \qquad H_2(x_i) = \frac{n_{\mathrm{H_2}i}}{V_i}. \tag{1.72}$$

The results for a model system show a remarkable difference in the degradation between the microscopic and the macroscopic RD model (see Fig. 1.26). In addition, a comparison to experimental data (Fig. 1.27) clearly shows that only unreasonable interstitial sizes would allow the application of the 1D RD model to obtain the experimentally observed $t^{1/6}$ aging behavior, using the parameters in Table 1.1. Also, a justification of the diffusion coefficients was investigated, but shows only for the unphysical limit $D_\mathrm{I} \to \infty$ a match with the macroscopic model. Hence, “the apparent match of the RD models with the experimental data must [...] be considered a mathematical artifact without any physical background”, Schanovsky and Grasser in [106].

parameter	value
k_f	$3\,\mathrm{s^{-1}}$
k_r	$6 \times 10^{-13}\,\mathrm{cm^3\,s^{-1}}$
k_H	$5.6 \times 10^{-11}\,\mathrm{cm^3\,s^{-1}}$
$k_\mathrm{H_2}$	$95.4\,\mathrm{s^{-1}}$
D	$10^{-13}\,\mathrm{cm^2\,s^{-1}}$
D_2	$1.8126 \times 10^{-14}\,\mathrm{cm^2\,s^{-1}}$
N_0	$5 \times 10^{12}\,\mathrm{cm^{-2}}$

Table 1.1.: For the simulations a parameter set based on the values published in [118] was used. In order to allow physically more appropriate k_r and k_H, the other parameters were slightly changed to result in the same degradation behavior. [106]

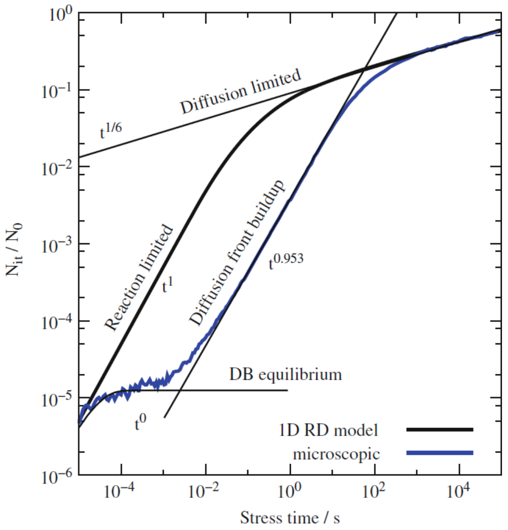


Figure 1.26.: A large discrepancy between the microscopic and the macroscopic RD model is observable. Furthermore, the physically more reasonable microscopic model cannot be confirmed by experimental data. [106]

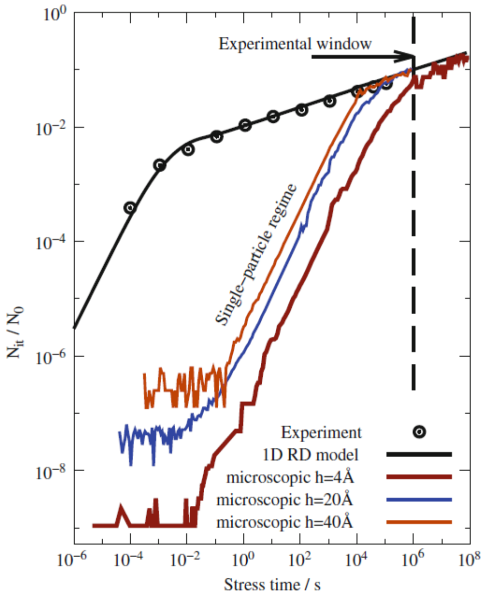


Figure 1.27.: The comparison between experimental data and the degradation transient, which is predicted by the microscopic and macroscopic RD models, demonstrates that even unphysical large interstitial sizes (40 Å) fail to reflect the $t^{1/6}$ behavior for stress times shorter than 10^4 s. [106]

1.2.3. Trap charging model

Until the last decade transistor degradation caused by bias temperature instability has been explained using the reaction-diffusion (RD) theory [53, 104, 105, 110]. Nowadays, it has been shown in the literature that the RD framework is based on unphysical assumptions [114] and cannot reflect the experimental results for NBTI recovery [112]. Recent BTI modeling employs the non-radiative multiphonon process (NMP) as the trapping process, which is found to be the root cause of many related physical issues. Despite other NMP models, the here presented [123] model by Grasser et al. is based on the microscopic Franck-Condon theory and reveals generalized trapping rates. By that, all possible NMP transitions for the substrate and poly-gate are covered, including the quadratic electron-phonon coupling. To explain the experimental results of time-dependent dielectric spectroscopy (TDDS), which allows the detection of charge emissions of individual defects using recovery traces of p-MOSFETs after NBTI stress [124], metastable states are introduced to account for the trapping dynamics of the analyzed defect. The observed close model-hardware relationship facilitates a detailed analysis of the physical trapping mechanism of BTI for fixed and switching oxide hole traps in contrast to empirical explanations.

The first charge trapping models [125–128] were based on holes, able to tunnel elastically between the substrate and oxide defects. Nevertheless, the experimentally observed thermal activation could not be reproduced by them. To overcome this problem, the Schockley-Read-Hall (SRH) model [129] was adapted to take into account both the tunneling effect [130] as well as the temperature dependance of BTI [131, 132]. In this ansatz the thermal activation was considered to be caused by NMP transitions. Unfortunately, the resulting transition barriers were not described on a microscopic level but phenomenologically [133–135]. In contrast to phenomenologically models, the here presented multi-state model [123] is based on a rigorous NMP framework, which describes the charge transfer process, and also takes metastable states into consideration where bistable defects are modeled via two-step capture and emission processes. The Huang-Born approximation describes an atomic configuration by a system of electrons and nuclei whose dynamic is given by two separated Schrödinger equations:

$$\left\{\hat{T}_\mathrm{e} + \hat{V}_\mathrm{ee}\left(\boldsymbol{r}\right) + \hat{V}_\mathrm{en}\left(\boldsymbol{r};\boldsymbol{R}\right) + \hat{V}_\mathrm{nn}\left(\boldsymbol{R}\right)\right\}\varphi_i\left(\boldsymbol{r};\boldsymbol{R}\right) = V_i\left(\boldsymbol{R}\right)\varphi_i\left(\boldsymbol{r};\boldsymbol{R}\right) \tag{1.73}$$

$$\left\{\hat{T}_\mathrm{n} + V_i\left(\boldsymbol{R}\right)\right\}\eta_{i\alpha}\left(\boldsymbol{R}\right) = E_{i\alpha}\eta_{i\alpha}\left(\boldsymbol{R}\right) \tag{1.74}$$

All Coulomb contributions are included in Eq. 1.73, $\hat{V}_\mathrm{ee}$ for electron-electron, $\hat{V}_\mathrm{en}$ for electron-nucleus and $\hat{V}_\mathrm{nn}$ for nucleus-nucleus interactions. The degrees of freedom of the electrons and nuclei are given by $\boldsymbol{r}$ and $\boldsymbol{R}$, and the corresponding kinetic energies are represented by $\hat{T}_\mathrm{e}$ and $\hat{T}_\mathrm{n}$. Because it corresponds to a configuration energy and appears as a potential in the Schrödinger Eq. 1.74, the solution $V_i\left(\boldsymbol{R}\right)$ of 1.73 is also called the adiabatic potential energy. BTI related charge trapping is referred to as "charge transfer reaction" in the literature and has to be described by a comprehensive atomic system, including the defect-related atoms as well as the substrate atoms. This leads to a $3N$-dimensional space with N considered atoms. The surfaces of the adiabatic potential energy have a nearly parabolic shape, which can be approximated by harmonic quantum oscillators. Equal curvatures imply a single intersection point (IP) of both potential surfaces (linear electron-phonon coupling), whereas for different curvatures two intersection points exist (quadratic electron-phonon coupling).

Using the Franck-Condon approximation [136–139], the NMP transition rate is derived from first-order time-dependent perturbation theory to

$$k_{ij} = A_{ij} f_{if} \tag{1.75}$$

$$A_{ij} = \frac{2\pi}{\hbar}\left|\left\langle\varphi_i\left|V'\right|\varphi_j\right\rangle\right|^2 \tag{1.76}$$

$$f_{ij} = \operatorname*{ave}_{\alpha}\sum_{\beta}\left|\left\langle\eta_{i\alpha}\left|\eta_{i\beta}\right.\right\rangle\right|^2 \tag{1.77}$$

with the thermal average "ave" over the initial states α and the sum over the final states

β. For the matrix element A_{ij} a simple electronic transition is assumed. The Franck-Condon factor $|\langle\eta_{i\alpha}|\eta_{i\beta}\rangle|^2$ only contributes for the initial and final state having the same energy. In this case, the resulting overlap integral of the vibrational wavefunctions $i\alpha$ and $j\beta$ are in accordance with the appropriate transition probability. The lineshape function f_{ij} governs the temperature dependence and gate bias of the NMP transition rate. For hole capture it is defined by [140]

$$\begin{aligned} f_{0/+}\left(c_0, c_+, q_s, V_0, V_+\right) &= f_{0/+}\left(c_0, c_+, q_s, V_+\right) \\ &= \frac{1}{2}\sqrt{\frac{c_0\beta}{\pi}}\left(\frac{\mathrm{e}^{-\beta c_0 \Delta q_1^2}}{\left|c_0\Delta q_1 - c_+\left(\Delta q_1 - q_s\right)\right|} + \frac{\mathrm{e}^{-\beta c_0 \Delta q_2^2}}{\left|c_0\Delta q_2 - c_+\left(\Delta q_2 - q_s\right)\right|}\right). \end{aligned} \tag{1.78}$$

For NBTI in p-MOSFETs normally the "hole picture" [123] is used to describe the system. Here, the hole of unknown initial energy is emitted from a continuum of valence band states into the trap state with energy E_t. This leads to a new formulation of the energy minima,

$$V_0 = \tilde{V}_0 - E \tag{1.79}$$

$$V_+ = \tilde{V}_0 - E_\mathrm{t}, \tag{1.80}$$

and the energy difference of the adiabatic potentials reads as

$$V_\mathrm{s} = V_+ - V_0 = E - E_\mathrm{t}. \tag{1.81}$$

To take account of the interaction between the oxide defects and the conduction and valence band of the substrate, the NMP formulation has to be extended to a variety of band states considering all possible energies E. Due to the continuous spectrum of the band states a summation over their number n can be transformed into an integral over the DOS [141], leading to the NMP hole capture rate

$$k_{0/+}^{\mathrm{pc}} = \Omega \int_{-\infty}^{E_\mathrm{v}} D_\mathrm{p}(E)\, f_\mathrm{p}(E, E_\mathrm{f})\, A_{0/+}(E, x_\mathrm{t})\, f_{0/+}(c_0, c_+, q_\mathrm{s}, E - E_\mathrm{t})\, f_t \mathrm{d}E \tag{1.82}$$

Using the WKB approximation [123] the full set of NMP trapping rates is given by

$$k^{\mathrm{nc}} = k_0^{\mathrm{n}} \int_{E_\mathrm{c}}^{+\infty} D_\mathrm{n}(E)\, f_\mathrm{n}(E, E_\mathrm{f})\, \lambda(E, x_\mathrm{t})\, f_{+/0}(c_+, c_0, q_\mathrm{s}, E_\mathrm{t} - E)\, \mathrm{d}E \tag{1.83}$$

$$k^{\mathrm{ne}} = k_0^{\mathrm{n}} \int_{E_\mathrm{c}}^{+\infty} D_\mathrm{n}(E)\, f_\mathrm{p}(E, E_\mathrm{f})\, \lambda(E, x_\mathrm{t})\, f_{0/+}(c_0, c_+, q_\mathrm{s}, E - E_\mathrm{t})\, \mathrm{d}E \tag{1.84}$$

$$k^{\mathrm{pc}} = k_0^{\mathrm{p}} \int_{-\infty}^{E_\mathrm{v}} D_\mathrm{p}(E)\, f_\mathrm{p}(E, E_\mathrm{f})\, \lambda(E, x_\mathrm{t})\, f_{0/+}(c_0, c_+, q_\mathrm{s}, E - E_\mathrm{t})\, \mathrm{d}E \tag{1.85}$$

$$k^{\mathrm{pe}} = k_0^{\mathrm{p}} \int_{-\infty}^{E_\mathrm{v}} D_\mathrm{p}(E)\, f_\mathrm{n}(E, E_\mathrm{f})\, \lambda(E, x_\mathrm{t})\, f_{+/0}(c_+, c_0, q_\mathrm{s}, E_\mathrm{t} - E)\, \mathrm{d}E \tag{1.86}$$

"n" and "p" represent electrons and holes, "c" and "e" capture and emission. The pref-

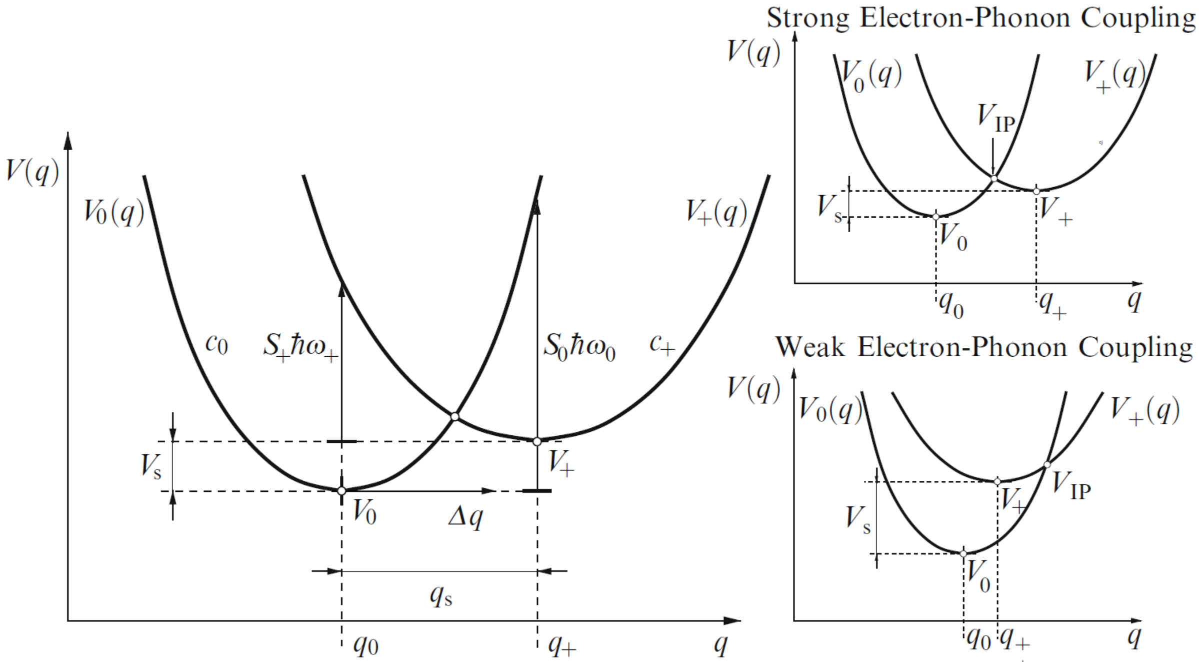


Figure 1.28.: The usage of the Huang-Rhys factors S_0 and S_+ allows a compact analytical solution of the hole capture and emission rate. On the **(left)** side the configuration coordinate diagram is shown. If the adiabatic potentials are defined as harmonic oscillators ($V_0\left(q\right) = {}^1\!/\!_2 M\omega_0^2\left(q-q_0\right)^2 + V_0$, $V_+\left(q\right) = {}^1\!/\!_2 M\omega_+^2\left(q-q_+\right)^2 + V_+$), an optical transition with energy difference $V_0\left(q_+\right) - V_+$ is equal to $S_0 \times \hbar\omega_0$. This is the same case for the energy difference $V_+\left(q_0\right) - V_0$, which is equal to $S_+ \times \hbar\omega_+$. **Right:** For the case of strong (top) electron-phonon coupling, the intersection point is located in between the minima of both parabolas, and for weak (bottom) electron-phonon coupling it is situated besides them. [123]

actors Ω and A_1 are substituted by $k_0^{\mathrm{n/p}}$. Because the integrands generally are sharply peaked, the numerical solution of these integrals is obtained via adaptive integration schemes to allow a good accuracy with respect to computation time.

The Huang-Rhys factor S [136] (the electron-phonon coupling constant) enables a compact analytical solution of the situation for fluorescence and absorption spectra of gases and solids. Here, the quadratic electron-phonon coupling can be depicted by two parabolas for the adiabatic potentials, having different curvatures and energy minima leading to the NMP transition barrier (Fig. 1.28)

$$V_{0/+}\left(V_{\mathrm{s}}\right) = \frac{S\hbar\omega}{\left(R^2-1\right)^2}\left(1 \pm R\sqrt{\frac{S\hbar\omega + V_{\mathrm{s}}\left(R^2-1\right)}{S\hbar\omega}}\right)^2 \tag{1.87}$$

with $S\hbar\omega = c_0 q_{\mathrm{s}}^2$ and $R^2 = c_0/c_+$.

Following [123], for the case of linear electron-phonon coupling where both parabolas have the same curvature, this simplifies to

$$V_{0/+}\left(V_{\mathrm{s}}\right) = \frac{\left(V_{\mathrm{s}} + S\hbar\omega\right)^2}{4S\hbar\omega} \tag{1.88}$$

with the prefactor $\xi_{0/+}\left(\Delta q\right)$ of the exponential term in the lineshape function (1.78)

$$\xi_{0/+}\left(\Delta q\right)=\sqrt{\frac{\beta}{4\pi}}\frac{1}{\sqrt{S\hbar\omega}}. \tag{1.89}$$

The second order expansion of the NMP transition barrier 1.87 yields

$$V_{0/+}\left(V_{\mathrm{s}}\right)\approx\frac{S\hbar\omega}{\left(1+R\right)^{2}}+\frac{R}{1+R}V_{\mathrm{s}}+\frac{R}{4S\hbar\omega}V_{\mathrm{s}}^{2} \tag{1.90}$$

and with

$$V_{\mathrm{s}}=V_{+}-V_{0}=E-E_{\mathrm{t}}=\underbrace{E-E_{\mathrm{v}}}_{=-\Delta E}+E_{\mathrm{v}}-E_{\mathrm{t}} \tag{1.91}$$

eq. 1.90 reads as

$$V_{0/+}\left(\Delta E\right)\approx\frac{S\hbar\omega}{\left(1+R\right)^{2}}+\frac{R}{1+R}\left(E_{\mathrm{v}}-E_{\mathrm{t}}-\Delta E\right)+\frac{R}{4S\hbar\omega}\left(E_{\mathrm{v}}-E_{\mathrm{t}}-\Delta E\right)^{2}. \tag{1.92}$$

Under the following assumptions [123],

- “Assuming the parabolic-band approximation, the valence band density of states [...] is given by $D_{\mathrm{p}}\left(E\right)=D_{\mathrm{p},0}\sqrt{\Delta E}$ with $D_{\mathrm{p},0}$ being an energy-independent prefactor.
- The occupancy $f_{\mathrm{p}}\left(E,E_{\mathrm{f}}\right)$ follows Boltzmann statistics.
- The WKB factor is approximated by the factor $exp(-x_{\mathrm{t}}/x_{0})$ with the tunneling length x_{0}.
- The lineshape function is dominated by the exponential barrier term so that the prefactor $\xi_{0/+}$ can be neglected to first order.”

the **hole capture rate** (Eq. 1.85) can be simplified for strong electron-phonon coupling ($S\hbar\omega\gg\left|E_{\mathrm{v}}-E_{\mathrm{t}}-\Delta E\right|$) to

$$\begin{aligned}k^{\mathrm{pc}}&=k_{0}^{\mathrm{p}}\int_{-\infty}^{E_{\mathrm{v}}}D_{\mathrm{p}}\left(E\right)f_{\mathrm{p}}\left(E,E_{\mathrm{f}}\right)\lambda\left(E,x_{\mathrm{t}}\right)\mathrm{e}^{-\beta V_{0/+}\left(\Delta E\right)}\mathrm{d}E\\&=k_{0}^{\mathrm{p}}\left(1+R\right)^{3/2}p\mathrm{e}^{-\frac{x_{\mathrm{t}}}{x_{0}}}\mathrm{e}^{-\beta\left(\frac{S\hbar\omega}{(1+R)^{2}}-\frac{R}{1+R}\Delta E_{\mathrm{t}}\right)},\end{aligned} \tag{1.93}$$

with hole density $p=D_{\mathrm{p},0}\mathrm{e}^{\beta\left(E_{\mathrm{v}}-E_{\mathrm{t}}\right)}\beta^{-3/2}\Gamma\left({}^{3}\!/\!_{2}\right).$

Compared to the rate equations from the standard SRH theory, k_{0}^{p} can be identified with the hole thermal velocity $\nu_{th,p}$ multiplied with a hole capture cross-section σ_{p}. Using the hole capture barrier,

$$\epsilon^{\mathrm{pc}}=V_{0/+}\Big|_{\Delta E=0}=\frac{S\hbar\omega}{\left(1+R\right)^{2}}+\frac{R}{1+R}E_{\mathrm{v}}-E_{\mathrm{t}}, \tag{1.94}$$

the hole capture rate can be simplified to

$$k^{\mathrm{pc}}=\nu_{\mathrm{th,p}}\sigma_{\mathrm{p}}\left(1+R\right)^{3/2}p\mathrm{e}^{-\frac{x_{\mathrm{t}}}{x_{0}}}\mathrm{e}^{-\beta\epsilon^{\mathrm{pc}}}. \tag{1.95}$$

Because the electron occupation function is replaceable by

$$f_{\mathrm{n}}\left(E,E_{\mathrm{f}}\right)=f_{\mathrm{p}}\left(E,E_{\mathrm{f}}\right)\mathrm{e}^{-\beta\left(E-E_{\mathrm{f}}\right)} \tag{1.96}$$

and in the hole picture the ratio of the exponential barrier terms for each band state reads as

$$\frac{\mathrm{e}^{-\beta V_{+/0}}}{\mathrm{e}^{-\beta V_{0/+}}} = \mathrm{e}^{-\beta(E_\mathrm{t}-E)}, \tag{1.97}$$

the **hole emission rate** can be expressed by

$$\begin{aligned} k^\mathrm{pe} &= v_\mathrm{th,p}\sigma_\mathrm{p}\int_\infty^{E_\mathrm{v}} D_\mathrm{p}(E)\, f_\mathrm{n}(E,E_\mathrm{f})\,\lambda(E,x_\mathrm{t})\,\mathrm{e}^{-\beta V_{0/+}}\mathrm{d}E \\ &= v_\mathrm{th,p}\sigma_\mathrm{p}(1+R)^{3/2}\,\mathrm{e}^{\frac{x_\mathrm{t}}{x_0}} p\mathrm{e}^{-\beta\epsilon^{pc}}\mathrm{e}^{-\beta(E_\mathrm{t}-E_\mathrm{f})}. \end{aligned} \tag{1.98}$$

Although several approximations were used, (1.95) and (1.98) reflect the main physics for charge trapping and allow gate bias and temperature dependent modeling of strong electron-phonon coupling [123].

Time-Dependent Defect Spectroscopy (TDDS, see Section 1.2.6) experiments show that defects can disappear during stress and recovery phases from the spectral maps. In addition to that observation, transitions were found which are gate-bias independent and cannot be understood with charge transfer reactions. These can only be explained with an activation over thermal barriers. To account for those two observations, metastable states (1' and 2') were introduced in the NMP model, in which certain oxide defects can abide for some time. The latest version of the NMP transition rates based multi-state model is presented in Fig. 1.29. Here, a NMP transition occurs between states with an opposite charge ($1 \leftrightarrow 2'$ and $1' \leftrightarrow 2$) and a thermal transition between states with the same charge ($1 \leftrightarrow 1'$ and $2 \leftrightarrow 2'$).

Due to the complex dynamics of the defect state transitions, homogeneous continuous-time Markov chain theory must be employed. Here, a future transition is independent from the past of the analyzed system. This assumption is only valid for a defect, which relaxed completely by interaction with its environment. This is the case for pure thermal as well as NMP transitions.

The time evolution of the system is described by the following Master equation:

$$\delta_t\pi_\mathrm{i}(t) = \sum_{\mathrm{j}\neq\mathrm{i}}\pi_\mathrm{j}(t)\,k_\mathrm{ji} - \sum_{\mathrm{i}\neq\mathrm{j}}\pi_\mathrm{i}(t)\,k_\mathrm{ij} \tag{1.99}$$

with the time-dependent occupation probability $\pi_\mathrm{i}(t)$ of a defect being in state i and the transition rate k_ij between state i and j. To take into account that the metastable states are not observable during the experiment, a two-step process needs to be considered which transition times read as

$$\tau_{\alpha\gamma} = \frac{k_{\alpha\beta}+k_{\beta\gamma}+k_{\beta\alpha}}{k_{\alpha\beta}k_{\beta\gamma}}. \tag{1.100}$$

The four-state NMP model incorporates four possible transition pathways with the first-passage times (1.101-1.104) with the gate bias independent thermal transitionbarri-

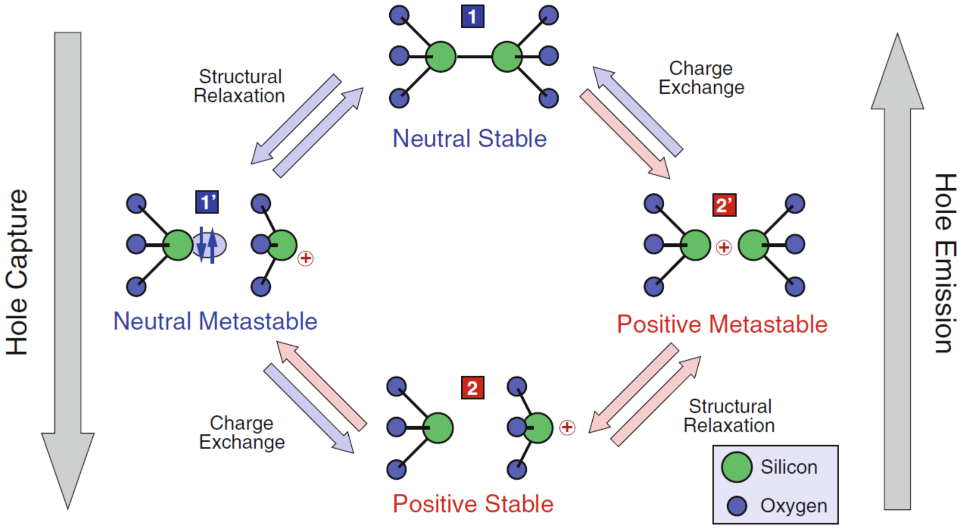


Figure 1.29.: The state diagram of the multi-state model consists of two stable (1, 2) and two metastable (1’, 2’) states. The NMP transitions ($1 \leftrightarrow 2'$ and $2 \leftrightarrow 1'$) are valid for states with different charges, the thermal transitions ($1 \leftrightarrow 1'$ and $2 \leftrightarrow 2'$) occur for states with the same charge. The experimental data correspond to the capture and emission times of the transition between the two stable states. The metastable states are needed to explain the gate bias and temperature dependence of the capture and emission times. [123]

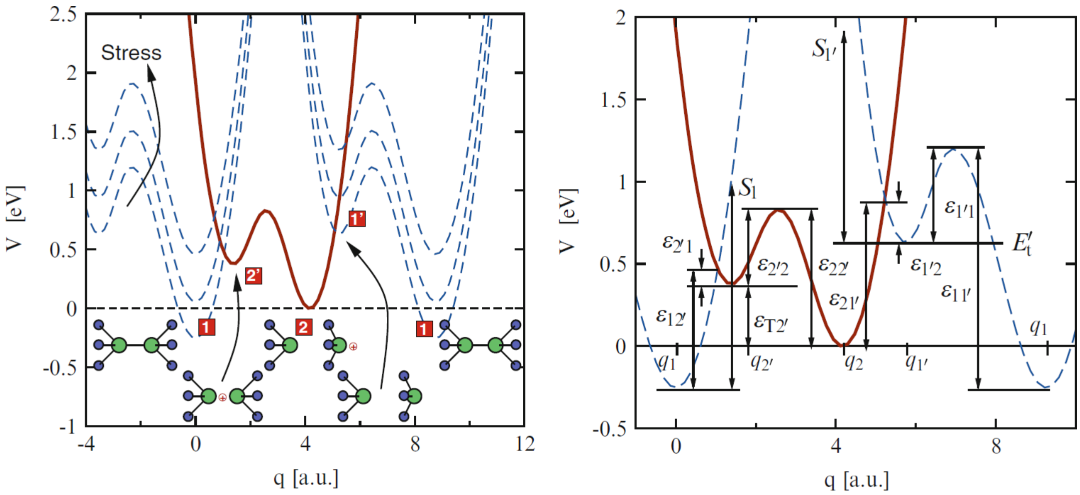


Figure 1.30.: Left: In this schematic configuration coordinate diagram of a bistable defect the solid/dashed line represents the adiabatic potential of its positive/neutral charge state. To illustrate the various stable and metastable defect configurations, stick-and-ball models are inserted. Right: The used energies and barriers for the multi-state model. [123]

ers ϵ_{ij} (see Fig. 1.30) (1.105-1.108) having an attempt frequency ν_0 of the order $10^{13}\,\mathrm{Hz}$:

$$\tau_{\mathrm{c}}^{2'} = \frac{1}{k_{12'}} + \frac{1}{k_{2'2}} + \frac{1}{k_{2'2}}\frac{k_{2'1}}{k_{12'}} \tag{1.101}$$

$$\tau_{\mathrm{c}}^{1'} = \frac{1}{k_{11'}} + \frac{1}{k_{1'2}} + \frac{1}{k_{1'2}}\frac{k_{1'1}}{k_{11'}} \tag{1.102}$$

$$\tau_{\mathrm{e}}^{2'} = \frac{1}{k_{22'}} + \frac{1}{k_{2'1}} + \frac{1}{k_{2'1}}\frac{k_{2'2}}{k_{22'}} \tag{1.103}$$

$$\tau_{\mathrm{e}}^{1'} = \frac{1}{k_{2'1}} + \frac{1}{k_{1'1}} + \frac{1}{k_{1'1}}\frac{k_{1'2}}{k_{21'}} \tag{1.104}$$

$$k_{11'} = \nu_0 \mathrm{e}^{-\beta\epsilon_{11'}} \tag{1.105}$$

$$k_{1'1} = \nu_0 \mathrm{e}^{-\beta\epsilon_{1'1}} \tag{1.106}$$

$$k_{22'} = \nu_0 \mathrm{e}^{-\beta\epsilon_{22'}} \tag{1.107}$$

$$k_{2'2} = \nu_0 \mathrm{e}^{-\beta\epsilon_{2'2}} \tag{1.108}$$

In the hole picture the adiabatic potentials are given by

$$V_1 = \tilde{V}_0 - E \qquad V_2 = \tilde{V}_0 - E_{\mathrm{t}}' \tag{1.109}$$

$$V_{1'} = \tilde{V}_0 - E \qquad V_{2'} = \tilde{V}_0 + \epsilon_{\mathrm{T2'}} - E_{\mathrm{t}} \tag{1.110}$$

Because the minimum energy of state $2'$ is $\tilde{V}_0 + \epsilon_{\mathrm{T2'}}$, it follows that:

$$V_{12'} = V_{2'} - V_1 = E - E_{\mathrm{t}} + \epsilon_{\mathrm{T2'}} \qquad V_{1'2} = V_2 - V_{1'} = E - E_{\mathrm{t}}' \tag{1.111}$$

Hence, the NMP transition rates are described by:

$$\begin{aligned} k_{12'} = {} & \nu_{\mathrm{th,n}}\sigma_{\mathrm{n}} \int_{E_{\mathrm{c}}}^{+\infty} D_{\mathrm{n}}(E)\, f_{\mathrm{p}}(E, E_{\mathrm{f}})\, \lambda(E)\, f_{0/+}\left(c_0, c_+, q_{\mathrm{s}}, \underbrace{E - E_{\mathrm{t}} + \epsilon_{\mathrm{T2'}}}_{=V_{12'}}\right) \mathrm{d}E \\ & + \nu_{\mathrm{th,p}}\sigma_{\mathrm{p}} \int_{-\infty}^{E_{\mathrm{v}}} D_{\mathrm{p}}(E)\, f_{\mathrm{p}}(E, E_{\mathrm{f}})\, \lambda(E)\, f_{0/+}(c_0, c_+, q_{\mathrm{s}}, E - E_{\mathrm{t}} + \epsilon_{\mathrm{T2'}})\, \mathrm{d}E \end{aligned} \tag{1.112}$$

$$\begin{aligned} k_{2'1} = {} & \nu_{\mathrm{th,n}}\sigma_{\mathrm{n}} \int_{E_{\mathrm{c}}}^{+\infty} D_{\mathrm{n}}(E)\, f_{\mathrm{n}}(E, E_{\mathrm{f}})\, \lambda(E)\, f_{+/0}\left(c_+, c_0, q_{\mathrm{s}}, \underbrace{E_{\mathrm{t}} - \epsilon_{\mathrm{T2'}} - E}_{=-V_{12'}}\right) \mathrm{d}E \\ & + \nu_{\mathrm{th,p}}\sigma_{\mathrm{p}} \int_{-\infty}^{E_{\mathrm{v}}} D_{\mathrm{p}}(E)\, f_{\mathrm{n}}(E, E_{\mathrm{f}})\, \lambda(E)\, f_{+/0}(c_+, c_0, q_{\mathrm{s}}, E_{\mathrm{t}} - \epsilon_{\mathrm{T2'}} - E)\, \mathrm{d}E \end{aligned} \tag{1.113}$$

$$\begin{aligned} k_{12'} = {} & \nu_{\mathrm{th,n}}\sigma_{\mathrm{n}} \int_{E_{\mathrm{c}}}^{+\infty} D_{\mathrm{n}}(E)\, f_{\mathrm{p}}(E, E_{\mathrm{f}})\, \lambda(E)\, f_{0/+}\left(c_0, c_+, q_{\mathrm{s}}, \underbrace{E - E_{\mathrm{t}}'}_{=V_{1'2}}\right) \mathrm{d}E \\ & + \nu_{\mathrm{th,p}}\sigma_{\mathrm{p}} \int_{-\infty}^{E_{\mathrm{v}}} D_{\mathrm{p}}(E)\, f_{\mathrm{p}}(E, E_{\mathrm{f}})\, \lambda(E)\, f_{0/+}(c_0, c_+, q_{\mathrm{s}}, E - E_{\mathrm{t}}')\, \mathrm{d}E \end{aligned} \tag{1.114}$$

$$\begin{aligned} k_{2'1} = {} & \nu_{\mathrm{th,n}}\sigma_{\mathrm{n}} \int_{E_{\mathrm{c}}}^{+\infty} D_{\mathrm{n}}(E)\, f_{\mathrm{n}}(E, E_{\mathrm{f}})\, \lambda(E)\, f_{+/0}\left(c_{+}, c_{0}, q_{\mathrm{s}}, \underbrace{E_{\mathrm{t}}' - E}_{=-V_{1'2}}\right) \mathrm{d}E \\ & + \nu_{\mathrm{th,p}}\sigma_{\mathrm{p}} \int_{-\infty}^{E_{\mathrm{v}}} D_{\mathrm{p}}(E)\, f_{\mathrm{n}}(E, E_{\mathrm{f}})\, \lambda(E)\, f_{+/0}(c_{+}, c_{0}, q_{\mathrm{s}}, E_{\mathrm{t}}' - E)\, \mathrm{d}E. \end{aligned} \tag{1.115}$$

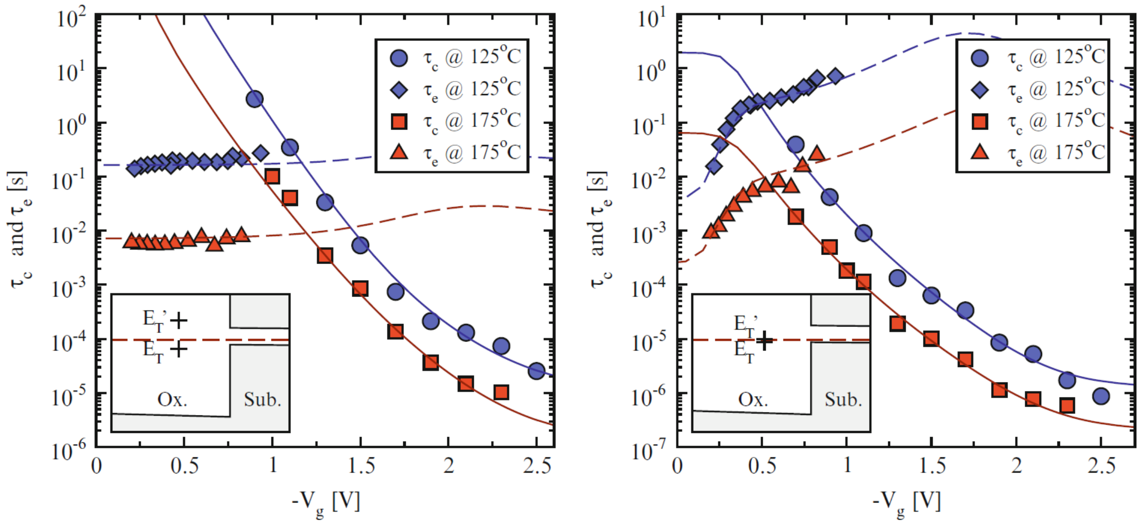


Figure 1.31.: Model evaluation for fixed **(left)** and switching **(right)** oxide hole traps. The capture (solid lines) and emission (dashed lines) times are in good agreement to the experimental data (symbols). For the fixed oxide hole trap E_{t}' is much higher compared to the substrate Fermi level and, therefore, no gate bias dependence is observed for τ_{e}. In contrast to that, for a switching hole trap E_{t}' is very close to the Fermi level, resulting in a distinct dependence of the emission time on V_{gs}. [123]

If the thermal transition rates (Eq. 1.105-1.108) together with the above NMP transition rates are inserted into the equations for the capture and emission times (Eq. 1.101-1.104), one can evaluate them against TDDS data as shown in Fig. 1.31.

1.2.4. Volatility of oxide defects

The defects which are activated by NBTI cannot be described completely by two-state models which only consider the neutral and charged state. Recent research [142] employing TDDS measurements (see Section 1.2.6) show that oxide defects can additionally be volatile. This means that they can disappear and reappear at the same site with same properties. In addition, it is possible, that such defects transform, while changing their properties. Fig. 1.33 shows an example of the disappearance of one defect after many stress and recovery cycles.

The defect did not reappear even after 4×10^{5} s at 398 K. With the help of TDDS measurements traps are identified during the recovery phase by their step height and

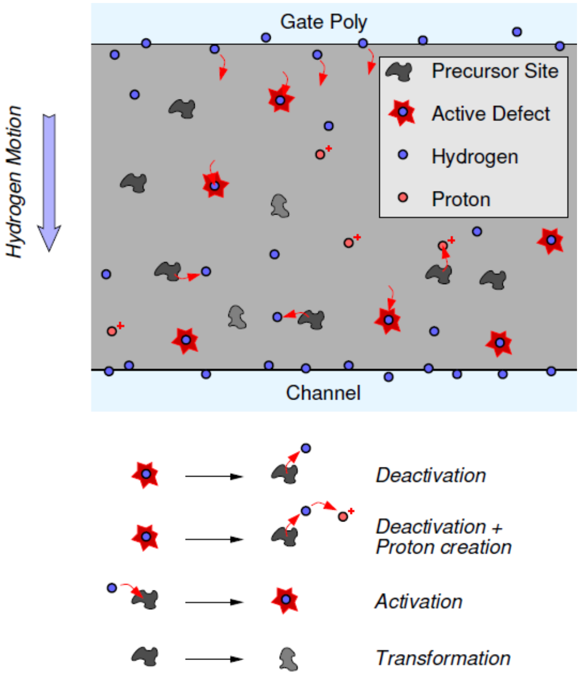


Figure 1.32.: The schematic description of the defect volatility considers the deactivation (+proton creation), activation and transformation of oxide defects. The transport of hydrogen during NBTI stress is supported by nuclear reaction analyses [143, 144]. Note: The channel interface states, which are created by long-term stress, are excluded from this figure. [142]

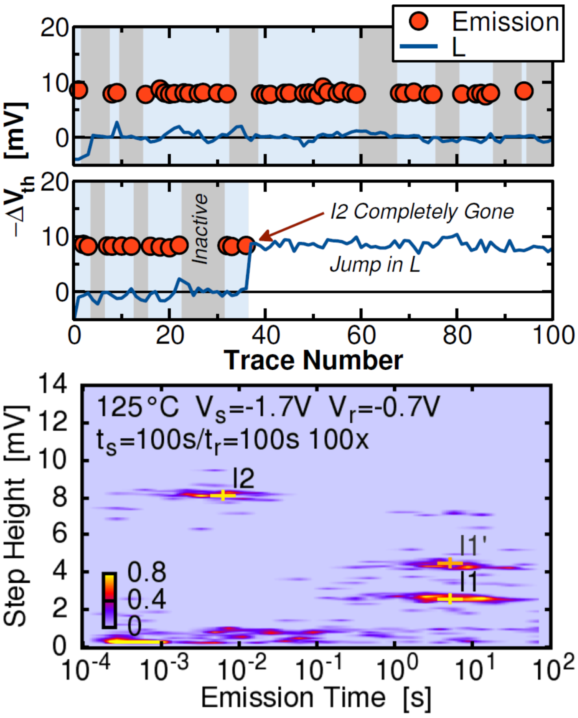


Figure 1.33.: **Top:** The emission pattern of defect I2 during the first 1.5 days shows its volatility. After 100 s stress a recovery trace of 100 s was recorded. As the capture time of I2 is about 4×10^{-3} s, the capture probability is about 1. If the defect emitted a charge during the recovery phase, a red circle is plotted. L (blue line) is the threshold voltage value at the end of the trace. **Middle:** After 1.5 days I2 transformed into a fixed positive charge, thus contributing to L with the same magnitude. **Bottom:** The TDDS spectral map shows the fingerprint of device I with several marked defects. I1 was stable for the whole measurement duration of 4.5 days, while I2 showed a high volatility. [142]

emission time τ_e as presented at the bottom of Fig. 1.33. Another possible transition of a defect can lead to the depassivation of a $Si-H$ bond. To reactivate the defects, baking steps at 623 K can be performed. By that, the capture time of the defect may change by a few orders of magnitude, but not the emission time or step height. Additionally, the reactivation of a defect by a pulse into accumulation, which injects electrons, is possible. The performed DFT calculations with hydroxyl E'-centers are in good accordance with classic models about hydrogen-release [145–148] and support the hypothesis about hydrogen transport from the gate towards the channel [143, 149, 150] during NBTI stress.

The results indicate that atomic hydrogen is able to passivate the dangling Si bond at the hydroxyl E'-center [151] and a further approaching hydrogen atom can reactivate that defect [152]. Due to the large barrier size about 1 to 2 eV the process is considered to be rather reaction- than diffusion-limited. Fig. 1.32 schematically describes the relevant processes behind the volatility of oxide defects. Here, defects can either be present or be activated by hydrogen incorporation at a precursor site (e.g. a strained $Si - O$ bond). It has to be noted that the hydrogen-release should not be mistaken as the reaction-diffusion model, which is presented in Section 1.2.1, because the measured degradation is caused by a collection of first-order processes [153] and the hydrogen release takes place at the gate side, which is consistent to classic oxide breakdown models [154].

1.2.5. The "permanent" component of NBTI

In the past the scientific emphasis about the recovery after NBTI stress was on its recoverable component (R). But, besides that, NBTI consists also of a "permanent" component (P), which investigation is far more difficult due to the involved very long emission times and its superposition by R [155].

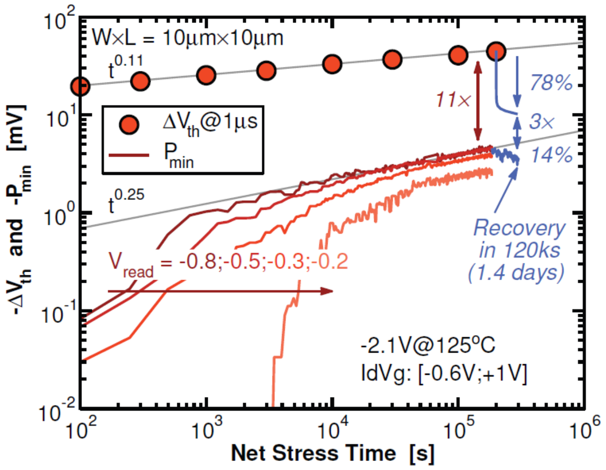


Figure 1.34: The "permanent" component of NBTI is about 10% of the total threshold voltage shift 1 µs after stress. The influence of the chosen readout voltage on P_{min} can be explained by the bias dependent occupancy of the involved interface states. P shows a very small recovery signal compared to R (log scale). [155]

During recovery of NBTI in a p-MOSFET the application of a positive gate bias sweep ("accumulation condition"), as presented in Section 4.3, Fig. 4.40, can be used to recover a huge fraction of R. The remaining 10% [155] of the total threshold voltage degradation with a recovery time of 1 µs after 2×10^{5} s of stress are referred to as P (see Fig. 1.34). The size of P strongly depends on the chosen sweep range - the more the sweep enters the accumulation region the smaller the remaining P is. Additionally, P can be removed with a bake at $V_{\mathrm{gs}} = 0\,\mathrm{V}$ and $T = 623\,\mathrm{K}$, while this condition can also lead to a buildup of positive charges in the oxide. The amount of these positive charges can be reduced by a positive gate bias during the bake. Fig. 1.35 shows the physical model [155] explaining the origin of P. Here, a proton binds to the bridging oxygen dependent on the Fermi level and configuration. If the trap energy is higher than the Fermi level during stress, it can be neutralized and be released as H^0. This hydrogen diffuses quickly from the gate to the channel and meanwhile occupies new traps which were not available at the lower gate bias V_{th}. This occupation process creates P. If the local strain is high enough, the neutral hydrogen cannot be detached from the bridging oxygen. Therefore, the hydrogen survives in this configuration during the gate bias sweep. Additionally, (few) by the migrating H^0 depassivated P_{b} centers contribute to P. An increase of P during the bake can be related to the release of additional hydrogen from the poly-gate over a barrier of 2.5 eV ($Si - H$ bond), which can occupy further H-traps. Of course, the system response is reaction-limited due to the high diffusivity of hydrogen [156]. The statistical information about R and P was also investigated in this work for the core class p-MOSFET with $A_{\mathrm{gate}} = 1\,\mu\mathrm{m}^2$ and can be found in Section 4.7.

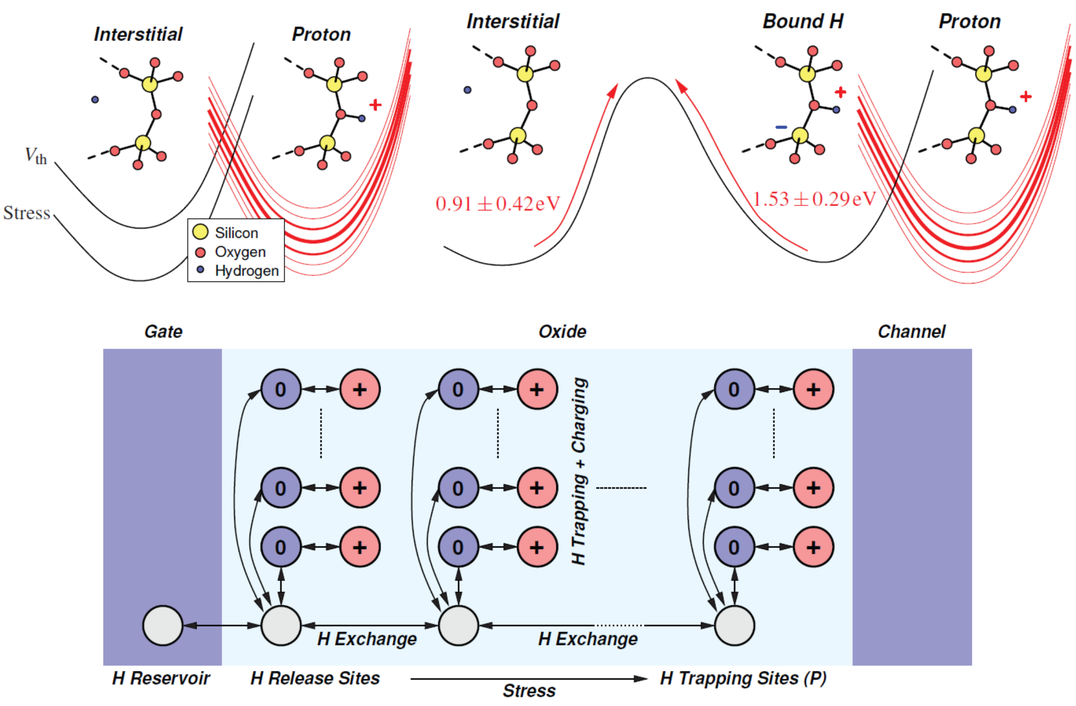


Figure 1.35.: Left: Hydrogen can capture a hole and attach to the bridging oxygen with a broad distribution of energy levels ($\sigma \approx 1\,\mathrm{eV}$). For a specific Fermi level the hole can be emitted and the hydrogen becomes interstitial again. During stress the trap energy level is higher than the Fermi level and, therefore, the hydrogen becomes neutralized. Then, this H^0 quickly migrates from the gate to the channel. The other migration direction is possibly under an accumulative gate bias, where the hydrogen moves from the channel to the gate side. **Middle:** Only Si-O bonds with a high strain disable the release of the attached hydrogen, because of the high energy barriers. That is why, these configurations contribute to P. **Right:** A schematic representation of the hydrogen release model. The trapping sites of hydrogen are connected to the interstitial configuration via barriers and the hydrogen diffusion is reaction-limited. During the bake at $T \geq 623\,\mathrm{K}$ an exchange of additional hydrogen with the reservoir is possible ($Si - H$, 2.5 eV bond barrier). [155]

1.2.6. Time dependent defect spectroscopy (TDDS) and Capture and Emission Time (CET) maps

To statistically analyze the single defects, which constitute NBTI, a special measurement technique was employed by Reisinger in 2010 [157]. This time dependent defect spectroscopy (TDDS) setup enables the charging (stress) of the contributing defects and temporally monitoring their discharging (recovery). The recorded data of small p-MOSFETs ($W \times L \approx 0.01\,\mu\mathrm{m}^2$) can be fit to a physical model, which assumes RC-constants (see Fig. 1.36) for each defect. By that, the electric field as well as the activation energy of the defects can be identified with great precision. The usage of spectral maps of defect classes, rather than single defects for larger devices ($W \times L \geq 1\,\mu\mathrm{m}^2$), facilitates a quantitative degradation and recovery NBTI model of application-like transistors.

To precisely determine the capture and emission time constants τ_c and τ_e, repetitive measurements (typically 100 or more) need to be performed for each examined defect and condition (gate voltage, excitation pulse length, temperature). For instance, the device geometry has a huge impact on the threshold voltage shift caused by a single interface trap (see Table 1.2). A measurement, which is performed with the ultra-fast measurement equipment described in Section 2.5, consists of a charging pulse with $t_\mathrm{s} = 2 \times 10^{-7}\,\mathrm{s}..1 \times 10^{2}\,\mathrm{s}$ stress followed by a continuous measurement of the threshold voltage with $t_\mathrm{r} = 1 \times 10^{3}\,\mathrm{s}$ recovery. Examples of the recorded recovery traces are shown in Fig. 1.37 and Fig. 1.38, which directly demonstrate the stochastic nature of the discharging interface traps.

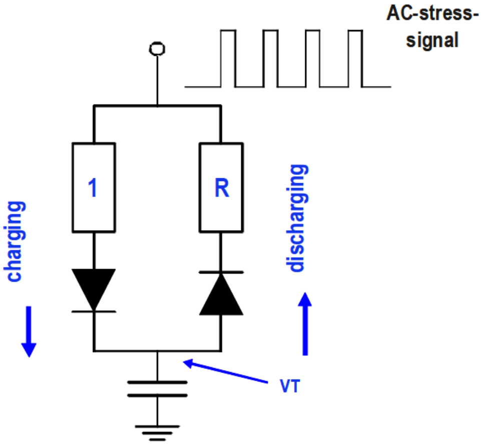


Figure 1.36.: Each defect in a MOSFET constituting NBTI can be represented by a single RC-element with capture and emission time constants τ_c and τ_e. The ratio between both is $R = \tau_\mathrm{e}/\tau_\mathrm{c}$. The stress at the gate of the transistor is also present at the input of the RC-element and leads to a bias marked as VT. A summation of all those biases leads to the total shift of the threshold voltage, ΔV_th. [157]

The occupation probability of a single defect is described by

$$P\left(t_\mathrm{s}\right) = P\left(t_\mathrm{s} = 0\right) \times \left(1 - \mathrm{e}^{\frac{t_\mathrm{s}}{\tau_\mathrm{c}}}\right) \tag{1.116}$$

for the increasing occupancy caused by capture during stress and

$$P\left(t_\mathrm{r}\right) = P\left(t_\mathrm{r} = 0\right) \times \left(\mathrm{e}^{\frac{-t_\mathrm{r}}{\tau_\mathrm{e}}}\right) \tag{1.117}$$

for its decay caused by emission during recovery.

FET-name and dimension /µm	**wide**		**narrow**		**small**	
	W	L	W	L	W	L
	10	0.1	0.2	0.12	0.11	0.1
number of carriers in channel at $V_g = V_{\text{th}} - 0.2\,\text{V}$	15000		370		170	
number #Nit at a densitiy $D_{\text{Nit}} = 1 \times 10^{11}\,\text{cm}^{-2}$	1000		24		11	
ΔNit causing a $\Delta V_{\text{th}} = 50\,\text{mV}$ /cm^{-2}	4.9×10^{11}		4.9×10^{11}		4.9×10^{11}	
makes a number Δ#Nit	4900		120		50	
ΔV_{th} caused by a single trapped carrier (at interface)	0.01 mV		0.43 mV		1.0 mV	

Table 1.2.: The geometry has a huge impact on the threshold voltage shift a single interface trap can cause. The real measured step-heights for the small p-MOSFET vary between 0.2 mV and 5 mV. The charged interface-states density is abbreviated with D_{Nit}. [157]

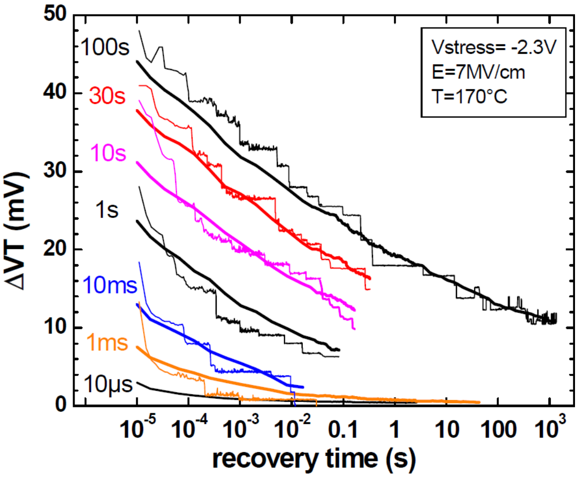


Figure 1.37.: A comparison of a single recovery trace from a narrow device with discrete steps and 25 averaged traces of the same device type. The different step heights for the same recovery time imply that each defect has a specific recovery time depending on the charging pulse time. Because the charging and recovery process are stochastic, the averaged data include all variations of τ_{c} and τ_{e}. [157]

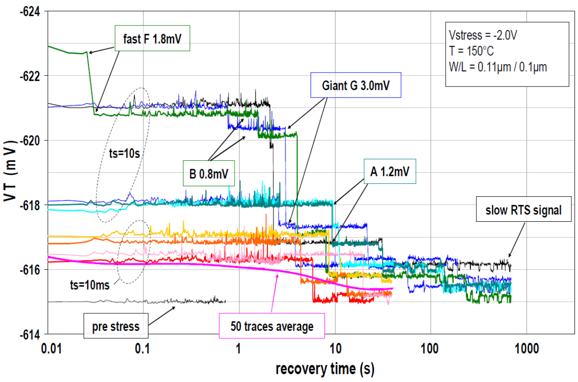


Figure 1.38.: The recovery traces after two stress pulses of different duration show the stepwise recovery of the recorded threshold voltage. An individual defect is identified by its unique step height. The reddish colored traces were measured after $t_{\text{s}} = 10\,\text{s}$, the bluish after $t_{\text{s}} = 10\,\text{ms}$. The trace of 50 averages (pink) depicts the $\mathrm{e}^{\frac{-t_{\text{r}}}{\tau_{\text{e}}}}$ behavior of the occupation probability of a single defect. [157]

The statistical determination of the capture and emission time is shown graphically in Fig. 1.39 and Fig. 1.40. It should be noted that, contrary to the conventional determination of the thermal activation of NBTI in wide MOSFETs [111], the plot for a specific defect exactly follows the Arrhenius behavior. Reisinger et al. further state that the strong thermally activated behavior of individual defects cannot be observed in “large FETs due to the mixing of all the defects” and “all conclusions about non-thermally-

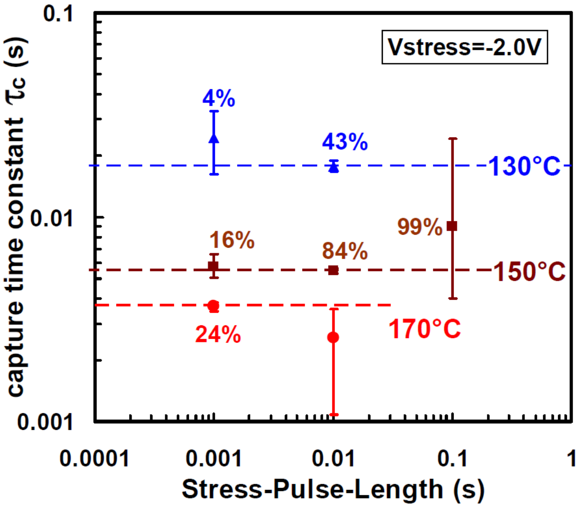


Figure 1.39.: The capture time constant statistics shown for a single defect at three different temperatures. The capture probability ($r = \#\,(\text{capture events})\,/\#\,(\text{attempts})$) is given by the labels. The statistical error for τ_c is minimal for a pulse length with a capture probability between 15% and 90%. [157]

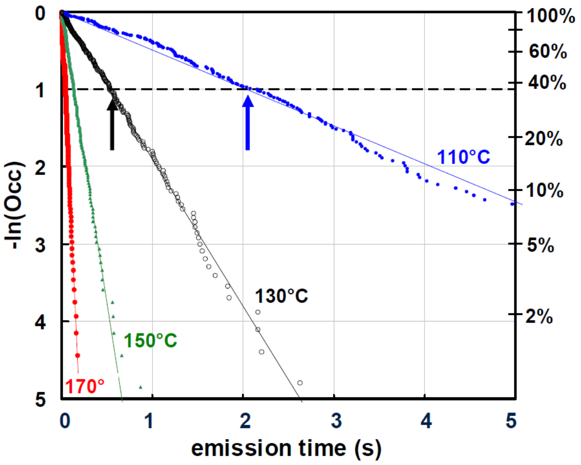


Figure 1.40.: The straight line behavior (according to Eq. 1.117) of the emission time constant statistics for a single defect at four different temperatures ensures that only one defect is contained in each curve. The emission time constants τ_e are denoted by arrows. [157]

activated behavior for short times in large FETs were wrong"[157].

Because the number of defects and their properties remain unchanged during stress [157], the assumptions for single defects in narrow devices can be translated into wide MOSFETs using defect classes (see Fig. 1.41 to Fig. 1.43). In contrast to the measurement of single defects of narrow transistors, now full recovery traces of wide devices are recorded and analyzed to characterize the density of defect classes. As shown in Fig. 1.42, the difference of two recovery traces with different stress times is calculated, following:

$$\frac{\delta \Delta V_\mathrm{th}}{\delta t_\mathrm{s}} = \Delta V_\mathrm{th}\left(t_\mathrm{s,i}\right) - \Delta V_\mathrm{th}\left(t_\mathrm{s,i-1}\right), i = 2 \ldots \#\left(\text{stress times}\right) \tag{1.118}$$

The resulting trace contains all defects which have a capture time constant between 10 ms and 100 ms. Within the next step these traces are divided into one-decade-wide bins, following:

$$\begin{aligned}\frac{\delta^2 \Delta V_\mathrm{th}}{\delta t_\mathrm{s} \delta t_\mathrm{r}} = {}& \left[\Delta V_\mathrm{th}\left(t_\mathrm{s,i}, t_\mathrm{r,j}\right) - \Delta V_\mathrm{th}\left(t_\mathrm{s,i-1}, t_\mathrm{r,j}\right)\right] - \left[\Delta V_\mathrm{th}\left(t_\mathrm{s,i}, t_\mathrm{r,j-1}\right) - \Delta V_\mathrm{th}\left(t_\mathrm{s,i-1}, t_\mathrm{r,j-1}\right)\right], \\ & i = 2 \ldots \#\left(\text{stress times}\right), j = 2 \ldots \#\left(\text{recovery time bins}\right)\end{aligned} \tag{1.119}$$

This leads to a defect class density which consists of all defects having a capture time

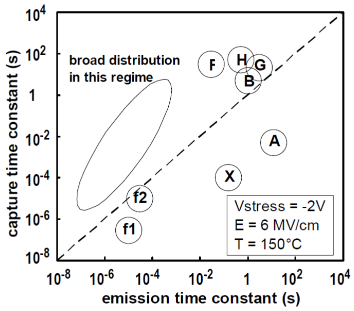


Figure 1.41.: A capture and emission time map of a narrow transistor showing selected single defects. The time constants were determined statistically using TDDS measurement data. [157]

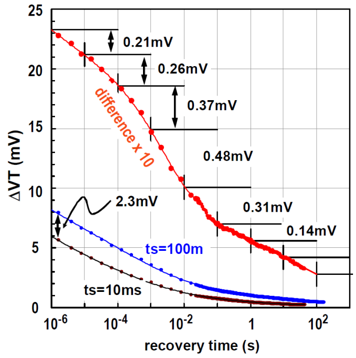


Figure 1.42.: To extract the spectral trap densities, the second order derivative of the threshold voltage shift is determined. This graph is an example for the calculation of the capture time slice between 10 ms and 100 ms in Fig. 1.43. [157]

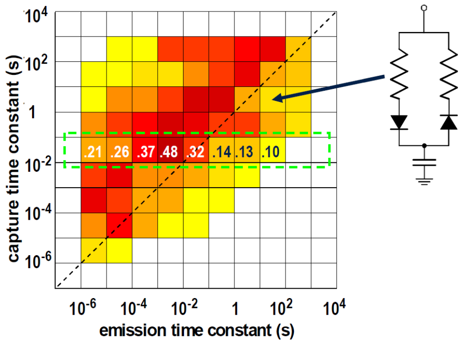


Figure 1.43.: The capture and emission time map of a wide transistor consists of so called defect classes which correspond to an equivalent RC-circuit. Each defect class is a decade-wide bin in capture and emission time space and includes all defects which constitute the calculated second order derivative of the threshold voltage shift, $\delta^2 \Delta V_{\mathrm{th}} / (\delta t_{\mathrm{s}} \delta t_{\mathrm{r}})$. The color is representing the spectral density (in mV) mapped from white to dark red. [157]

between $t_{\mathrm{s,i}}$ and $t_{\mathrm{s,i-1}}$ and a recovery time between $t_{\mathrm{r,j}}$ and $t_{\mathrm{r,j-1}}$, e.g.

$$0.48\,\mathrm{mV} = [\Delta V_{\mathrm{th}}\,(t_{\mathrm{s}} = 100\,\mathrm{ms}, t_{\mathrm{r}} = 10\,\mathrm{ms}) - \Delta V_{\mathrm{th}}\,(t_{\mathrm{s}} = 10\,\mathrm{ms}, t_{\mathrm{r}} = 10\,\mathrm{ms})] \tag{1.120}$$

$$- [\Delta V_{\mathrm{th}}\,(t_{\mathrm{s}} = 100\,\mathrm{ms}, t_{\mathrm{r}} = 1\,\mathrm{ms}) - \Delta V_{\mathrm{th}}\,(t_{\mathrm{s}} = 10\,\mathrm{ms}, t_{\mathrm{r}} = 1\,\mathrm{ms})] \tag{1.121}$$

as shown in Fig.1.43, which is also called a “Capture and Emission Time map” (CET map). In Section 2.5 the further post-processing of the experimental data of this work is presented in detail.

Because all defects are included in the calculated CET map, it is possible to calculate the threshold voltage shift for a given stress signal as an integral over the applied capture and emission time.

1.3. Non-Conducting Stress (NCS)

Contrary to “on” state current stress, which has been studied in great detail during the last decades (see Section 1.1), “off” state current stress (OSS), also known as non-conducting stress (NCS), is reported rather sparsely in the literature [158]. Similar to HCD, with $|V_{\mathrm{ds}}| > V_{\mathrm{nom}}$, during NCS hot-carriers are generated via impact ionization (see Fig. 1.44) at the drain side of the device. The main difference to HCD is that for NCS the gate-drain field ($V_{\mathrm{gs}} = 0\,\mathrm{V}$) leads to a prevailing hot electron injection into the oxide as shown in Fig. 1.45. This degradation mechanism causes a positive shift of the threshold voltage of a p-MOSFET, a changed sub-threshold slope and also off-current degradation.

Because the main focus of this work is on HCD, NBTI and their interaction, the chosen stress voltage regimes cause primarily one of these two degradation mechanisms. Of course, NCS is possible for $|V_{\mathrm{ds}}| > V_{\mathrm{nom}}$, but the maximum drain-source voltages (see Section 3.1 for details) are kept below typical NCS conditions. That is why, only little NCS caused degradation is expected to be contained within the result data.

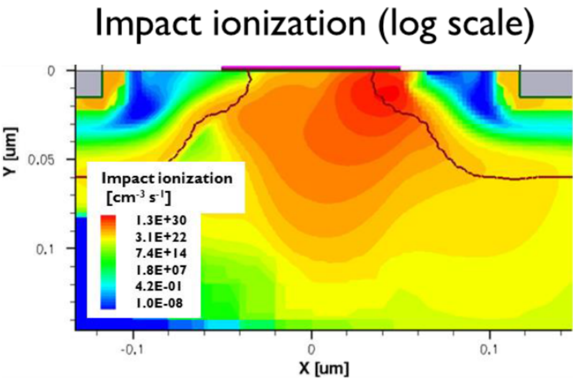


Figure 1.44.: The TCAD simulation of a 65 nm short channel transistor shows severe impact ionization inside the drain junction caused by a strong drain-source field during off state ($V_{\mathrm{gs}} = V_{\mathrm{bs}} = 0\,\mathrm{V}$, $V_{\mathrm{ds}} = -3\,\mathrm{V}$). [158]

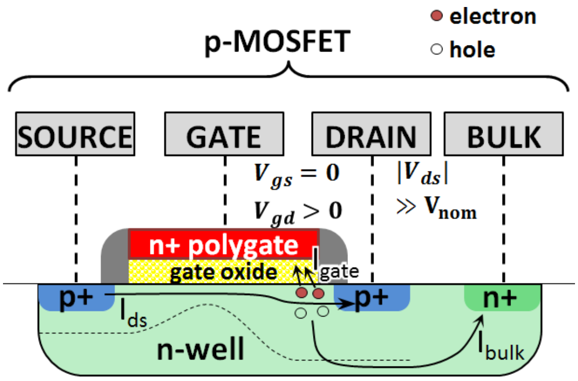


Figure 1.45.: Non-conducting stress: The strong drain-source field leads to impact ionization at the drain side of the device. The generated hot electrons are accelerated towards the oxide by the positive gate-drain field. The electrons are injected into the oxide, if their energy exceeds the SiO_2 barrier of 3.1 eV [158].

2. Experimental setup

2.1. Sample and technology information

All measurements for this work were performed on automotive production quality transistors of a 130 nm smart power technology from Infineon Technologies AG [159]. Although the examined devices are placed together with power devices on the same wafer, they are well isolated via grounded substrates from electromagnetic interference (EMI) with them and show no leakage currents via parasitic transistors (see Fig. 2.1).

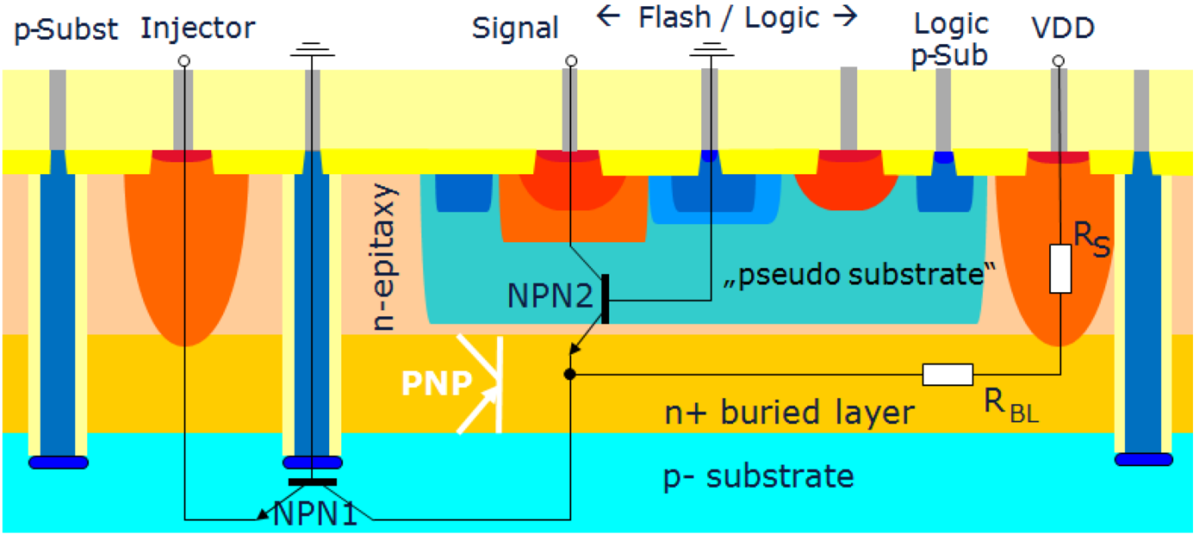


Figure 2.1.: Schematic cross-section of logic, flash and power areas which are embedded on the same wafer with deep trench isolation. The substrate of each device is grounded in order to minimize EMI. [159]

Two different device classes of that technology have been characterized thoroughly, namely 2.2 nm single gate oxide and 5.2 nm dual gate oxide n- and p-MOSFETs with a nitrided dielectric material. The (core) devices with the thin gate oxide are mainly used for digital applications between 1.2 V and 1.5 V, while the thick oxide transistors can be found in analog and I/O circuitry at nominal voltages ranging from 2.5 V up to 3.3 V. The devices provide 10 years operational lifetime with a temperature range between 233 K and 398 K. A schematic cross-section of the CMOS process, which is typical for the analyzed technology [160], is given in Fig. 2.2.

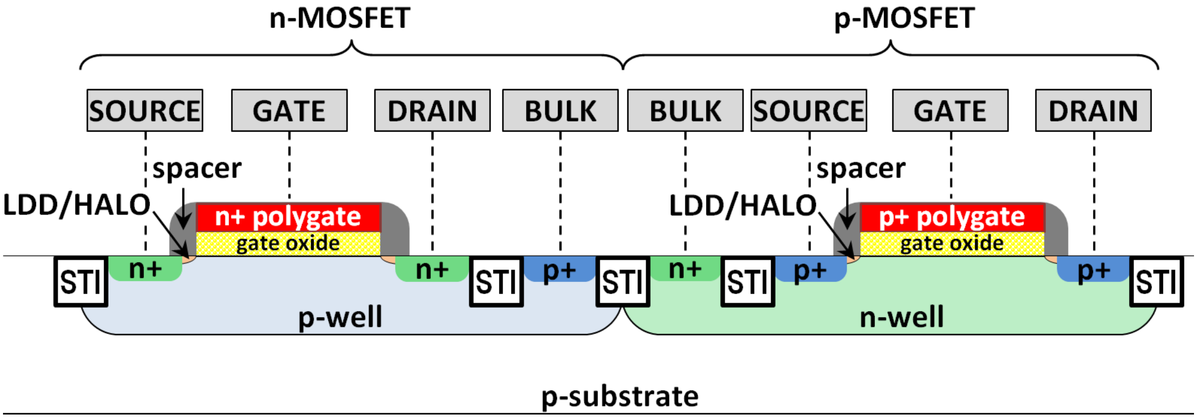


Figure 2.2.: Representative schematic cross-section of the n- and p-MOSFET of the analyzed technology. The p-substrate is embedded as a "pseudo-substrate" into the whole smart power technology process shown in Fig. 2.1.

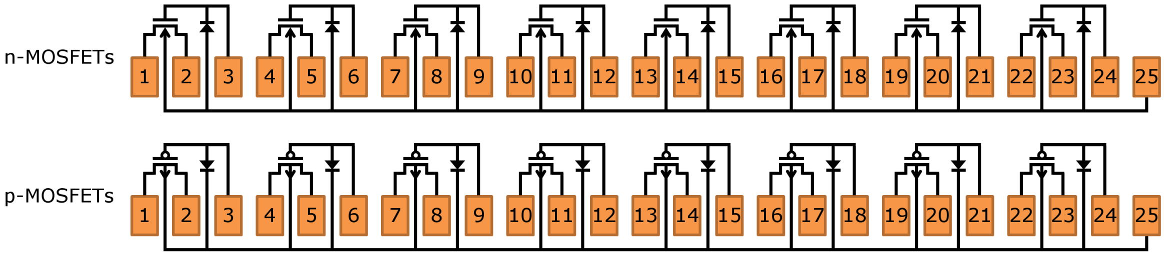


Figure 2.3.: Schematic of the n- and p-MOSFET test modules with 25 pads each for the probe card contact. Every transistor has single gate, drain and source pads and shares a common bulk pad with the other devices. For PID protection a diode is placed between the gate and the substrate.

Several geometries of each device class are available in the examined test structures with different channel lengths, $L = 100\,\mathrm{nm} \ldots 30\,\mu\mathrm{m}$, and widths, $W = 150\,\mathrm{nm} \ldots 30\,\mu\mathrm{m}$, for the core device as well as $L = 370\,\mathrm{nm} \ldots 30\,\mu\mathrm{m}$ and $W = 400\,\mathrm{nm} \ldots 30\,\mu\mathrm{m}$ for the analog device.

To allow a parallel stress and sequential measurement of the devices under test, the test structure has 25 copper pads, which are wired to eight transistors, with single gate, drain, source pads and a common bulk terminal (see Fig. 2.3). All devices can be contacted individually or simultaneously using a probe card with 25 needles. To minimize the risk of plasma induced damage (PID) [161, 162], a protection diode is placed between the gate and the substrate, which discharges plasma process related currents. This and further counter measures ensure that the gate oxides of the transistors were not degraded prior to the reliability measurements.

2.2. Probe station

An Ultracision 880e semi-automatic analytical 200 mm wafer prober, which is also deployed for standard reliability qualifications, was used for all measurements in this work. To minimize the EMI of the probe station the internal power supplies were removed and substituted by external power sources. With that measure the initial noise of about 10 mV was reduced below the detection limit of a Tektronix DPO with a looped standard test probe.

To contact the test module with the utilized source measurement units (SMUs) a probe card with 25 copper-beryllium needles was used. For the so called *standard setup*, used for qualification measurements, the probe card is wired inside the probe station box with BNC wires to an interface board. This board is connected to guarded wires which transmit the signal over a switching matrix to the SMUs inside the HP-4155C analyzer. An ultra-fast measurement equipment [112] was used to record recovery data after stress with a measurement delay of about 1 µs. For that purpose the wiring between the device under test and the ultra-fast measurement equipment had to be very short to minimize the wire capacitance. Therefore the ultra-fast measurement equipment was placed on top of the probe card stage and directly wired to the single needles of the probe card (see Fig. A.1). A multi-purpose probe card which allows either a connection to the standard lab equipment or the ultra-fast measurement equipment is shown in Fig. A.2.

The hot-chuck of the probe station provides temperatures up to about 463 K and is controlled by a Temptronic ThermoChuck® controller. Because no dew eliminating compressed air nozzles are attached to the setup only temperatures above the dew point were used for all tests.

2.3. Standard measurement equipment

To perform hot-carrier degradation measurements, which are independent of the delay between stress and measure mode, the standard lab equipment was used. It consists of the probe station described in Section 2.2 an HP-E5250A switching matrix, an HP-4155C analyzer with four SMUs and two voltage measure units (VMUs), a control PC and a department-wide used software suite for technology qualifications. The software sends the commands needed to run the stress and measurements via a General Purpose Interface Bus (GPIB) connection to the hardware. A typical sequence for a stress-measure-stress experiment is shown in Fig. 2.4.

Because the switching matrix can connect several in- and output signals together it is possible to apply the stress signal to the devices under test in parallel. For a given test structure with n devices this reduces the stress time to $t_{\mathrm{stress}} = 1/n$ and enables

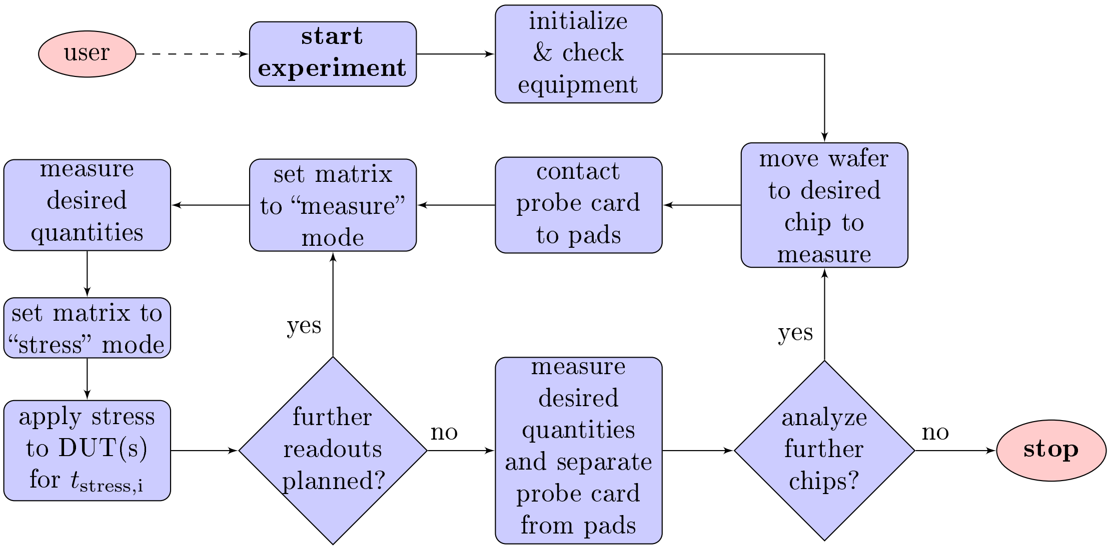


Figure 2.4.: A typical measure-stress-measure sequence employing the ultra-fast measurement equipment. After the user has started the experiment all shown steps are run fully automatically by the measurement software.

a performant setup with respect to time and accuracy. The readouts, of course, must be carried out sequentially whereas the needed readout time t_{readout} depends on the measured quantities. A single operating point, for instance, can be recorded much faster compared to a full linear characteristic of a transistor. The complete time of a measure-stress-measure experiment for k readouts is then given by

$$t_{\text{exp,msm,par}} = t_{\text{readout}} kn + \sum_{i=1}^{k-1} t_{\text{stress,i}}. \tag{2.1}$$

Without a parallel stress applied to the devices under test this would become

$$t_{\text{exp,msm,seq}} = t_{\text{readout}} kn + \sum_{i=1}^{k-1} t_{\text{stress,i}} \times n. \tag{2.2}$$

The resulting time-saving benefit of $\sum_{i=1}^{k-1} t_{\text{stress,i}} \times (n-1)$ is about 20 hours per measured chip for the experiments performed in this work which is about $5 \times t_{\text{exp,msm,par}}$.

The recorded data are saved at two places, a text file and a database. The text-file can either be opened with a viewer available in the lab software suite or be imported by other software. Due to the huge amount of measured data for this work a software environment based on open-source tools (e.g. gawk, gnuplot), named PlotIt, has been programmed to post-process the gathered information (see Fig. A.3).

PlotIt can generate 2-dimensional, 3-dimensional and animated plots to visualize the complex multi-dimensional space of device aging. This mainly includes plots dependent on chosen device parameters (e.g. threshold voltage, linear operating point, characteristics) and axes (e.g. stress time, temperature, stress voltages) as well as wafer maps. Compared to a setup in Origin where a semi-automatic plot (using a template and a little manual fine-tuning) of the drift of two parameters over stress time for a given stress condition takes about 90 seconds PlotIt needs less than three seconds. A further benefit of PlotIt is its flexible program structure which allows an easy adaption to new experimental setups. To generate exponential or power law based drift and lifetime assessments of the recorded data a different department-wide available software package can be used which reads the experimental data from the mentioned database.

2.4. Necessity of short delays during recovery measurements

As demonstrated by Reisinger et al. in [163] the delay between NBTI stress and the measurement of the threshold voltage can lead to huge errors in the lifetime prediction of MOSFETs (see Fig. 2.5). The main reason is that data recorded with standard equipment do not include information of fast recovering traps ($\tau_\mathrm{e} < 300\,\mathrm{ms}$, [112]). This can be illustrated (Fig. 2.6) with recovery data from a W/L = 10 µm/0.1 µm transistor of the 130 nm technology examined in this work.

Therefore all data which underlie recovery effects are recorded with the ultra-fast measurement equipment described in Section 2.5.

2.5. Ultra-fast measurement equipment

Because it is very important to keep the measurement delay $t_\mathrm{delay,msr}$ of recovery traces after NBTI stress as short as possible (Section 2.4) an ultra-fast measurement equipment (after Reisinger, [112]) as shown in Fig. A.1 was assembled and modified for the purposes of this work. The four layer printed circuit board (PCB) consisting of more than 100 single components (see Fig. A.4) was implemented into the probe station (see Fig. A.5).

The measurement principle for homogeneous and inhomogeneous ($V_\mathrm{ds,str} < 0\,\mathrm{V}$) NBTI stress is shown in Fig. 2.7. The core circuit for the measurement consists of an operational amplifier, a sensing resistor and two digital switches. To reach a desired drain-source current the setpoint has to be set to,

$$V_\mathrm{setpt.} = R_\mathrm{s} \times I_\mathrm{ds} \tag{2.3}$$

Then the operational amplifier will zero the difference between its positive and negative input by applying the needed gate bias to the device under test. This gate bias, which is the threshold voltage corresponding to the drain-source current of the measured transistor, is acquired on a logarithmic timebase by an ADC circuit connected to the gate. To ensure a very small load to the gate, the ADC circuit has an additional operational amplifier ("PostAmp") implemented as input stage. The resulting settling time of the feedback loop is about 1 µs (depending on the measured transistor, set-point and sensing resistor). If the transistor would not cause a delay the shortest achievable delay was in the range of the switching speed, in this case for the chosen digital switches about 200 ns.

For inhomogeneous NBTI several changes had to be implemented into the ultra-fast measurement equipment, because the negative signal during stress at the inverting input of the operational amplifier drives its output to the positive operation voltage as depicted in Fig. 2.8. This leads to a strong and about 35 µs long accumulative stress of the device under test after switching to the measurement mode. In addition, $V_{\mathrm{ds,str}}$ has to be grounded directly at the source side of the transistor instead of using the sense resistor. The first version of a solution to fix this problem was to insert a drain-source current simulating resistor in the feedback loop during stress to keep the output of the operational amplifier at a known (negative) value. With two more digital switches it

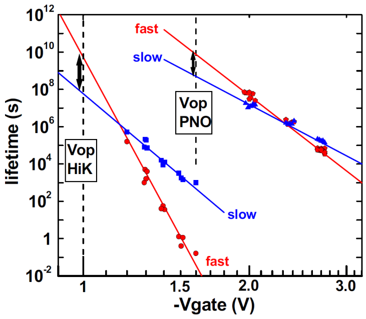


Figure 2.5.: A different delay between NBTI stress and the measurement of the threshold voltage can lead to a several order of magnitudes large wrong prediction of the MOSFET lifetime. The double arrows show the additional lifetime which is obtained employing an ultra-fast measurement technique with a delay about 1 µs compared to standard equipment data. [163]

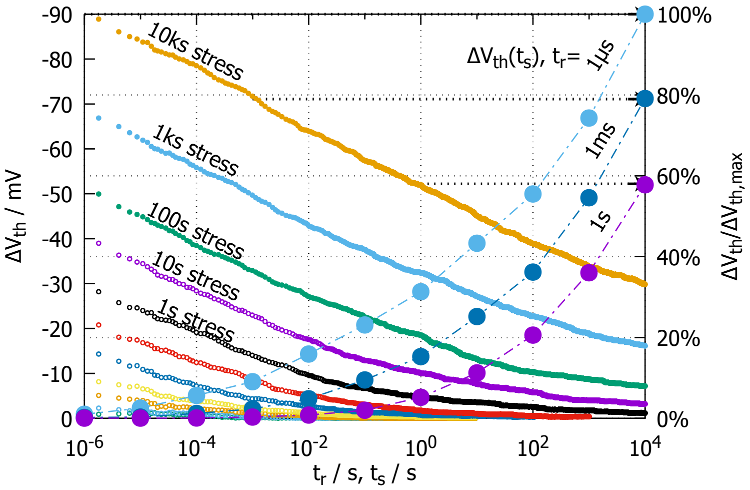


Figure 2.6.: Recovery traces of the 130 nm technology examined in this work with nitrided 2.2 nm gate oxides and W/L = 10 µm/0.1 µm show the huge impact of the chosen measurement equipment on the recorded threshold voltage. More than 40% of the threshold voltage shift were lost if the readout delay was 1 s. Standard setups have a delay about several 100 ms. Dashed lines beneath recovery traces are S-function fits (see Eq. 2.4 on page 63 for details) of the raw data (symbols). The degradation emulating different measurement delays is plotted with dashed-dotted lines.

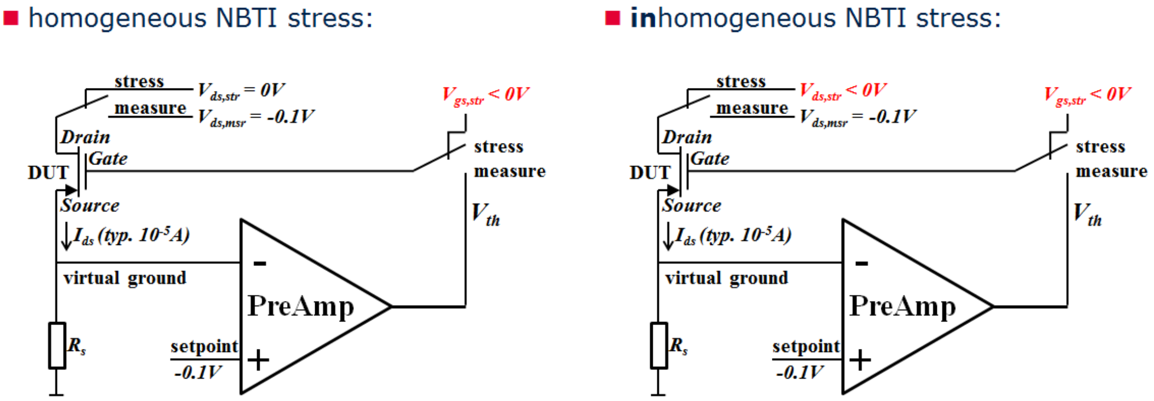


Figure 2.7.: Experimental setup to measure the threshold voltage of a transistor after NBTI stress within 1 µs. Left: homogeneous NBTI with $V_{\mathrm{ds,str}} = 0\,\mathrm{V}$, right: inhomogeneous NBTI with $V_{\mathrm{ds,str}} < 0\,\mathrm{V}$. The driven gate bias of the operational amplifier “PreAmp” is acquired by an ADC circuit (not shown) on a logarithmic timebase.

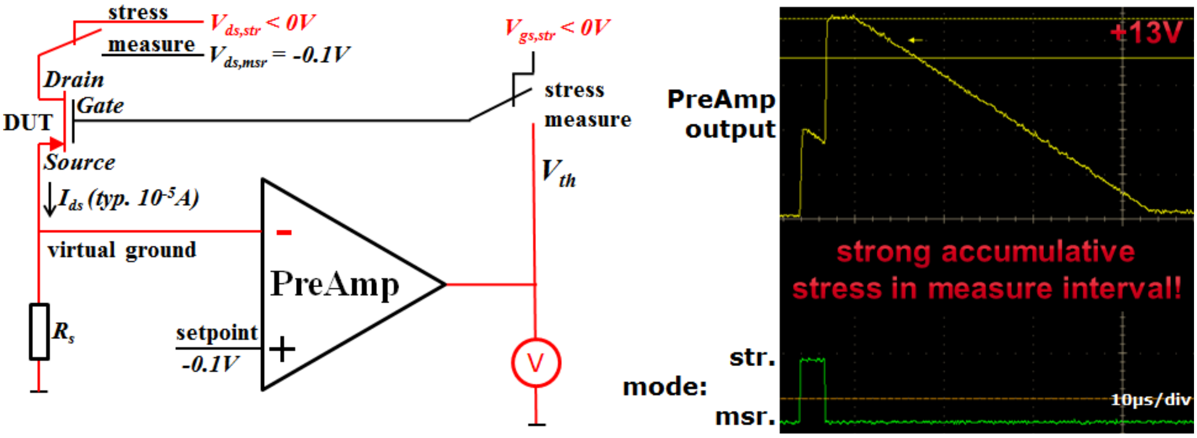


Figure 2.8.: For an inhomogeneous NBTI stress with $V_{\mathrm{ds,str}} < 0\,\mathrm{V}$ the original design of the ultra-fast measurement equipment shows a positive signal equal to the positive operation voltage at the output of the operational amplifier after switching from stress to measure mode. The resulting gate bias causes an about 35 µs long accumulative stress signal to the device under test.

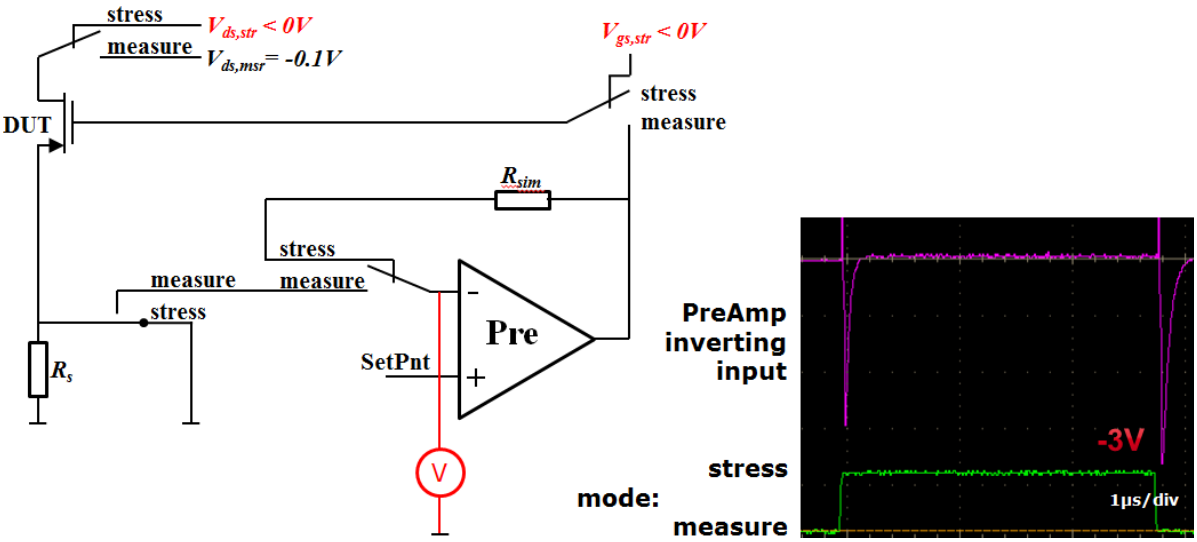


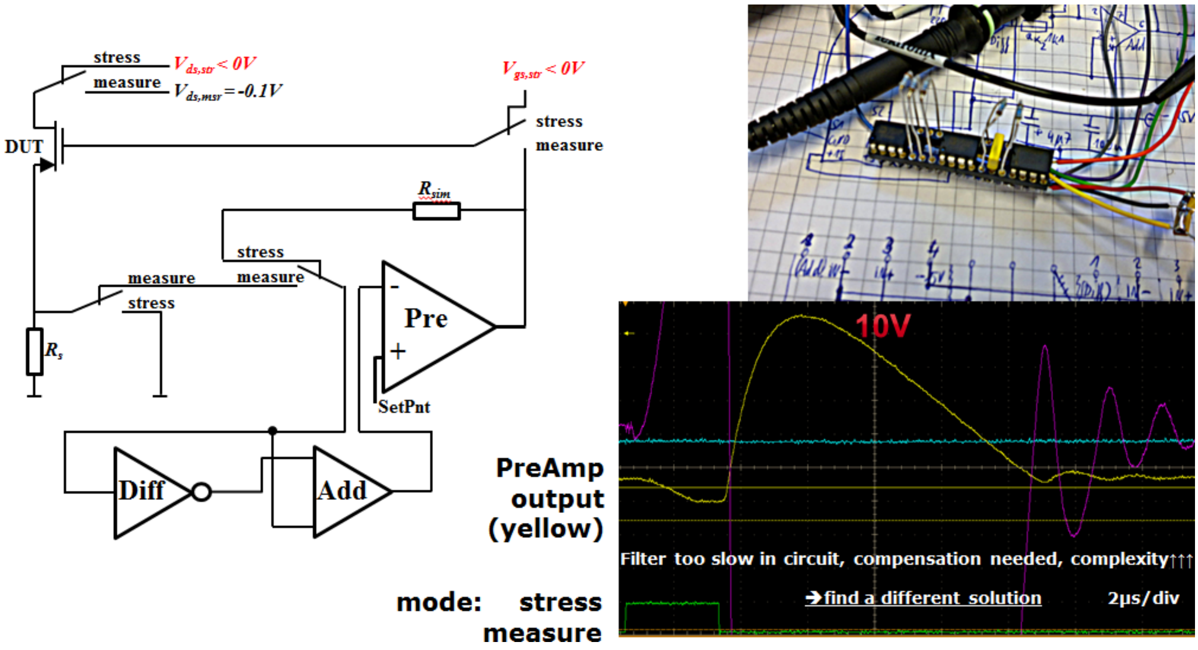
Figure 2.9.: First improved experimental setup to measure the threshold voltage of a transistor after inhomogeneous NBTI stress within 1 µs. The insertion of two switches enables both stress to the device under test as well as keeping the operational amplifier at a known value. Anyhow, the inverting input signal of the operational amplifier shows 200 ns long spikes about −3 V after a change of the operation mode from stress to measure. This is mainly caused by the timing of the switches and their resulting unsynchronized switching behavior.

was possible to realize this setup but the resulting signal at the inverting input of the operational amplifier showed severe spikes about −3 V as represented in Fig. 2.9. Those spikes mainly originate from the unsynchronized switching procedure of the additional switches. Because the timing depends on the loads, which must be anticipated to change during or after stress, the accurate synchronization of the timing of the switches is very complicated.

The first improvement of the initial solution introduces active filters into the circuit design. Here the spikes and their inverted signal are fed into an additional operational amplifier. In contrast to the fast operation of the core circuit the active filter correction unfortunately is too slow and causes a different positive signal at the output of the operational amplifier "PreAmp". As such, the preferred solution would be to use passive parts substituting the active filter.

Based on this attempt, two RC elements with $\tau = 220\,\text{ns}$ were installed to replace the active filter according to Fig. 2.11 to minimize the height of the spikes. The resulting signal of the "PreAmp" settles within 2 µs when switching from measure to stress mode and within 1.5 µs vice versa.

In contrast to the standard lab hardware with $t_{\text{delay,msr}} \approx 0.1s$ this highly specialized measurement equipment minimizes $t_{\text{delay,msr}}$ to about 1 µs. As described in Section 2.4 measurement data for stress times below 300 ms can only be recorded with such an equipment. In the following, accurate CET maps which are in the focus of this work with



Filter too slow in circuit, compensation needed, complexity↑↑↑

→find a different solution

Figure 2.10.: An active filter is employed to eliminate the spikes caused by the switching behavior of the initial solution for inhomogeneous NBTI stress. Unfortunately, this ansatz needs further compensation with a disproportionate effort compared to more practical solutions.

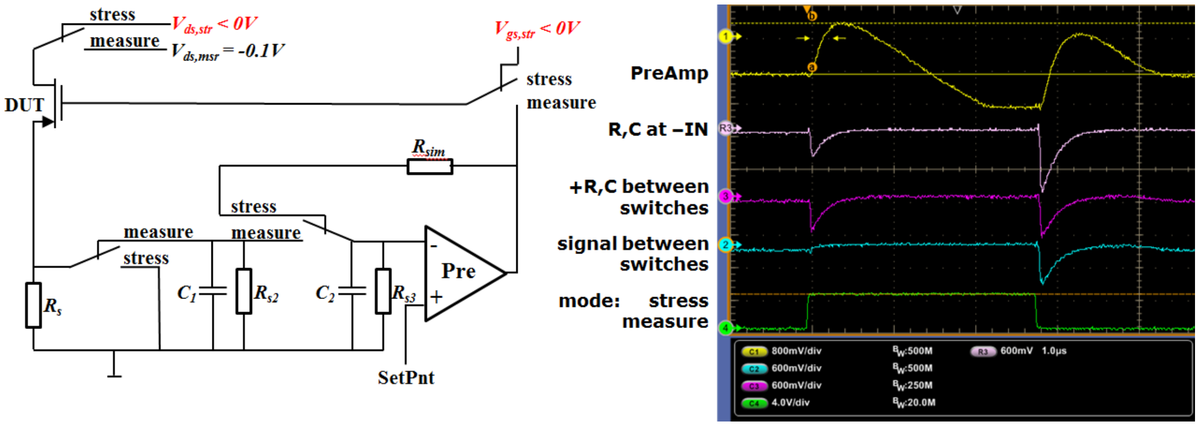


Figure 2.11.: Two RC elements with $\tau = 220\,\text{ns}$ minimize the spikes caused by the switching behavior of the initial solution for inhomogeneous NBTI stress. With respect to the measured transistor type (analog p-MOSFET with $V_{\text{nom}} = 3.3\,\text{V}$) and compared to the unfiltered initial version the "PreAmp" signal shows a highly acceptable signal quality.

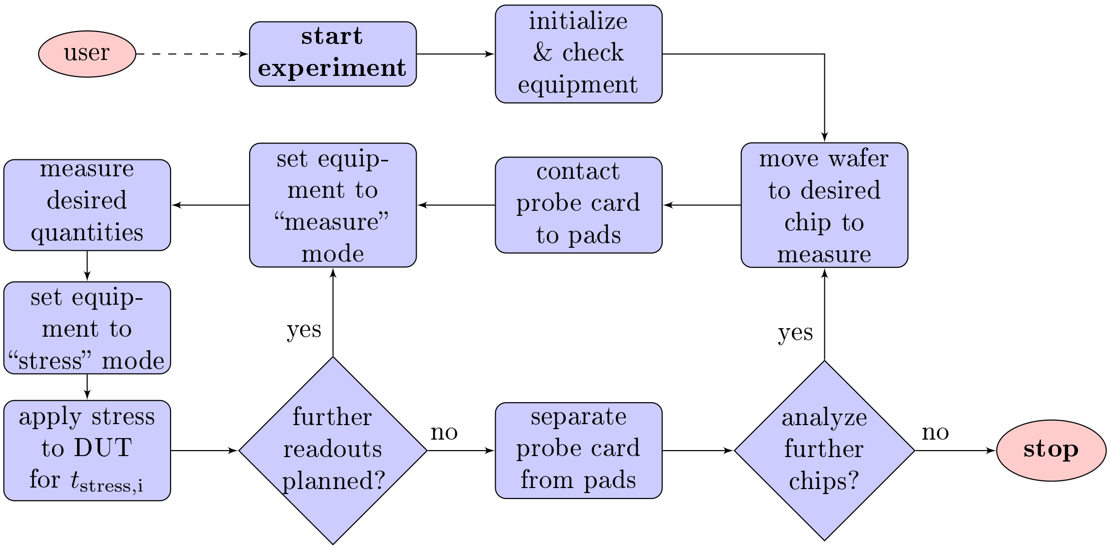


Figure 2.12.: A typical measure-stress-measure sequence employing the ultra-fast measurement equipment. After the user has started the experiment all shown steps are executed fully automatically by the measurement software.

t_s starting at 1 µs can only be measured using this ultra-fast measurement equipment.

Several additional requirements must be met to ensure a correct measurement. This includes for instance the selection of a correct sensing resistor at the source side, a suitable set-point for the readout of the threshold voltage and a known settling behavior of the amplifier circuit for the measured transistor.

A typical measure-stress-measure experiment is following a sequence as shown in Fig. 2.12 and is executed fully automatically after the user has started it. In addition, the ultra-fast measurement equipment has the possibility to switch to the so called "relax mode" where a third voltage - besides the preset measure and stress biases - can be applied to the gate of the transistor. By that, for instance, accumulative pulses can be realized between stress and measurement. An additional feature of the ultra-fast measurement equipment is a built-in arbitrary waveform generator which allows the imitation of application-like stress conditions (see Section 2.8 for further details).

2.6. Data post-processing and CET map generation

To generate a CET map the raw measurement data first needs to be post-processed for the examined transistor geometry of the core device ($A_{\text{gate}} = 1\,\mu\text{m}^2$). The main reason is the noise which is rather distinct for small transistors [164, 165] due to the stochastic nature of charge exchange. The CET map which is depicted in Fig. 2.13 was generated from the raw data without any post-processing. The result shows many artifacts which hamper a qualitative interpretation of the map.

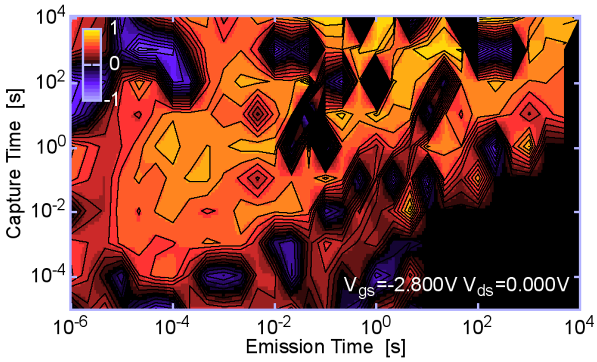
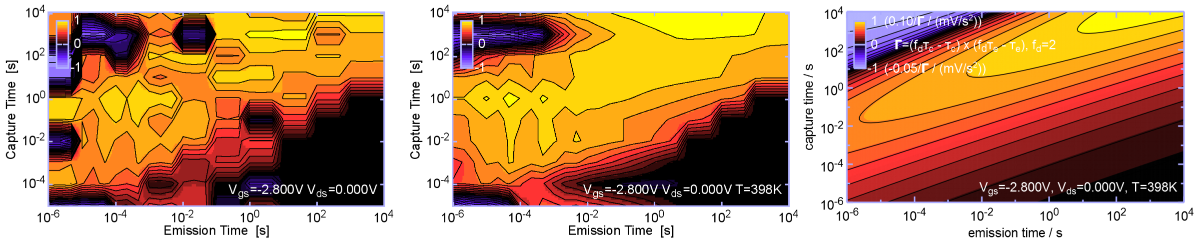


Figure 2.13: The CET map of the core device under homogeneous NBTI stress generated from raw data without any post-processing. The various artifacts originate from the noise of the transistor which depends reciprocally on the gate area ($A_{\text{gate}} = 1\,\mu\text{m}^2$).

Therefore the following steps were introduced to generate a CET map from the measured recovery traces:

1. read measurement information (chip, module, geometry, voltages, temperature)
2. extract recovery traces from raw data
3. compress recovery traces to a reduced logarithmic temporal resolution
4. fit the raw data with
   a) a logarithmic function for each stress and recovery time decade
   b) an S-shape function (data smoothing) per stress time decade
   c) an S-shape function with stress time dependent parameters
5. calculate $\delta^2\Delta V_{\text{th}}/\left(\delta t_{\text{s}}\delta t_{\text{r}}\right)$ from the fitted data
6. plot normalized CET map with $\operatorname{sign}(g)\log(1+\kappa\,|g/g_{\max}|)/\log(1+\kappa)$ [166]

The logarithmic and normalized CET map with $\kappa = 100$ shows all important details about the relevant defect classes $g\,(t_{\text{s}}, t_{\text{r}})$ and is presented in Fig. 2.14 for the three different fit algorithms employing either a logarithmic (Fig. 2.14(a)) or an S-shape fit function (Fig. 2.14(b) and (c)). The S-shape fit can be formulated independent (Section 2.6.1) as well as dependent (Section 2.6.2) of the stress time.



(a) logarithmic fit per stress time decade
(b) S-shape function fit per recovery trace
(c) S-shape function fit, stress time dependent parameters

Figure 2.14.: CET maps generated from post-processed measurement data using either a (a) logarithmic fit function per recovery time decade, (b) S-shape fit function for the whole recovery trace or (c) employing stress time dependent fit parameters to smooth out the noise of the examined transistor.

2.6.1. S-shape function based stress time independent fit

If the data are fitted per recovery time decade with a logarithmic function the calculated CET map results in a very irregular picture (see Fig. 2.14(a)). Because the discernible irregularities cannot be identified in the raw-data they must be considered fitting artifacts. That is why, a log-logistic distribution, called “S-shape function” in the following, is preferred to generate the CET maps presented in this work. This function is given by:

$$\Delta V_{\mathrm{th}}\left(t_{\mathrm{r}}\right)=\frac{a}{1+b\mathrm{e}^{\left[ct_{\mathrm{r}}^{d}\right]}} \tag{2.4}$$

with the four fit parameters a, b, c and d. For an infinite recovery time and $c<0$ it follows:

$$\Delta V_{\mathrm{th},\infty}=\lim_{t_{\mathrm{r}}\to\infty}\frac{a}{1+b\mathrm{e}^{\left[ct_{\mathrm{r}}^{d}\right]}}=a \tag{2.5}$$

By setting t_{r} to 0 parameter b can also be expressed by the following relation,

$$\Delta V_{\mathrm{th},0}=\frac{\Delta V_{\mathrm{th},\infty}}{1+b\mathrm{e}^{\left[c0^{d}\right]}}\Leftrightarrow b=\frac{\Delta V_{\mathrm{th},\infty}}{\Delta V_{\mathrm{th},0}}-1, \tag{2.6}$$

which allows a physical interpretation of this parameter.

Applying those two substitutions, the resulting fit formula is given by:

$$\Delta V_{\mathrm{th}}\left(t_{\mathrm{r}}\right)=\frac{\Delta V_{\mathrm{th},\infty}}{1+\left(\frac{\Delta V_{\mathrm{th},\infty}}{\Delta V_{\mathrm{th},0}}-1\right)\mathrm{e}^{\left[ct_{\mathrm{r}}^{d}\right]}} \tag{2.7}$$

The most significant function adjustment by each fit parameter is shown in Fig. 2.15. According to Eq. 2.5 the non-recoverable (“permanent”) part (after $1\times10^{4}\,\mathrm{s}$) of the threshold voltage drift is covered by parameter $\Delta V_{\mathrm{th},\infty}$ which also increases with the applied drain bias for inhomogeneous NBTI or the mixture of NBTI and HCD. The extrapolation from the starting value of ΔV_{th} to $t_{\mathrm{r}}=0$ is given by parameter $\Delta V_{\mathrm{th},0}$ and its value depends on the chosen gate stress bias and the stress time. Unfortunately, due to the present measurement delay this data point cannot be measured by a measure-stress-measure setup. The longer the temporal offset of the recovery trace is (which is the case for inhomogeneous NBTI as well as for the mixture of NBTI and HCD) the smaller is the absolute value of the exponential parameter “c” and the steeper the slope of the recovery trace is the higher the exponential parameter “d” becomes.

2.6.2. S-shape function based stress time dependent fit

Because the four fit parameters as depicted in Fig. 2.15 show a strong stress time dependence, a third method to generate CET maps is formulated via an extension of the

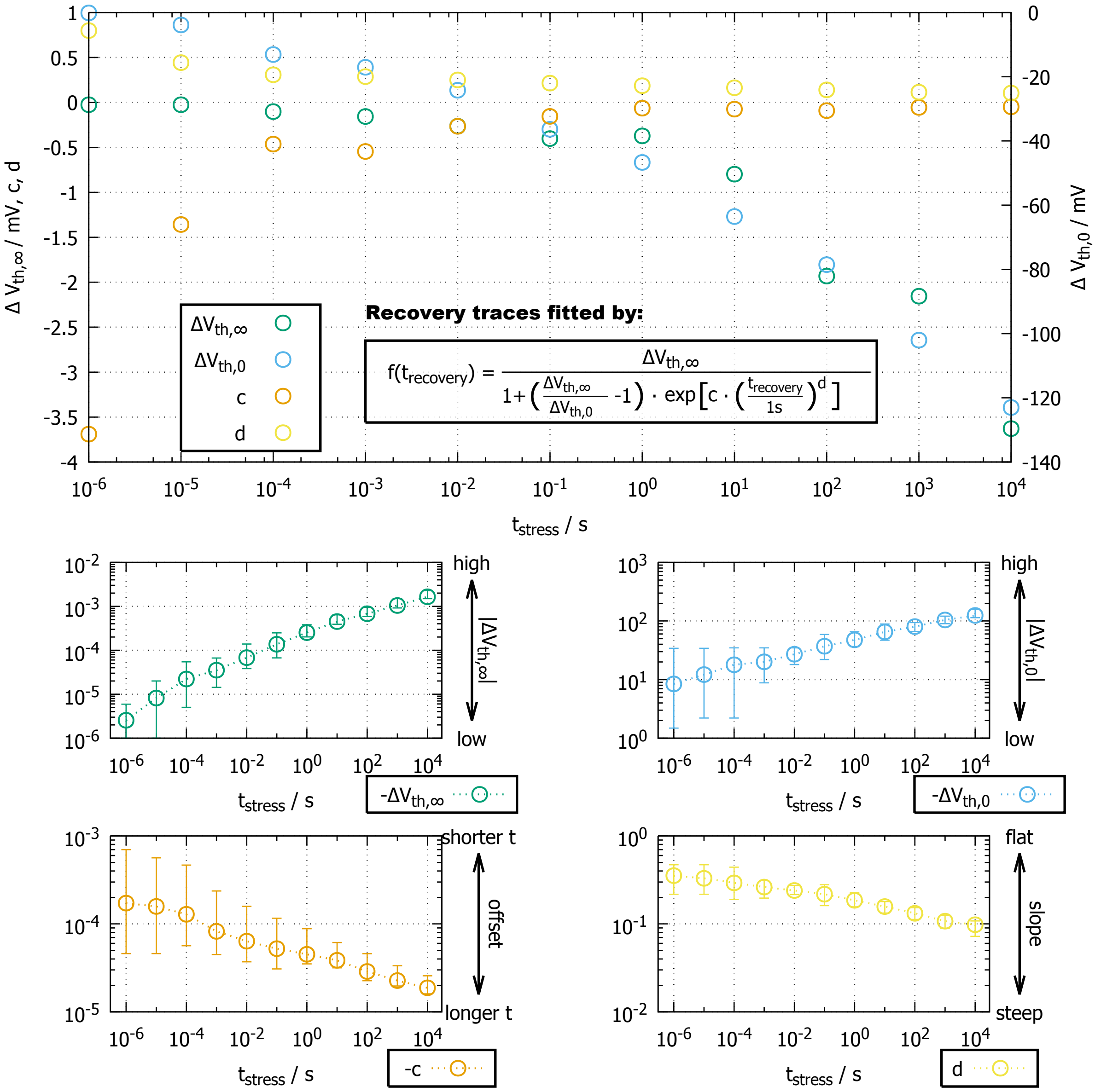


Figure 2.15.: **Top:** The four parameters of the S-shape function which are used to fit the raw-data of each measured recovery trace strongly depend on the stress time.
Bottom: The error bars indicate the value range of each parameter for 11 recovery traces (after $t_s = 10^{-6}\,$s...$10^4\,$s) averaged over 10 CET map measurements with the same stress conditions on adjacent chips. The mean value shows the general trend of each parameter value over stress time.

second method by employing stress time dependent parameters $\Delta V_{\mathrm{th},\infty}(t_\mathrm{s})$, $\Delta V_{\mathrm{th},0}(t_\mathrm{s})$, $c(t_\mathrm{s})$ and $d(t_\mathrm{s})$. This time dependent approximation of the fit parameters should be considered to generate qualitatively interpretable CET maps. It is known that the degradation (ΔV_th) of the devices under test follows a power law with t_s. This implies that the first two parameter, $\Delta V_{\mathrm{th},\infty}(t_\mathrm{s})$ and $\Delta V_{\mathrm{th},0}(t_\mathrm{s})$, may be approximated with a power law as well. For the sake of simplicity, all four fit parameters are approximated by power laws, i.e.

$$P_i(t_\mathrm{s}) = f_i t_\mathrm{s}^{g_\mathrm{i}}. \tag{2.8}$$

This yields

$$\Delta V_\mathrm{th}(V_\mathrm{gs,s}, t_\mathrm{r}, t_\mathrm{s}) \approx \frac{f_1 t_\mathrm{s}^{g_1}}{1 + \left(\frac{f_1 t_\mathrm{s}^{g_1}}{f_2 t_\mathrm{s}^{g_2}} - 1\right) e^{\left(f_3 t_\mathrm{s}^{g_3}\right) t_\mathrm{r}^{\left(f_4 t_\mathrm{s}^{g_4}\right)}}} \tag{2.9}$$

for the threshold voltage shift being dependent on the stress and recovery time.

As shown in Fig. 2.16 the value of each f_i and g_i is limited to a narrow range for the investigated 10 transistors of adjacent chips and also the resulting fit parameters $P_i(t_\mathrm{s})$ show a narrow distribution. Compared to the stress time independent formulation the ranges of the parameters $P_i(t_\mathrm{s})$ show a steady decrease with increasing stress time. This decrease of the parameter ranges for long stress times is mainly due to the measurement uncertainty which decreases with the stress time while the determined threshold voltage shift increases. The stress time dependent formulation minimizes this effect because it takes all data points during the fit into account, thereby also smoothing the measurement uncertainty for short stress times. Of course, individual properties of the single recovery trace may be smoothed by that method, too. But for a qualitative interpretation of the CET maps this fitting method is highly advantageous, because it focuses on the main properties of the recovery trace ensemble and suppresses any residual noise due to its continuous formulation in recovery (Eq. 2.7) and stress time (Eq. 2.8).

Comparisons between the raw-data and the fit of recovery traces after homogeneous NBTI stress and the mixture of NBTI and HCD are depicted in Fig. 2.19 and 2.20 and reveal how well and flexibly the used S-shape function adapts to the measurement data employing either a stress time independent or stress time dependent fit. The benefit of the stress time dependent fit is the opportunity to extrapolate the results to any reasonable stress and recovery time. An extrapolated CET map from the data of Fig. 2.20 (a) is depicted in Fig. 2.18.

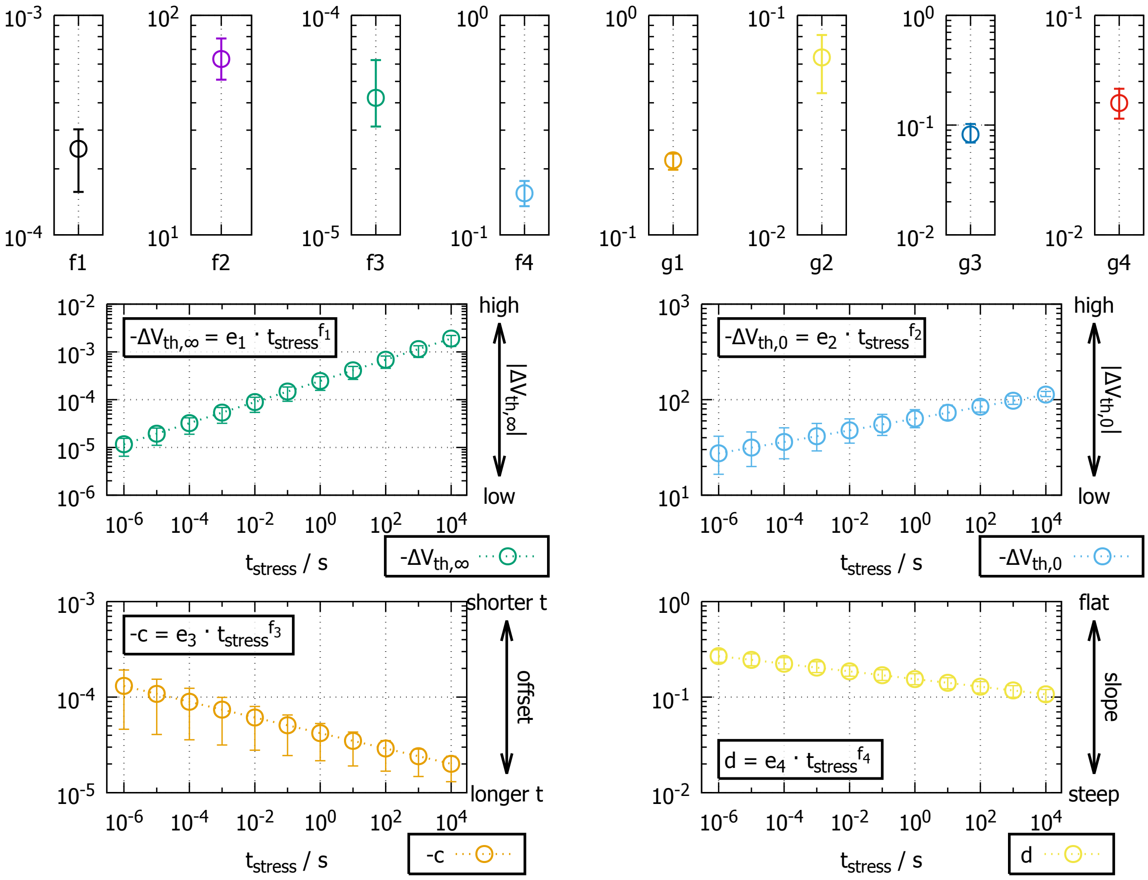


Figure 2.16.: **Top:** The eight input parameters which are needed to calculate the four S-shape parameters for **any** reasonable stress time show a narrow distribution for the 10 investigated p-MOSFETs stressed with homogeneous NBTI.
Bottom: The four calculated output S-shape parameters have a smaller value range compared to Fig. 2.15 (bottom) because the measurement uncertainty for short stress times is smoothed via the 3-dimensional fit over the whole recovery trace field.

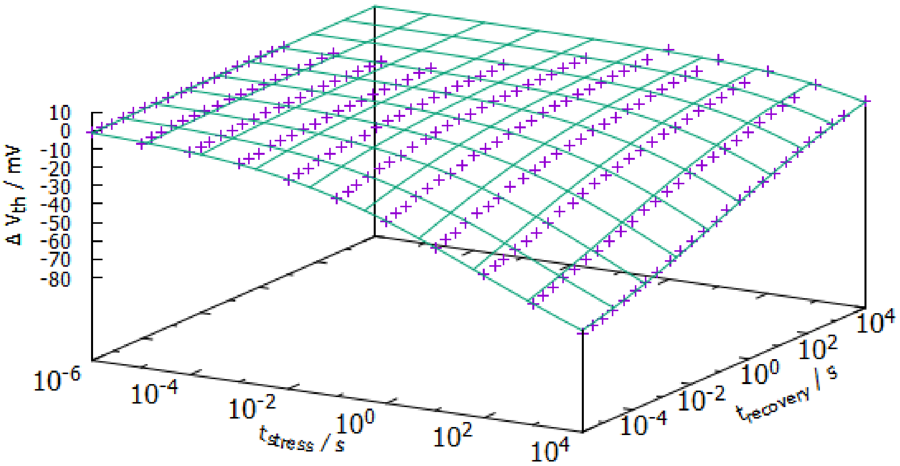


Figure 2.17.: The 3-dimensional fit (green mesh) of the recovery trace data (purple symbols) via the eight stress parameters presented at the top of Fig. 2.16 is typically performed within a few seconds.

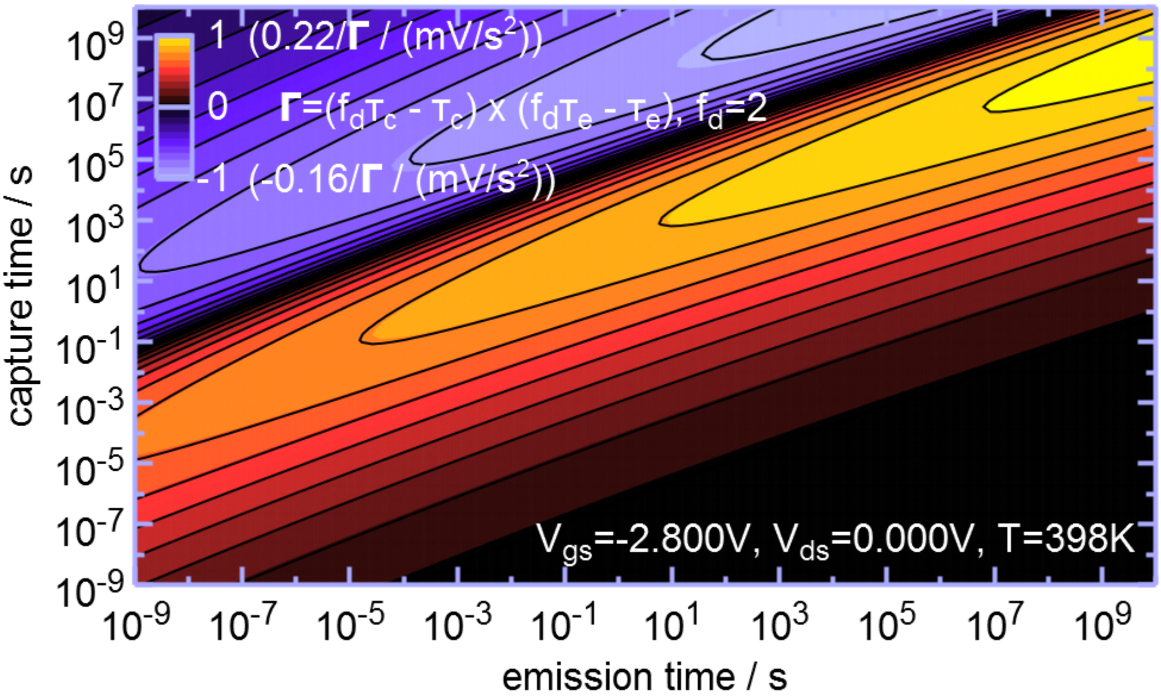
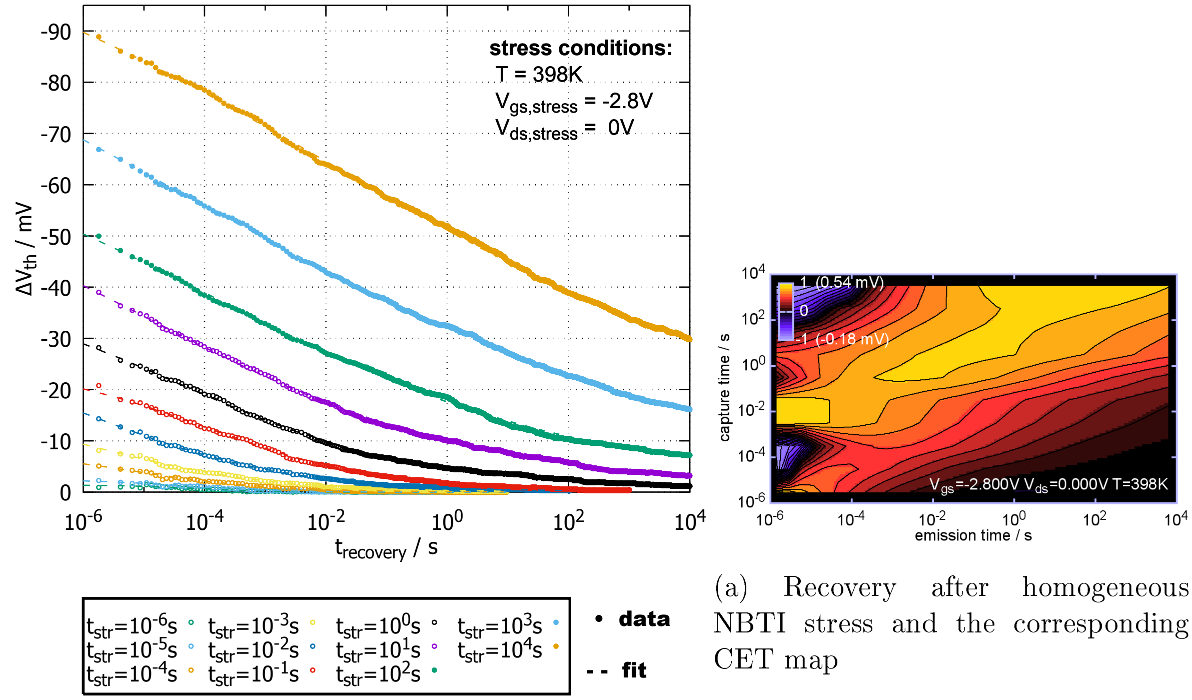
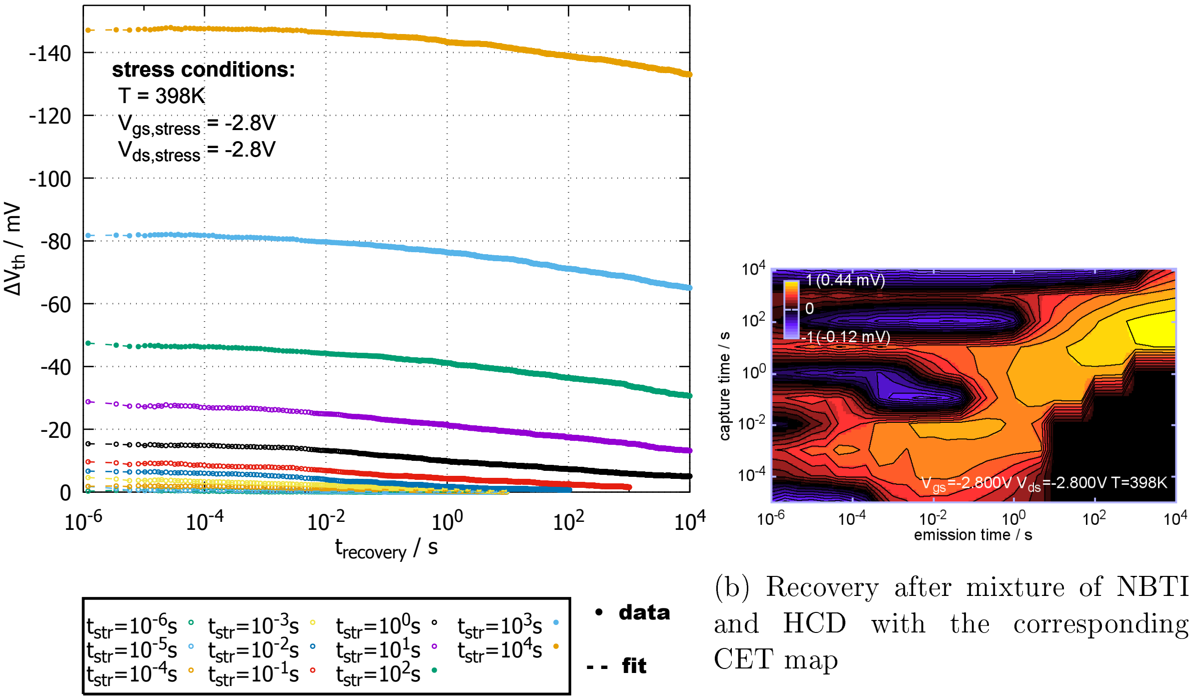


Figure 2.18.: The CET map of the core device under homogeneous NBTI stress extrapolated from the raw data fit (see Fig. 2.20 (a)) to a wide range of stress and recovery time.

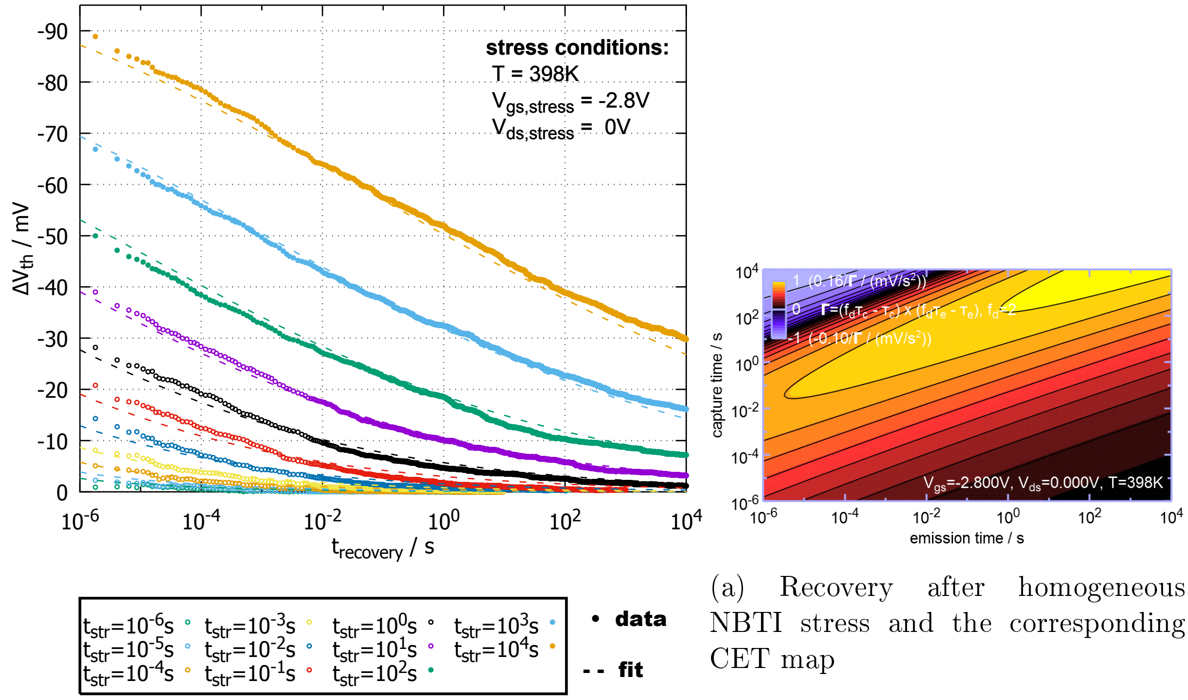


(a) Recovery after homogeneous NBTI stress and the corresponding CET map

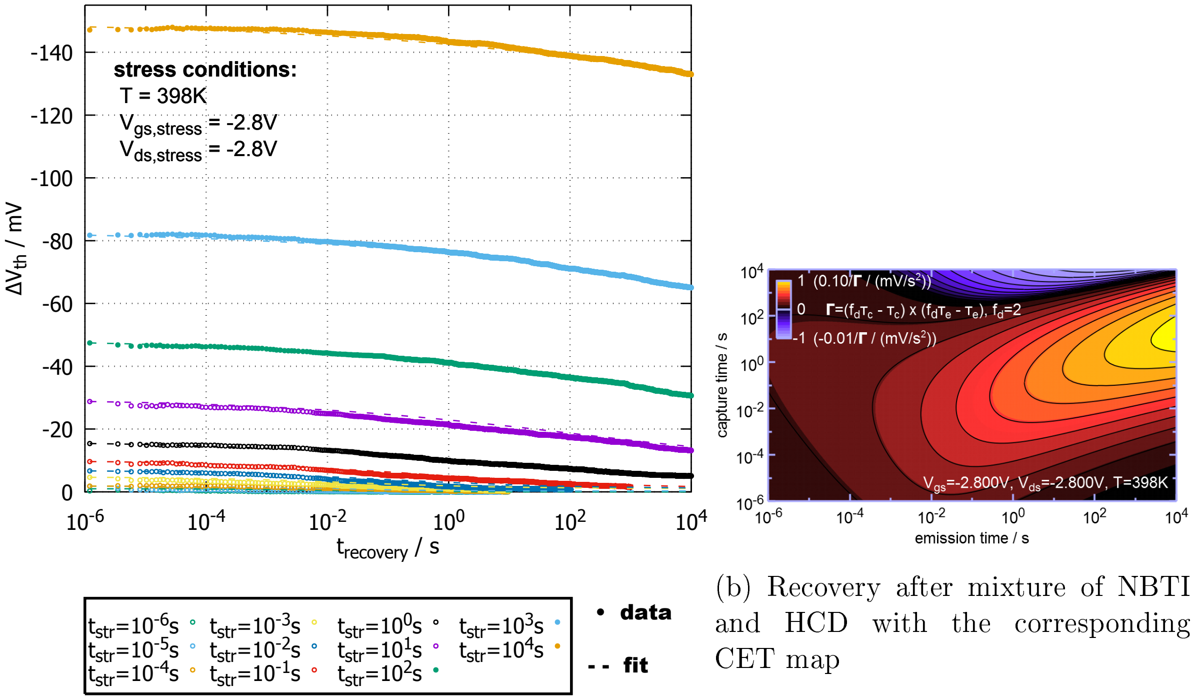


(b) Recovery after mixture of NBTI and HCD with the corresponding CET map

Figure 2.19.: Comparison between raw-data (symbols) and the fit (curves) with stress-time independent parameters of recovery traces demonstrating the high flexibility of the used S-shape fit function. The fits adapt to the recovery traces after different degradation mechanisms (a)/(b) very well and successfully smooth the residual noise of the recorded recovery traces. The CET maps still include some variability of the transistor. This is mainly caused by the independence of the fit parameters of each single recovery trace thus resulting in a discontinuous second order derivative (following Eq. 1.119).



(a) Recovery after homogeneous NBTI stress and the corresponding CET map



(b) Recovery after mixture of NBTI and HCD with the corresponding CET map

Figure 2.20.: Comparison between raw-data (symbols) and the fit (curves) of recovery traces demonstrating the high flexibility of the used fit method employing stress-time dependent parameters. The fits adapt to the recovery traces caused by different degradation mechanisms (a)/(b) very well and successfully smooth the residual noise of the device under test. For the second derivative picture only this fit allows smooth graphs because the stress time dependence is formulated via continuous functions (Eq. 2.8).

2.7. Consideration of worst case conditions

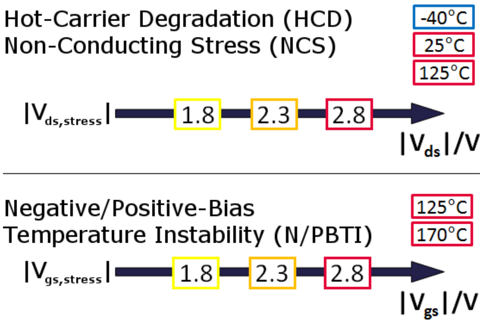


Figure 2.21.: Typical acceleration voltages and temperatures applied to a 1.5 V class transistor to activate different stress mechanisms. For HCD the $V_{\mathrm{gs,stress}}$ causing the maximum degradation has to be determined for each $V_{\mathrm{ds,stress}}$ by a pre-characterization to ensure that the device degradation is never underestimated during reliability assessment.

For a technology qualification it is of capital importance that the device degradation is never underestimated to ensure a faultless operation of the designed circuit over lifetime. Generally, this is achieved by using so called worst-case stress conditions for each stress mode (e.g. HCD, NBTI). At this point the terminal voltages and temperature, which cause the most significant degradation of a single mode, are applied to the device under test. In Fig. 2.21 typical acceleration voltages of a 1.5 V class transistor for HCD, N/PBTI and NCS can be found.

HCD is pronounced for a high drain-source voltage and a technology and device dependent gate-source voltage. To find the worst-case a high drain-source voltage is applied to the device under test while the gate-source voltage is sweeped. For long-channel devices the monitored substrate (n-MOSFET) or gate (p-MOSFET) current will show a maximum for a specific gate-source voltage which refers to the worst-case because the number of degradation causing carriers (impact ionization rate) is proportional to the measured current. Short-channel transistors are known to show the highest degradation for the maximum carrier flux which is obtained for an identical drain- and gate-source voltage. In addition to the gate voltage also the temperature activation can be roughly divided into two groups, long-channel devices show more degradation at low temperatures and short-channel devices at elevated temperatures.

The worst-case for BTI is found at elevated temperatures and a high gate bias whereas all other terminals are kept on the same electric potential (e.g. ground).

A third mechanism is called Non-Conducting Stress (NCS) and occurs for short-channel transistors in the off-state with an additional high drain-source bias. It is caused by carriers which originate from the drain-bulk area able to exceed the source-bulk potential barrier. If the channel is short enough or the drain-source voltage is high enough a merging of the source-bulk and drain-bulk depletion regions occurs and the resulting lowering of the potential barrier leads to a punch-through current. This effect is called Drain Induced Barrier Lowering (DIBL) [167] and the consequential high-energetic carriers can create interface states thus changing the device characteristics similar to HCD.

Although the worst-case concept was originally intended to emphasize a single stress mode and to allow a separated assessment of HCD and BTI, this idea fails for short channel devices. For these transistors HCD and BTI are both sensitive to high stress temperatures. Especially for hot-carrier degradation of short channel transistors, where the highest degradation occurs for equal drain- and gate-source voltages, the worst-case ansatz leads to a mixture of HCD and BTI (including recovery). In addition to that, the transistor rarely faces this worst-case condition in the circuit (see Fig. 2.22 and 2.23). Here as an example, the simulated behavior of an inverter is presented. The circuit design of the inverter is depicted in Fig. 2.24. In Fig. 2.22 one can nicely see, that the transistors never face the conditions of the worst case aging model (WCAM). But due to the HCD reliability assessment, which estimates the transistor lifetime based on the WCAM, the transistors will be implemented with a longer channel than needed for the application. This results in an additional costly reliability margin because the WCAM based circuit design will need more chip area. Anyhow, the WCAM approach is used for many circuit designs because of its simplicity and generally prevents a failure which is related to the assessed stress modes.

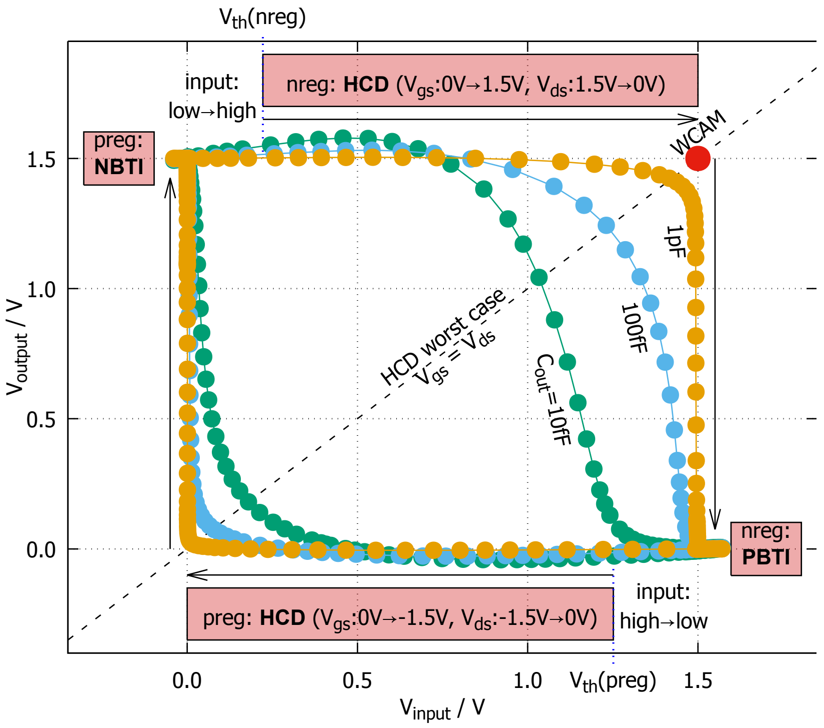
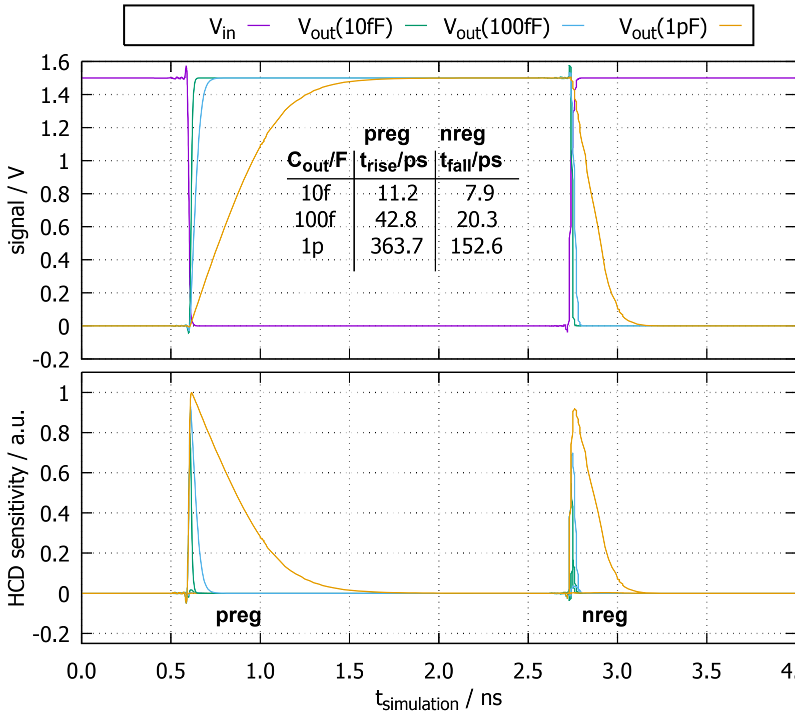


Figure 2.22.: The simulated terminal voltages of an inverter, consisting of the core n-MOSFET (nreg) and p-MOSFET (preg) examined in this work, show that the worst-case HCD condition $V_{gs} = V_{ds}$ is reached only for very high output loads (e.g. 1 pF which is equal to the load of about 32 inverter gates). Typical circuit design rules only allow much smaller loads without buffer structures (e.g. 6 gates). Therefore, the aging condition assumed by a worst-case aging model (WCAM) is never reached in the real application.



	preg	nreg
C_{out}/F	t_{rise}/ps	t_{fall}/ps
10f	11.2	7.9
100f	42.8	20.3
1p	363.7	152.6

Figure 2.23.: Simulated transient voltages of the inverter with different loads at the output. The HCD sensitivity ($\propto V_{gs}V_{ds}$) reveals that the single transistors only face little HCD stress during the transients between $|V_{ds}| = 0\,V$ and $1.5\,V$ compared to the whole simulated period of 4 ns. Both for nreg and preg the HCD related degradation increases with the output load. The given rise and fall times correspond to the slope between 20% and 80% of the output signal levels.

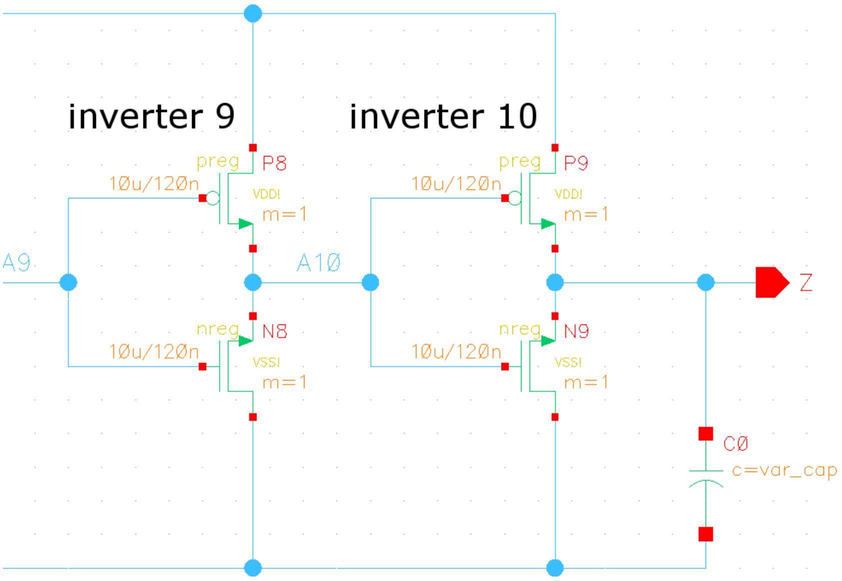
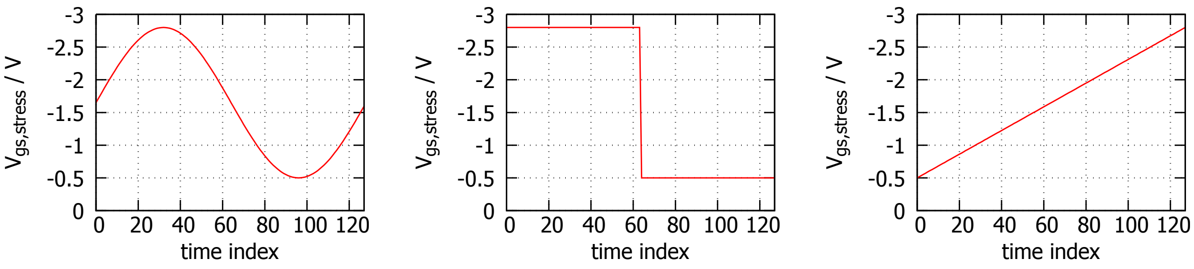


Figure 2.24.: End section of the design of a 10 stage inverter chain. The input of the chain (A1) is toggled between 0 and 1.5 V. In order to condition the input slope of the analyzed inverter nine preceding inverters are used. The inverter of interest, with input A10 and output Z, is located in the last stage and has a variable output capacitance (value: “var_cap”, e.g. 1 pF) to simulate the impact of various loads on the HCD stress the pull-up preg and pull-down nreg transistors face.

2.8. From DC to AC and arbitrary stress signals

In general most reliability assessments are based on drift data originated from worst-case direct current (DC) stress experiments (see Section 2.7). One way to build a bridge from the gathered reliability information to the real application case are SPICE simulations where the transistor is operated under arbitrary signals. This idea is presented in Section 5.1. A different possible approach is the application of alternating current (AC) or arbitrary stress to the device under test. AC stress is a special mode which can be chosen in the measurement recipe of the ultra-fast measurement equipment and does not need any additional module to be executed.

If one knows which signals are present in the real circuit at the transistor terminals these can be generated by an arbitrary waveform generator and applied to the tested transistor. To emphasize one aging mechanism the voltage levels may be elevated to a stress level (e.g. the drain-source voltage for hot-carrier degradation or the gate-source voltage for NBTI). An arbitrary waveform generator is built into the ultra-fast measurement equipment and enables this setup with an arbitrary signal at the gate of the device and a DC signal at its drain. Contrary to direct and alternating current gate bias stress a further module needs to be used in the software. This module pre-loads the information about the arbitrary stress signal from the experimental recipe sheet into the memory of the ultra-fast measurement equipment from which it will be interpreted and applied periodically to the gate of the transistor. Typical signal patterns for NBTI stress can be found in Fig. 2.25. The arbitrary signal consists of 128 time index points and can be applied with frequencies ranging from several Hertz up to about 4×10^3 Hz to the device under test. Measurement results of this approach and their comparison to aging simulations employing a compact aging model from [168] and a microscopic TCAD aging model (discussed in Section 1.2.3) are presented in Section 4.3.3.



Sine signal, amplitude: 2.3 V, offset −1.65 V. The minimum is set to −0.5 V ≈ V_{th0} of core p-MOSFET

Square-wave signal, V_h = −2.8 V, V_l = −0.5 V ≈ V_{th0} of core p-MOSFET

Saw-tooth signal, ramp ranging from −0.5 V ≈ V_{th0} to −2.8 V of core p-MOSFET

Figure 2.25.: Typical signal patterns which are applied periodically to the gate terminal of the transistor employing the arbitrary waveform generator of the ultra-fast measurement equipment. All signals are sampled with 128 time index points.

2.9. Circuits

For the purpose of this work schematics and layouts for a special ring oscillator circuits including wired-out transistors were created and realized in an advanced technology. The general design of those circuits is shown in Fig. 2.26. Each ring oscillator consists of 51 NAND gates. To switch between measurement and oscillation mode the input of one NAND gate is connected to a control pad ("enable PAD") while the enable signal of two transmission gates (T-Gates) is connected to the inverted signal. In the oscillation mode ("enable PAD" = high) the output of the oscillator is divided to a frequency below 1 MHz, decoupled from the ring with a buffer and can be probed at the "OSCI PAD" with an oscilloscope (an example signal is shown in Fig. 2.27). In the measurement mode ("enable PAD" = low) the ring is opened by the 1st NAND gate while the T-Gates conduct the input and output of the 26th NAND gate to two measurement pads ("M1/M2 PAD"). "M1 PAD" drives the gates of the n- and p-MOSFET of the NAND gate whereas "M2 PAD" is connected to the drain terminals. This (and the utilization of the VDD and VSS signal) allows a measurement of the characteristics of the two MOSFETs inside the probed NAND gate. The measurement mode can also be used to imprint a certain aging on the transistors of the 26th NAND gate which allows for instance the study of the NBTI recovery of the p-MOSFET during circuit operation.

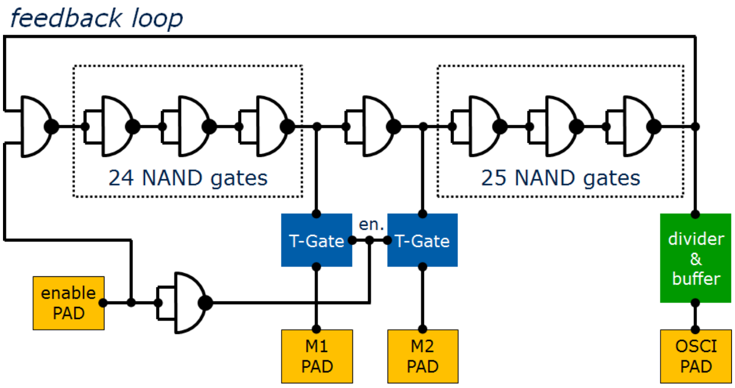


Figure 2.26.: The special design of a ring oscillator with wired-out transistors to measure degradation effects inside one NAND stage. The "enable PAD" is used to switch between oscillation and measurement mode. The output signal of the oscillator is measured indirectly after a divider and buffer block. The two pads "M1" and "M2" are connected to an SMU to measure the characteristics of the n- and p-MOSFET of the 26th NAND gate.

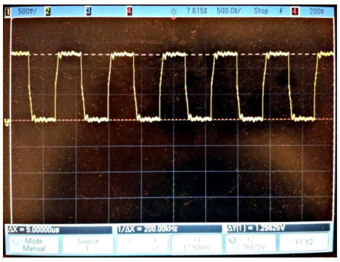


Figure 2.27.: Output signal of the ring oscillator in operation. The measured frequency of the unstressed circuit is about 125 kHz after the divider and buffer stage.

3. Measurements

In this chapter the setup of the performed measurements will be explained including the characterization of the unstressed devices, the applied stress voltage regimes and the thermal activation of each examined degradation mechanism.

3.1. Single transistors

Within the scope of this work n- and p-MOSFETs of two device classes with different geometries and gate oxide thicknesses (see Chapter 2.1 for further details) were characterized at room and elevated temperature in detail, the core ($t_{\mathrm{ox}} = 2.2\,\mathrm{nm}$) and analog ($t_{\mathrm{ox}} = 5.2\,\mathrm{nm}$) class transistors. All devices were analyzed for their HCD behavior using the standard measurement equipment presented in Section 2.3 whereas the NBTI degradation and recovery properties of the p-MOSFETs were measured with the ultra-fast measurement equipment (see Section 2.5). In addition to that the stress condition for "pure HCD" (*definition:* no recovery of the aged parameters is observable after stress) was determined with the ultra-fast measurement equipment including long-term recovery monitoring.

Because one of the main aspects of this work was to find voltage dependent degradation models for HCD, NBTI and their mixture, 2-dimensional stress voltage matrices as shown in Fig. 3.1 and Fig. 3.2 were defined. Each matrix contains the stress conditions for the gate-source and drain-source voltage for most of the performed measurements. The stress conditions of the core device include non-conducting and HCD stress conditions for the n-MOSFET, homogeneous, inhomogeneous NBTI, NCS, HCD and the mixture of NBTI and HCD for the p-MOSFET employing both the standard and ultra-fast measurement equipment. For the analog transistor HCD and the mixture of HCD and BTI were characterized using the standard equipment. In addition to that, measurements covering homogeneous NBTI were performed with the ultra-fast measurement equipment. To obtain information about the thermal activation of all stress modes each voltage condition was measured at 298 K and 398 K. Additional CET map data were recorded at 443 K with the ultra-fast measurement equipment. Low temperature measurements were not carried out due to the limitations of the installed wafer chuck.

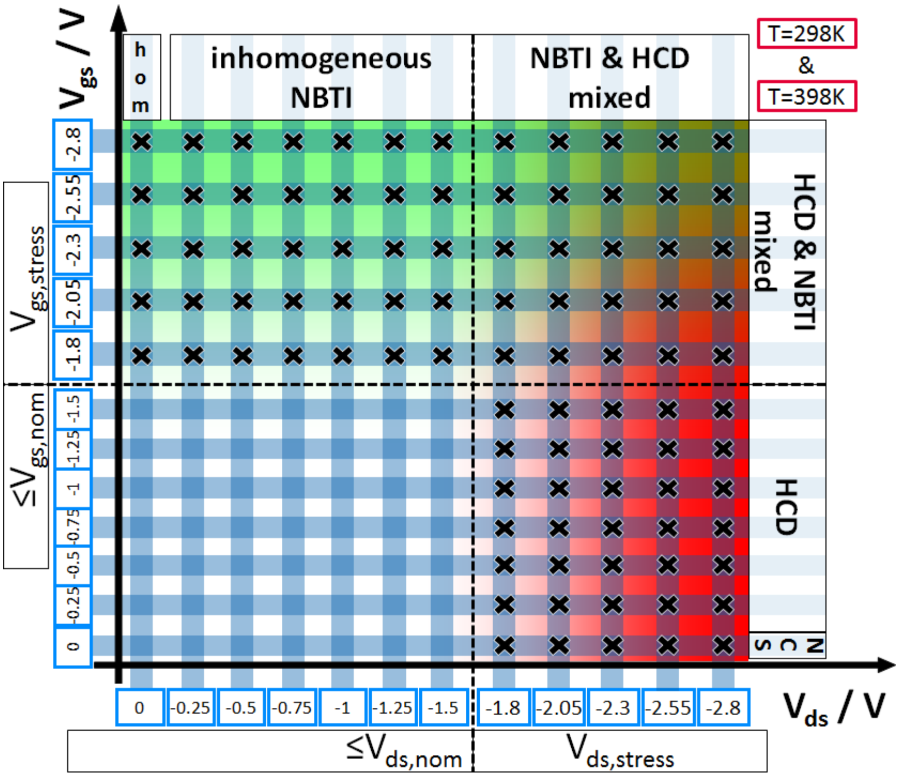


Figure 3.1.: The 2-dimensional stress voltage matrix of the core p-MOSFET defines the relevant terminal voltages for the examined degradation mechanisms. For $V_{\mathrm{gs}} = 0\,\mathrm{V}$ and $|V_{\mathrm{ds}}| \geq V_{\mathrm{nom}}$ non-conducting stress (NCS) is expected. The opposite condition, $V_{\mathrm{ds}} = 0\,\mathrm{V}$ and $|V_{\mathrm{gs}}| \geq V_{\mathrm{nom}}$, will cause homogeneous NBTI degradation, while increasing $|V_{\mathrm{ds}}|$ will lead to inhomogeneous NBTI stress. For $|V_{\mathrm{gs}}| \geq V_{\mathrm{th}}$ HCD will take place and for additional $|V_{\mathrm{gs}}| \geq V_{\mathrm{nom}}$ a mixture of HCD and NBTI is possible. The transition between nominal and stress voltages has to be considered as a formal definition, being aware that this only applies to lab stress times in the range of several kilo-seconds. All measurements were performed at room and elevated temperature. The stress conditions for the core n-MOSFET are at the lower right voltage regime of the matrix for NCS and HCD with inverted signs for V_{gs} and V_{ds}.

3.1.1. Standard equipment

The advantage of the standard equipment (described in Section 2.3) is is that it can be used to apply stress in parallel to different transistors. For the core and the analog device all geometries listed in Chapter 2.1 are located in the same test row and, therefore, stressed at once employing a switching matrix and a probe card with 25 needles. Of course, this setup can only be used if the degradation of the devices will not recover after stress because the serial measurement of the MOSFET parameters requires quite a long time (the last transistor is measured about 3 minutes after the first, see also Section 2.4). The programmed stress times are logarithmic equidistant, t_{s} =(0, 1, .. 10000)s, but are subject to uncertainties caused by the switching time of the installed matrix. To accurately account for this effect, the real stress times were also recorded.

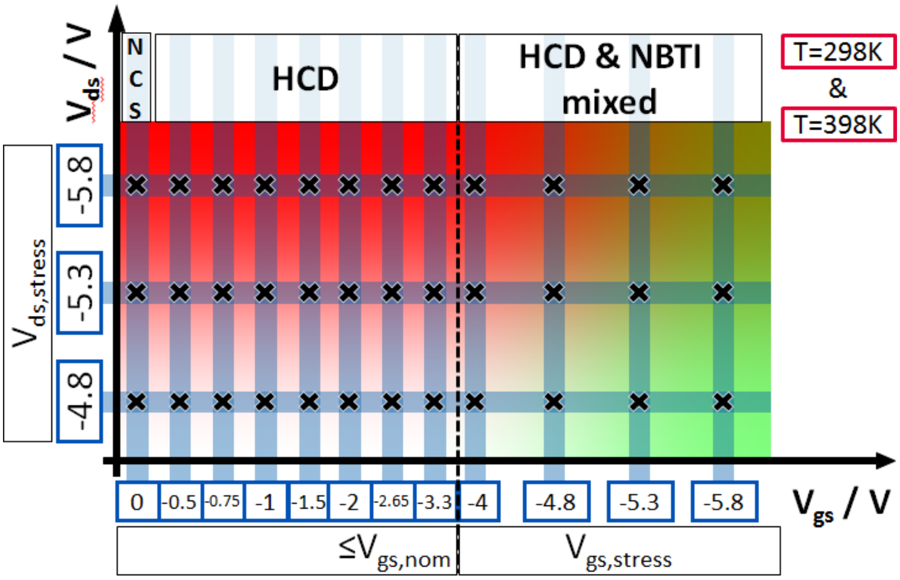


Figure 3.2.: The 2-dimensional stress voltage matrix of the analog p-MOSFET consists of 36 stress conditions at 298 K and 398 K. NCS is present at $V_{gs} = 0$ and increases with V_{ds}. HCD starts if $V_{gs} > 0$ and for $V_{gs} > V_{gs,nom}$ a mixture of HCD and NBTI is obtained. As with the core device, the transition from nominal to stress voltages has to be understood as a formal definition. The same stress voltages were applied to the n-MOSFET with inverted signs for V_{gs} and V_{ds}.

Before the first and after each stress time transistor parameters according to Table 3.1 were measured (measure-stress-measure sequence).

parameter	core device	analog device
threshold voltage	$V_{th}\,(\lvert V_{ds}\rvert = V_{nom}, I_{ds} = (70(\text{pMOS}), 300(\text{nMOS}))\,\text{nA} \times W/L)$	
operating point (forward)	$\lvert V_{ds}\rvert = \lvert V_{gs}\rvert = 0.8\,\text{V}$, $\lvert V_{ds}\rvert = \lvert V_{gs}\rvert = 1.2\,\text{V}$, $\lvert V_{ds}\rvert = \lvert V_{gs}\rvert = 1.5\,\text{V}$	$\lvert V_{ds}\rvert = \lvert V_{gs}\rvert = 1.8\,\text{V}$, $\lvert V_{ds}\rvert = \lvert V_{gs}\rvert = 2.5\,\text{V}$, $\lvert V_{ds}\rvert = \lvert V_{gs}\rvert = 3.3\,\text{V}$
operating point (reverse)	$\lvert V_{sd}\rvert = \lvert V_{gs}\rvert = 0.8\,\text{V}$, $\lvert V_{sd}\rvert = \lvert V_{gs}\rvert = 1.2\,\text{V}$, $\lvert V_{sd}\rvert = \lvert V_{gs}\rvert = 1.5\,\text{V}$	$\lvert V_{sd}\rvert = \lvert V_{gs}\rvert = 1.8\,\text{V}$, $\lvert V_{sd}\rvert = \lvert V_{gs}\rvert = 2.5\,\text{V}$, $\lvert V_{sd}\rvert = \lvert V_{gs}\rvert = 3.3\,\text{V}$
lin. input char.	$I_{ds}\,(\lvert V_{ds}\rvert = 0.1\,\text{V}, \lvert V_{gs}\rvert = 0 \ldots V_{nom})$	
saturation input characteristic	$I_{ds}\,(\lvert V_{ds}\rvert = (V_{nom}, V_{nom} - 0.2\,\text{V})$, $\lvert V_{gs}\rvert = 0 \ldots V_{nom})$	$I_{ds}\,(\lvert V_{ds}\rvert = (V_{nom}, V_{nom} - 0.4\,\text{V})$, $\lvert V_{gs}\rvert = 0 \ldots V_{nom})$
saturation output characteristic	$I_{ds}\,(\lvert V_{gs}\rvert = (V_{nom}, V_{nom} - 0.2\,\text{V})$, $\lvert V_{ds}\rvert = 0 \ldots V_{nom})$	$I_{ds}\,(\lvert V_{gs}\rvert = (V_{nom}, V_{nom} - 0.4\,\text{V})$, $\lvert V_{ds}\rvert = 0 \ldots V_{nom})$
leakage currents	$I_{ds}\,(\lvert V_{ds}\rvert = V_{nom}, V_{g0} = V_{b0} = 0)\,, I_{sd}\,(\lvert V_{ds}\rvert = V_{nom}, V_{g0} = V_{b0} = 0)$, $I_{gs}\,(\lvert V_{gs}\rvert = V_{nom}, V_{d0} = V_{b0} = 0),\ I_{gd}\,(\lvert V_{gs}\rvert = V_{nom}, V_{s0} = V_{b0} = 0)$	

Table 3.1.: List of the measured parameters of the analog and core class transistors.

3.1.2. Ultra-fast equipment

In order to record the recovery information starting at about 1 µs after stress, the ultra-fast measurement equipment described in Section 2.5 was employed for all stress modes with accelerated gate-source voltage. Contrary to the standard equipment a switching matrix cannot be used because the long connection wires have too high capacities and the matrix has too long switching times. Therefore, only single transistor measurements were performed with this instrument. Although several NBTI degradation measurements were carried out also for the analog device, the detailed NBTI analysis of the core p-MOSFET was in the focus of this work. Therefore, the following information is related to that device class.

To record a full CET map dataset the stress and recovery times as shown in Table B.1 were chosen. In general, a constant bias at a pre-defined value is used as a stress signal, e.g. $V_{\text{gs,stress}} = -2.8\,\text{V}$. In addition to that also periodic signals as sine, square-wave or sawtooth can be used to stress the gate of the transistor. To additionally activate HCD a constant drain-source stress voltage can be applied to the device under test. After each stress phase either the threshold voltage or an operating point is measured. To capture both parameter shifts with a single measurement the sequence as shown in Table B.2 can be used. Here, after each stress time the first data point of the threshold voltage is measured followed by an alternating sequence of the linear ($V_{\text{gs}} = -1.5\,\text{V}$, $V_{\text{ds}} = -0.1\,\text{V}$), analog ($V_{\text{gs}} = -0.8\,\text{V}$, $V_{\text{ds}} = -1.5\,\text{V}$) and saturation ($V_{\text{gs}} = -1.5\,\text{V}$, $V_{\text{ds}} = -1.5\,\text{V}$) operating point. This method allows an attachment of the threshold voltage recovery data to CET map data, where only V_{th} has been measured, featuring a minimum recovery of the operating point. Additionally, the recovery of the second and third operating point can be subtracted out by that and the real degradation after about 10 ms (time to set the ultra-fast measurement equipment to the SMU mode) calculated. The measurement delay is limited to the settling time of the instrument for each operating point ($t_{\text{set.,lin}} \approx 18\,\mu\text{s}$, $t_{\text{set.,ana}} \approx 13\,\mu\text{s}$, $t_{\text{set.,sat}} \approx 11\,\mu\text{s}$).

4. Experimental results

In this chapter the characterization of the fresh and degraded MOSFETs will be presented. The measurement data cover all important dimensions which may influence the operating point of the transistor including the source, drain, gate and bulk terminal voltages as well as the contact temperature. All of these parameters are accessible during SPICE simulations and can therefore be used for a direct comparison between the measured and simulated results.

4.1. Characterization of unstressed transistors

4.1.1. Analog device class

Due to the large minimal length ($L_{\text{min}} = 370\,\text{nm}$), the analog device is considered to age comparably to a long-channel transistor when subjected to hot-carrier stress. The substrate and gate current were measured in dependence of the gate-source voltage to find the most critical stress conditions at room and elevated temperature. This applies for an accelerated drain-source voltage and the gate-source voltage for which the substrate (n-MOSFET) or gate current (p-MOSFET) has its maximum. In order to characterize the entire current flow of the analog p-MOSFET also the substrate current has been recorded in dependence of the gate-source voltage. The results of the pre-characterization are shown in Figs. 4.1 to 4.3 for the analog n- and p-MOSFET at room and elevated temperature.

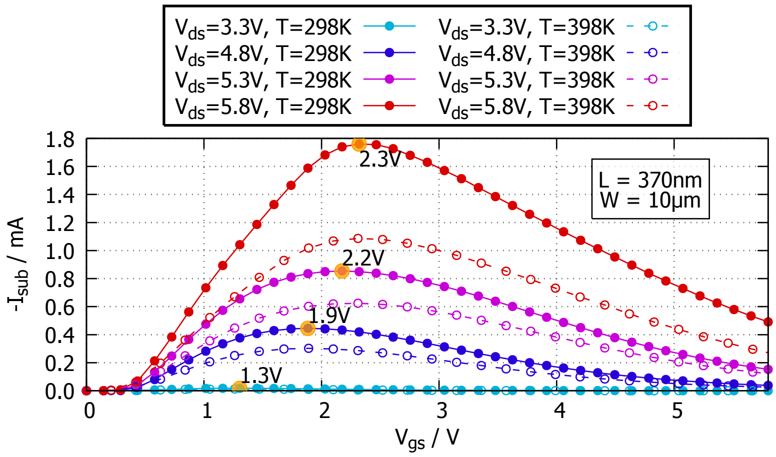


Figure 4.1.: The position of the substrate current maximum (orange highlight) of the analog n-MOSFET shows a strong dependence on the chosen drain-source voltage. The substrate current is lower at elevated temperature (dashed line) for each drain-source voltage (plotted with reduced data density).

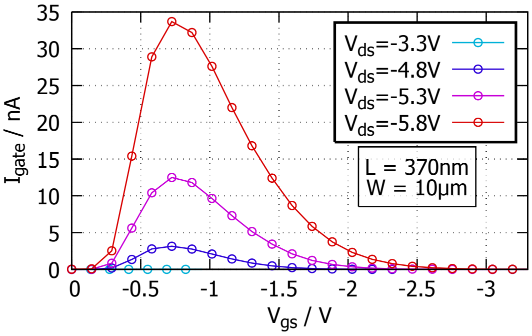


Figure 4.2.: The gate current maximum of the analog p-MOSFET is located at $V_{gs} = -0.75\,\mathrm{V}$. Contrary to the substrate current maximum (Fig. 4.3) the position of the gate current maximum is independent of the chosen drain-source voltage.

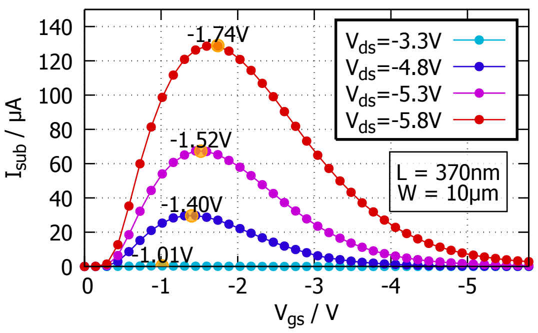


Figure 4.3.: The position of the substrate current maximum (orange highlight) of the analog p-MOSFET shows a strong dependence on the chosen drain-source voltage.

For the n- and p-MOSFET an absolute gate-source stress voltage up to about 6.5 V and absolute drain-source stress voltage below 5.8 V are reasonable because the characteristics (Fig. 4.4 and 4.5) do not show any avalanche within that range. Both transistors show a 0.2 V...0.3 V earlier drain-source voltage caused breakdown at 398 K in contrast to the room temperature data. The current after breakdown is limited by the SMU to 0.1 A. The p-MOSFET can be stressed with an about 3.1 V higher drain-source voltage than the n-MOSFET. Because the pre-characterization test is very short, compared to a typical stress time for the aging characterization, the limits for the degradation test voltages are expected to be lower (especially the gate-source voltage).

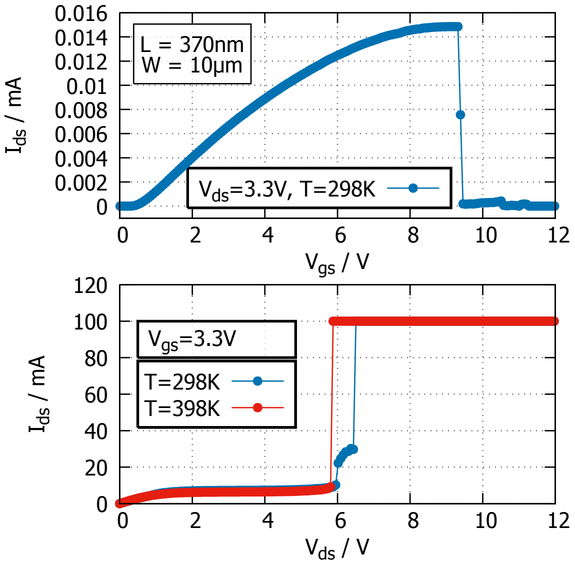


Figure 4.4.: Determination of the gate- (**top**) and drain-source (**bottom**) breakdown voltages of the analog n-MOSFET. The breakdown occurs at $V_{gs} > 9.4\,\mathrm{V}$ or $V_{ds} > 6.0\,\mathrm{V}$ ($T = 298\,\mathrm{K}$) resp. $V_{ds} > 5.8\,\mathrm{V}$ ($T = 398\,\mathrm{K}$).

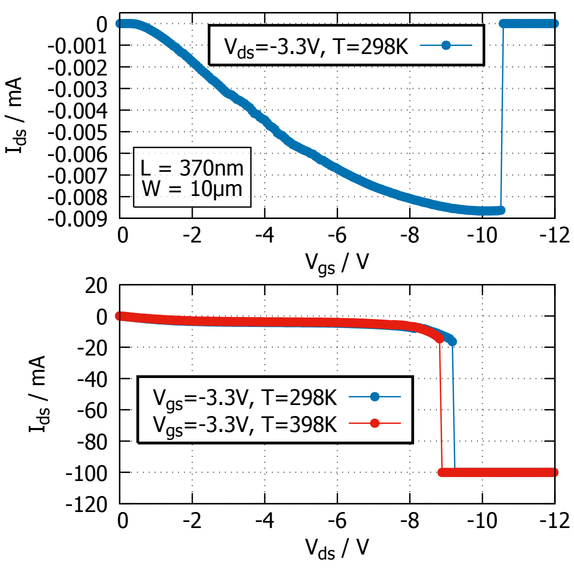


Figure 4.5.: Determination of the gate- (**top**) and drain-source (**bottom**) breakdown voltages of the analog p-MOSFET. The channel breaks down at $V_{gs} < -10.6\,\mathrm{V}$ or $V_{ds} < -9.2\,\mathrm{V}$ ($T = 298\,\mathrm{K}$) resp. $V_{ds} < -8.9\,\mathrm{V}$ ($T = 398\,\mathrm{K}$).

4.1.2. Core device class

Compared to the analog device the current-voltage characteristics and the worst case conditions are quite different for the core device because of its rather short channel ($L_{\text{min}} = 100\,\text{nm}$). The degradation is expected to be most severe for a large density of highly energetic carriers, which is the case for $V_{\text{gs}} = V_{\text{ds}}$, the well-known worst-case condition for short channel transistors. The corresponding measurements of the drain-source current of the core devices are shown in Fig. 4.6. The data show an increase of the drain-source current with increasing gate-source voltage, which indicates that the highest flux is present for the maximum absolute gate-source voltage.

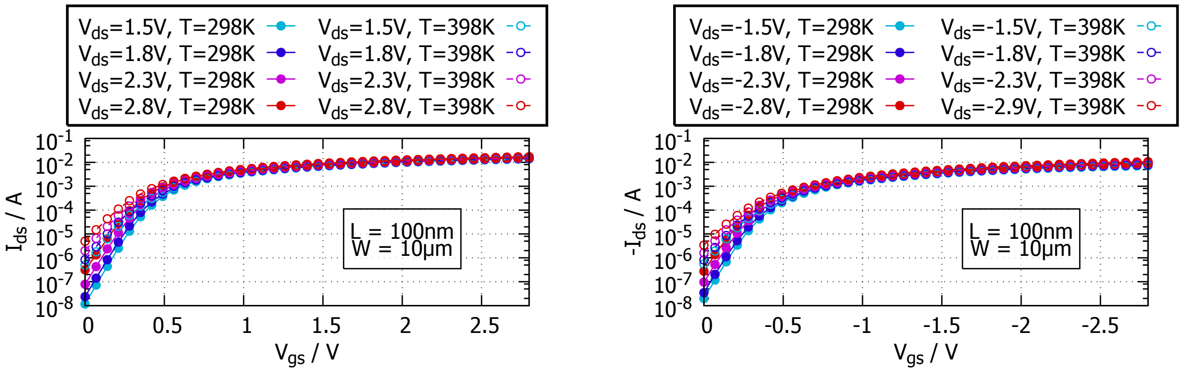


Figure 4.6.: The drain-source current of the core n- (**left**) and p-MOSFET (**right**) constantly increases with the gate-source voltage. Therefore the maximum carrier flux is present for the highest absolute gate-source voltage.

For the sake of completeness, the substrate and gate current of the core n- and p-MOSFET are depicted in Fig. 4.7 and Fig. 4.8. The substrate current of the n-MOSFET increases up to a maximum value which is dependent on the applied drain-source voltage and decreases after that. Here, with increasing gate-source voltage holes are injected into the substrate whereby the substrate current is decreasing. In addition to that, the absolute value of the room temperature data is always smaller compared to the high temperature data. The gate current of the p-MOSFET features two maxima. The first one is located at $V_{\text{gs}} = 0.35\,\text{V} \approx V_{\text{th}}$ and has a positive sign (electrons are injected into the oxide) whereas the second one is located at the maximum gate-source voltage and has a negative sign indicating a strong current of holes towards the channel-oxide interface.

The identified gate- and drain-source breakdown voltages limit the stress regime to $V_{\text{gs}} < 5.1\,\text{V}$ and $V_{\text{ds}} < 4.1\,\text{V}$. It has to be noted that the critical gate-source voltage is expected below the measured gate-source breakdown voltage because a typical stress time is several orders of magnitude longer compared to the applied signal of the pre-characterization test. Both devices show an earlier drain-source breakdown at room temperature. The p-MOSFET withstands an about 1.1 V higher drain-source voltage.

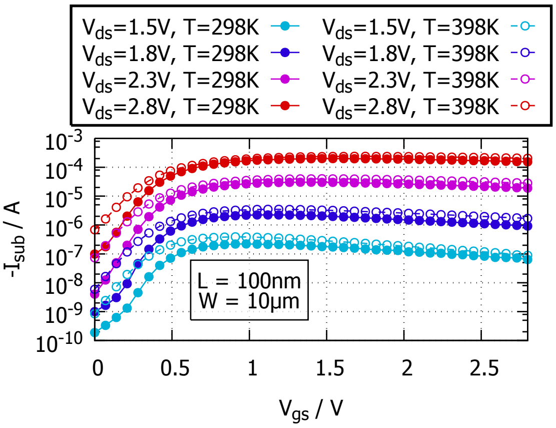


Figure 4.7.: The substrate current of the core n-MOSFET first increases with the gate-source voltage. After the maximum value, which is dependent on the chosen drain-source voltage, the injection of holes into the substrate leads to a decrease (up to a factor of 4.5 for $V_{\mathrm{ds}} = 1.5\,\mathrm{V}$ at 398 K) of the measured substrate current.

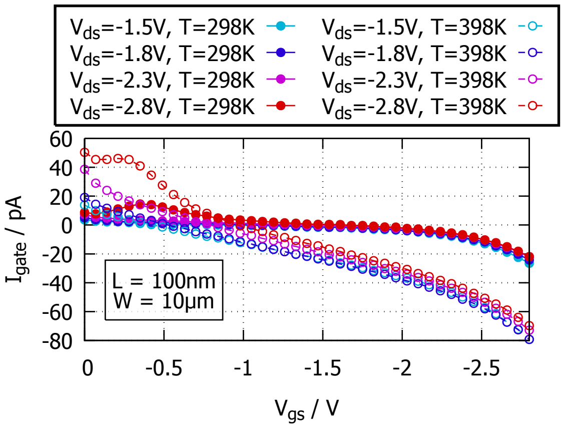


Figure 4.8.: Two maxima can be identified in the gate current data of the core p-MOSFET. The location of the first one (which is caused by electrons being injected into the oxide) is temperature dependent and most pronounced at $V_{\mathrm{gs}} = -0.35\,\mathrm{V}$ (298 K), resp. $V_{\mathrm{gs}} = -0.28\,\mathrm{V}$ (398 K). A second maximum is caused by the gate-source voltage dependent hole current towards the interface. The absolute values of both extrema strongly increase with the temperature.

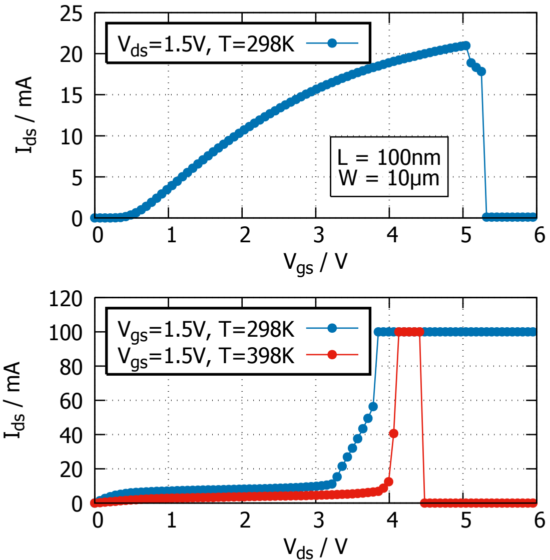


Figure 4.9.: Determination of the gate- (top) and drain-source (bottom) breakdown voltages of the core n-MOSFET. The channel breaks down at $V_{\mathrm{gs}} > 5.1\,\mathrm{V}$ or $V_{\mathrm{ds}} > 3.8\,\mathrm{V}$ ($T = 298\,\mathrm{K}$) resp. $V_{\mathrm{ds}} > 4.1\,\mathrm{V}$ ($T = 398\,\mathrm{K}$).

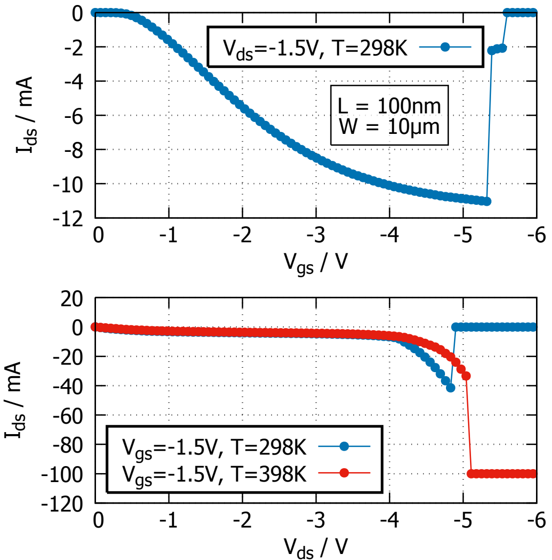


Figure 4.10.: Determination of the gate- (top) and drain-source (bottom) breakdown voltages of the core p-MOSFET. The breakdown occurs at $V_{\mathrm{gs}} < -5.4\,\mathrm{V}$ or $V_{\mathrm{ds}} < -4.9\,\mathrm{V}$ ($T = 298\,\mathrm{K}$) resp. $V_{\mathrm{ds}} < -5.1\,\mathrm{V}$ ($T = 398\,\mathrm{K}$).

4.2. Degradation measurements with standard equipment

For a better illustration only a subset of the recorded data was chosen for the result presentation. Concurrently high emphasis was put on the description of every observable aging effect to give a full picture of all relevant mechanisms.

4.2.1. Analog device class

If stress voltages (e.g. $|V_{\mathrm{ds,stress}}| = 5.8\,\mathrm{V} \gg V_{\mathrm{nom}} = 3.3\,\mathrm{V}$) are applied to the transistors the findings about the most critical gate voltage for the unstressed devices which were presented in Section 4.1 need to be adapted due to the field dependence of the accelerated carriers. Therefore, the maximum degradation of e.g. the threshold voltage or the drain-source current is expected at the well-known condition $V_{\mathrm{gs,stress}} \approx V_{\mathrm{ds,stress}}/2$ for the analog n-MOSFET. Because the gate current maximum of the p-MOSFET is always located at $V_{\mathrm{gs}} = -0.75\,\mathrm{V}$ (independent of V_{ds}, see Fig. 4.2) HCD is expected to be most pronounced at this gate-source voltage. The threshold voltage drift results for HCD (BTI activation for $V_{\mathrm{gs}} \geq V_{\mathrm{nom}}$) are shown in Fig. 4.11 to Fig. 4.14 for the analog n- and p-MOSFET stressed at room and elevated temperature. In addition to that, the drift of the forward and reverse drain-source current has been recorded for all stress conditions and is shown in Fig. 4.15 for one selected stress condition of the analog n-MOSFET. Here, one can see that the drift of the operating point at $V_{\mathrm{ds}} = V_{\mathrm{gs}} = 1.8\,\mathrm{V}$ has the strongest degradation, which is about 64%.

A comparison of the data at the worst-case conditions at two different stress temperatures (Fig. 4.16 and Fig. 4.17) shows that the HCD related drift of the analog device is always larger at 298 K than at 398 K. The maximal shifts of the threshold voltage after 10^4 s stress at room temperature of the n-MOSFET are located at $V_{\mathrm{ds}} = 4.8\,\mathrm{V}$, $V_{\mathrm{gs}} = 2\,\mathrm{V}$ and $V_{\mathrm{ds}} = 5.3\,\mathrm{V}$, $V_{\mathrm{gs}} = 2.65\,\mathrm{V}$. Although the analog n-MOSFET broke during stress at $V_{\mathrm{ds}} = 5.8\,\mathrm{V}$ for the applied most critical condition the detailed data of the measurements with different gate-source voltages reveal that the destruction of the device occurred a little earlier for $V_{\mathrm{gs}} = 2.65\,\mathrm{V}$ than for $V_{\mathrm{gs}} = 3.3\,\mathrm{V}$. Therefore the worst-case is situated between those two values. In summary, the analog n-MOSFET confirms the worst-case condition for long channel transistors having the most severe degradation at a low temperature and $V_{\mathrm{gs,worst\text{-}case}} \approx V_{\mathrm{ds}}/2$ (see Fig. 4.18). Additionally, the n-MOSFET shows a noticeable PBTI degradation for high gate-source voltages at room and elevated temperature.

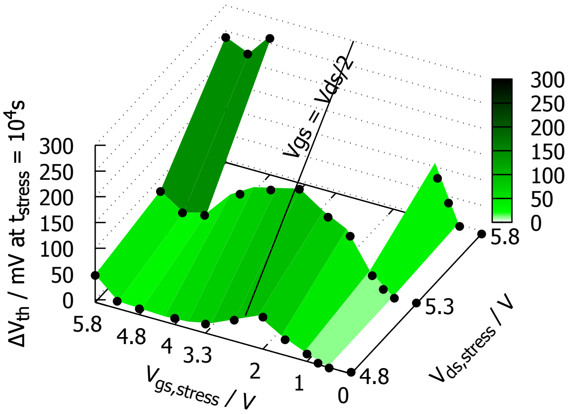


Figure 4.11.: Drift map of the analog n-MOSFET with $L = 370\,\text{nm}$, $W = 10\,\mu\text{m}$ stressed at 298 K. Due to the strong drift between $V_{\text{ds}} = 1\,\text{V}...4.8\,\text{V}$ the device broke for $V_{\text{ds}} = 5.8\,\text{V}$ leading to a gap in the map surface. As expected from the pre-characterization measurements the worst-case condition is located at $V_{\text{gs,stress}} \approx V_{\text{ds,stress}}/2$ (black line). The last data stripe at $V_{\text{gs,stress}} = 5.8\,\text{V}$ indicates a little PBTI activation.
Symbols: data points, surface: interpolated

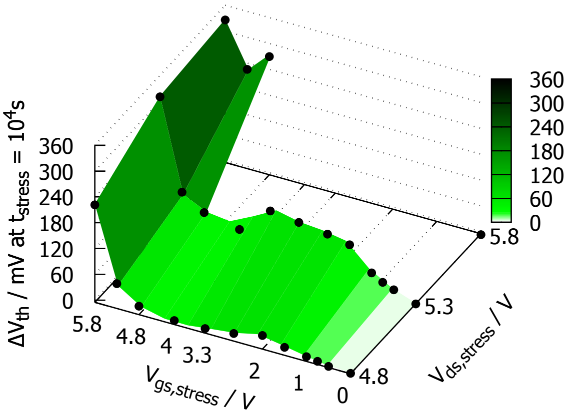


Figure 4.12.: Drift map of the analog n-MOSFET with $L = 370\,\text{nm}$, $W = 10\,\mu\text{m}$ stressed at 398 K. The device broke for several stress conditions at $V_{\text{ds}} = 5.8\,\text{V}$. The thermal activation of PBTI leads to a huge drift for $V_{\text{gs}} \geq 5.3\,\text{V}$. It should be noted that all data with $V_{\text{gs}} \geq V_{\text{nom}} = 3.3\,\text{V}$ potentially include BTI stress and therefore a recovery of several milliseconds before the value was recorded has to be taken into account.
Symbols: data points, surface: interpolated

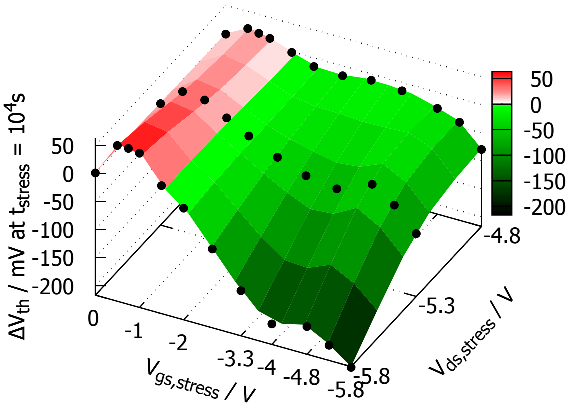


Figure 4.13.: Drift map of the analog p-MOSFET with $L = 370\,\text{nm}$, $W = 10\,\mu\text{m}$ stressed at 298 K. Two aging types are observable, one leading to an improvement of the device performance (red, $D_{\text{Igate,max}}$) and the other one causing a strong degradation of the threshold voltage (green, $D_{\psi,\text{max}}$).
Symbols: data points, surface: interpolated

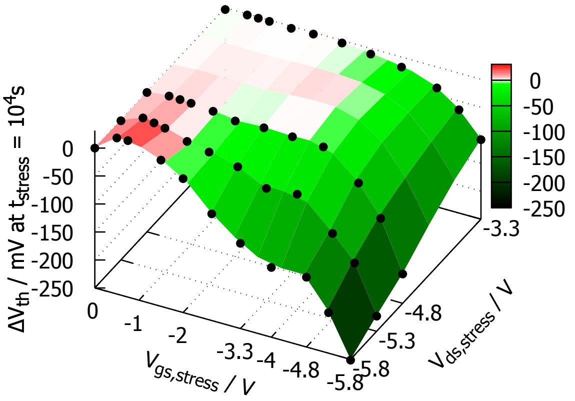


Figure 4.14.: Drift map of the analog p-MOSFET with $L = 370\,\text{nm}$, $W = 10\,\mu\text{m}$ stressed at 398 K. The same two aging types as found in the room temperature data (Fig. 4.13) can be identified but show a smaller parameter drift. An additional dataset has been recorded with $V_{\text{ds}} = -3.3\,\text{V}$ to quantify the NBTI activation which is about $-100\,\text{mV}$ at $V_{\text{gs}} = -5.8\,\text{V}$.
Symbols: data points, surface: interpolated

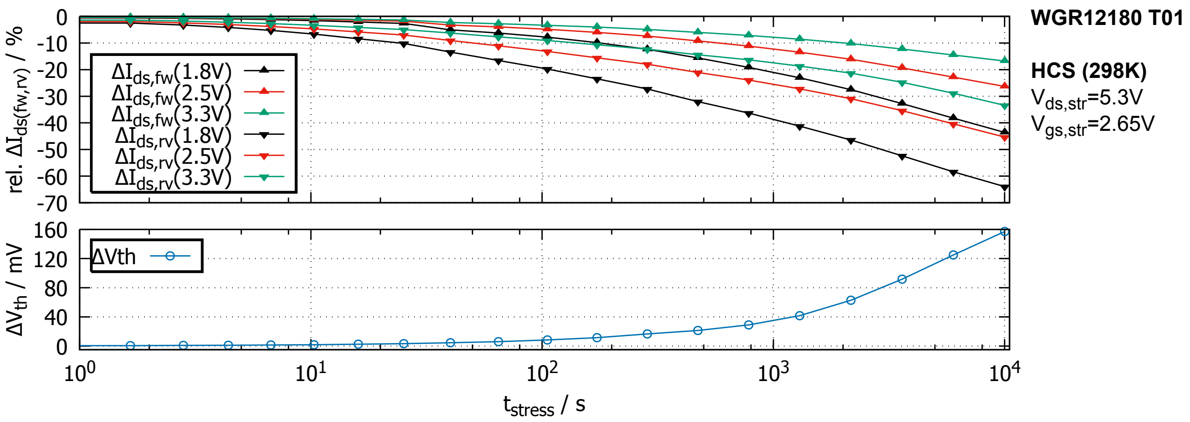


Figure 4.15.: Drift details of the analog n-MOSFET with $L = 370\,\text{nm}$, $W = 10\,\mu\text{m}$ stressed at 298 K and worst-case hot-carrier voltage conditions. The maximum current degradation is $\Delta I_{\text{ds,rv}}\ (V_{\text{ds}} = V_{\text{gs}} = 1.8\,\text{V}) \approx 64\%$ and the final threshold voltage drift is about 157 mV. Lines serve as guides to the eye.

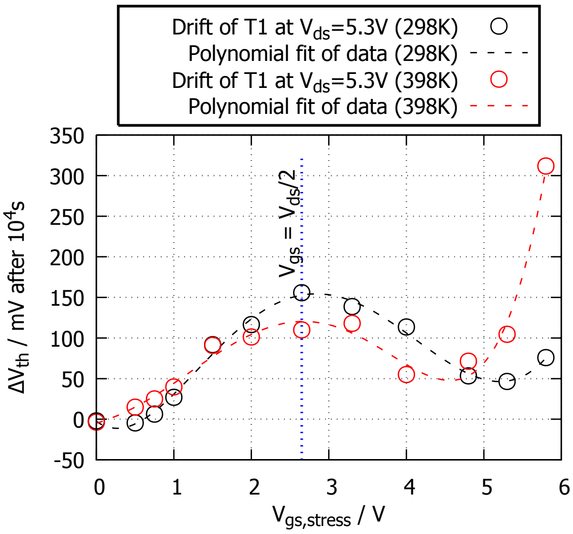


Figure 4.16.: The analog n-MOSFET with $L = 370\,\text{nm}$, $W = 10\,\mu\text{m}$ shows a higher drift ($\Delta V_{\text{th,298K}} = 157\,\text{mV}$) at 298 K compared to the 398 K data ($\Delta V_{\text{th,398K}} = 110\,\text{mV}$) at the worst-case hot-carrier voltage condition $V_{\text{gs,stress}} = V_{\text{ds,stress}}/2$. This is in accordance to e.g. the Hess model for degradation caused by solitary hot electrons (see also page 10). For $V_{\text{gs,stress}} = 5.8\,\text{V}$ the activation of PBTI is noticeable which leads to a very high threshold voltage drift at elevated temperature. Lines serve as guides to the eye and are based on polynomial fits of the drift data.

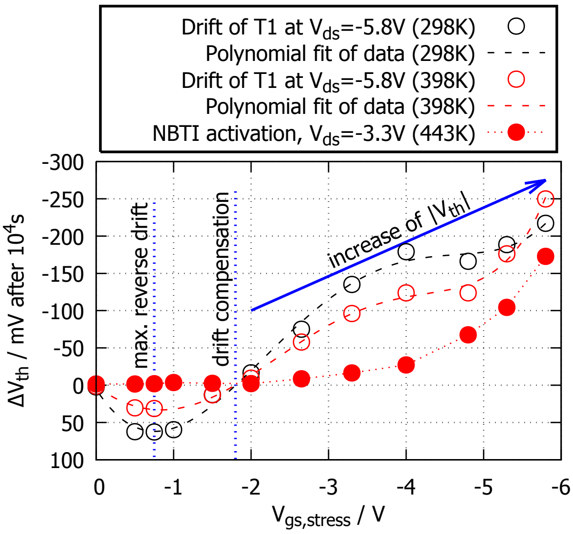


Figure 4.17.: The analog p-MOSFET with $L = 370\,\text{nm}$, $W = 10\,\mu\text{m}$ degrades stronger at 298 K than at 398 K. Two different degradation mechanisms are observable. For $V_{\text{gs,stress}} = -0.75\,\text{V}$ the performance of the transistor increases. At $V_{\text{gs,stress}} \approx -1.75\,\text{V}$ the two degradation mechanism compensate the threshold voltage drift after $10^4\,\text{s}$. Any higher $|V_{\text{gs,stress}}|$ leads to an increase of the absolute threshold voltage. At $V_{\text{gs,stress}} = -4.8\,\text{V}$ a local minimum of $|\Delta V_{\text{th}}|$ is present, which may be caused by the interplay of HCD and NBTI. A strong NBTI activation ($t_{\text{recovery}} \approx 100\,\text{ms}$) is taking place at $V_{\text{gs,stress}} = -5.8\,\text{V}$ and $T = 398\,\text{K}$ which can also be seen from the 443 K data at $|V_{\text{ds}}| = V_{\text{nom}} = 3.3\,\text{V}$.

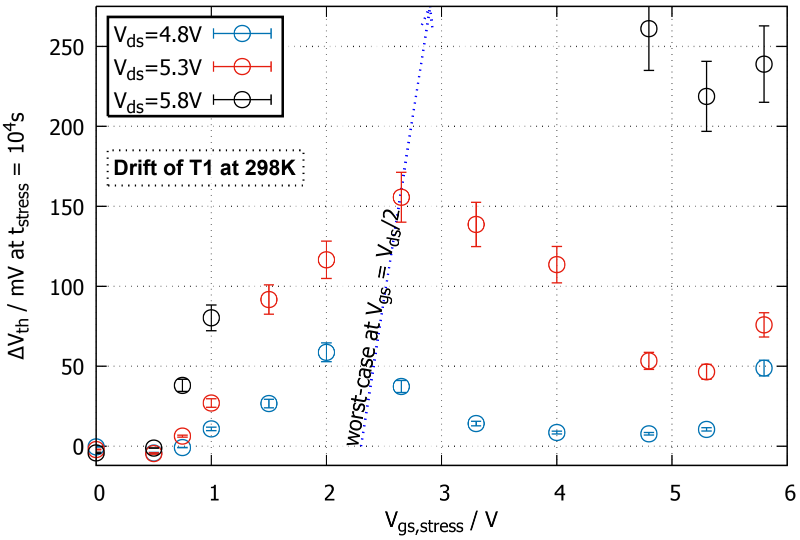


Figure 4.18.: The degradation behavior of the n-MOSFET follows the worst-case condition $V_{\mathrm{gs}} \approx V_{\mathrm{ds}}/2$ for long channel devices. The error bars of the room temperature data denote a typical device to device variation of 10%.

Contrary to the analog n-MOSFET, the p-MOSFET features two different HCD mechanisms, the first one, $D_{\mathrm{Igate,max}}$, leading to a positive drift of the threshold voltage ($0 \geq V_{\mathrm{gs}} \geq -1.75\,\mathrm{V}$) and the second one, $D_{\psi,\mathrm{max}}$, resulting in a negative drift ($-1.75\,\mathrm{V} \geq V_{\mathrm{gs}} \geq -3.3\,\mathrm{V}$). In-between those two degradation types the threshold voltage shift vanishes after about $10^4\,\mathrm{s}$ stress for $V_{\mathrm{gs}} \approx -1.75\,\mathrm{V}$. Elevated temperatures attenuate the drift type $D_{\mathrm{Igate,max}}$ and aggravate $D_{\psi,\mathrm{max}}$ to a 1:3 drift ratio (see Tab. 4.1 and Fig. 4.22). The observed drift compensation after $10^4\,\mathrm{s}$ is not influenced by the stress temperature and located at $V_{\mathrm{gs}} \approx -1.75\,\mathrm{V}$.

For $D_{\mathrm{Igate,max}}$ with the maximum at $V_{\mathrm{ds,stress}} = -5.8\,\mathrm{V}$ and $V_{\mathrm{gs,stress}} = -0.75\,\mathrm{V}$ the substrate current characteristics show a decrease of $\max(I_{\mathrm{bs}})$ with stress time. The fresh device has a substrate current maximum located at $-0.95\,\mathrm{V}$ whereas for the aged device the maximum is at $-0.85\,\mathrm{V}$ resulting in a shift of $100\,\mathrm{mV}$ (Fig. 4.19).

The maximum of $D_{\Psi,\mathrm{max}}$ is observed at $V_{\mathrm{ds,stress}} = -5.8\,\mathrm{V}$ and $V_{\mathrm{gs,stress}} = -3.3\,\mathrm{V}$ (the negative-bias temperature instability is negligible because of the low temperature and $|V_{\mathrm{gs,stress}}| = 3.3\,\mathrm{V} = V_{\mathrm{nominal}}$) and shows a completely different behavior (Fig. 4.20). Initially the maximum substrate current increases with stress time from the fresh device with a maximum located at $-0.95\,\mathrm{V}$ up to a turning point at $\approx 3.6 \times 10^3\,\mathrm{s}$ for $-1.2\,\mathrm{V}$ and ends after $10^4\,\mathrm{s}$ at $-1.4\,\mathrm{V}$. The total shift of the $V_{\mathrm{gs}}\,(\max(I_{\mathrm{bs}}))$ is $\approx -0.25\,\mathrm{V}$ (maximum of $D_{\Psi,\mathrm{max}}$) and $\approx -0.45\,\mathrm{V}$ after $10^4\,\mathrm{s}$ stress time.

The drift compensation ($\Delta V_{\mathrm{th}} = 0$ after $10^4\,\mathrm{s}$) due to both effects is best observable for $V_{\mathrm{ds,stress}} = -5.8\,\mathrm{V}$ and $V_{\mathrm{gs,stress}} = -1.5\,\mathrm{V}$. Also, the turning point at $1.3 \times 10^3\,\mathrm{s}$ is well correlated with the change of V_{th} (Fig. 4.21).

A third mechanism appears after the maximum of $D_{\Psi,\mathrm{max}}$ (Fig. 4.20) where $\max(I_{\mathrm{bs}})$ decreases and $V_{\mathrm{gs}}\,(\max(I_{\mathrm{bs}}))$ is shifted to a higher potential.

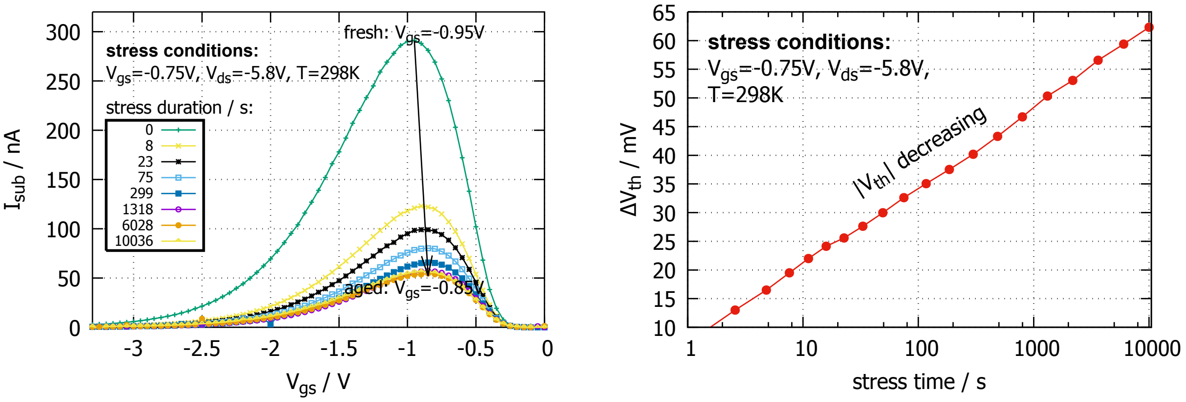


Figure 4.19.: Change of $I_{\mathrm{bs}}\,(V_{\mathrm{gs}})$ and V_{th} during aging at room temperature. **Left:** The drift type $D_{\mathrm{Igate,max}}$ is most pronounced at $V_{\mathrm{gs}} = -0.75\,\mathrm{V}$ and $V_{\mathrm{ds}} = -5.8\,\mathrm{V}$ after $10^4\,\mathrm{s}$ and shows a constant decrease of $\max(I_{\mathrm{bs}})$. **Right:** $|V_{\mathrm{th}}|$ decreases with stress time.

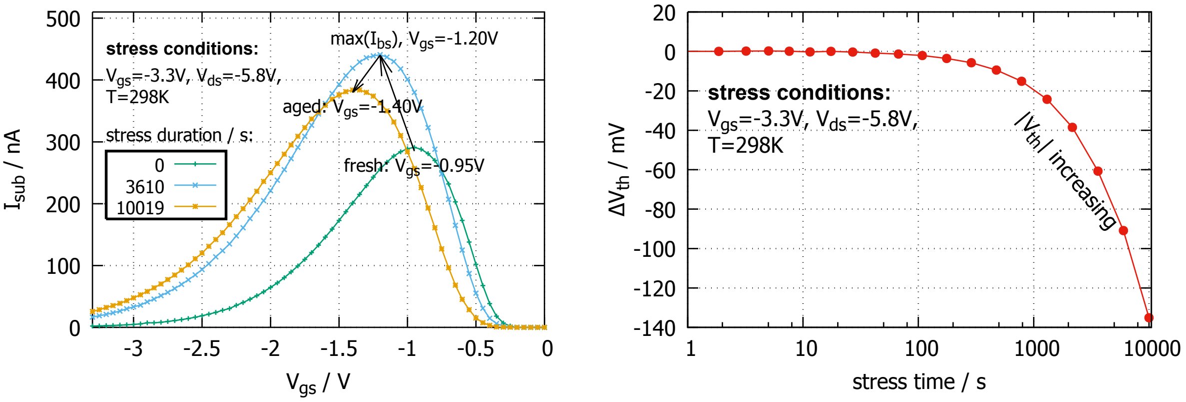


Figure 4.20.: Change of $I_{\mathrm{bs}}\,(V_{\mathrm{gs}})$ and V_{th} during aging at room temperature. **Left:** The drift type $D_{\Psi,\mathrm{max}}$ has its maximum after $3.6 \times 10^3\,\mathrm{s}$ and results in a constant increase of $\max(I_{\mathrm{bs}})$. An additional effect is noticeable for larger stress times decreasing $\max(I_{\mathrm{bs}})$. **Right:** $|V_{\mathrm{th}}|$ increases with stress time.

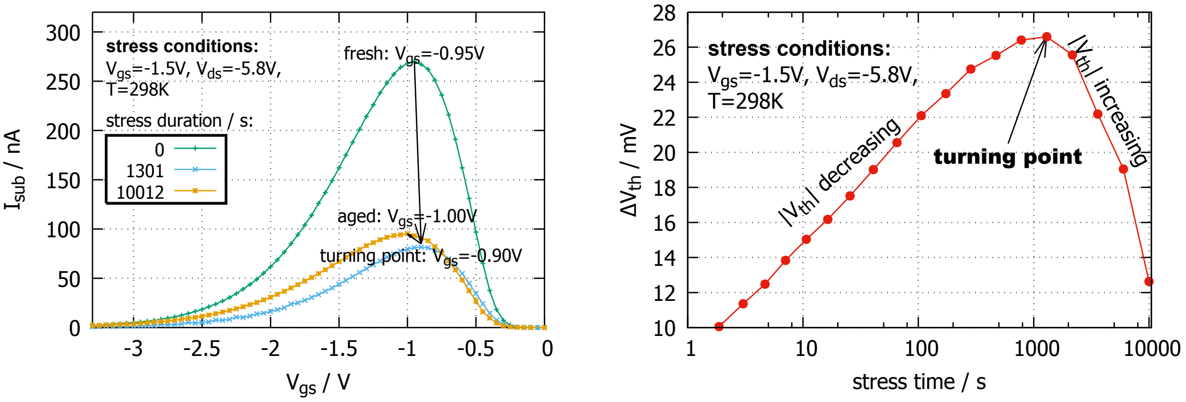


Figure 4.21.: The change of $I_{\mathrm{bs}}\,(V_{\mathrm{gs}})$ during aging shows a direct relation to the change of V_{th}, the common turning point is at $1.3 \times 10^3\,\mathrm{s}$ where $\Delta V_{\mathrm{th}} = \Delta\max(I_{\mathrm{bs}}) \approx 0$. **Left:** Measured combination of drift types $D_{\mathrm{Igate,max}}$ and $D_{\Psi,\mathrm{max}}$. **Right:** First drift type $D_{\mathrm{Igate,max}}$ decreases $|V_{\mathrm{th}}|$ with the stress time until a turning point after which $D_{\Psi,\mathrm{max}}$ dominates the aging process leading to an increase of $|V_{\mathrm{th}}|$.

A comparison of the drift data from transistors with different geometries indicates that the compensation is not influenced by the geometry and always located at $V_{\mathrm{gs}} \approx -1.75\,\mathrm{V}$. The drift ratio ($D_{\mathrm{Igate,max}}$:$D_{\Psi,\mathrm{max}}$) data (Tab. 4.1) show several groups of devices. On the one hand, there are short and wide channel transistors (1, 3, 4) with comparable drift ratio values. On the other hand, there are long and narrow devices (6, 7, 8) together with long and wide devices (2, 5) which have a very small drift in general for the applied stress time of 10^4 s. Only higher temperatures show a recognizable drift of the threshold voltage for that group. Both the room and elevated temperature data show an increase of the drift ratio up to 0.5 µm while the length is increasing. The 2 µm long device has the smallest drift ratio.

Analyzing the long term stress data at 3.6×10^5 s for $V_{\mathrm{gs}} = -0.75\,\mathrm{V}$ a compensation of the degradation around 7.7×10^4 s is observable (Fig. 4.23). This means that $D_{\mathrm{Igate,max}}$ will always be compensated by $D_{\Psi,\mathrm{max}}$ if a sufficient stress time was chosen.

Transistor	L/µm	W/µm	ratio, 298 K	ratio, 398 K
1	0.37	10	1:2.2	1:3.1
2	2	10	1:2.6	1:7.8
3	0.43	10	1:1.9	1:3.0
4	0.5	10	1:1.3	1:2.6
5	30	30	n.d.	1:12.1
6	1	0.5	n.d.	1:2.6
7	2	0.5	n.d.	1:7.4
8	2	0.4	n.d.	1:3.2

Table 4.1.: Drift ratio ($D_{\mathrm{Igate,max}}$:$D_{\Psi,\mathrm{max}}$) of several transistor geometries at room and elevated temperatures. n.d.: drift not detectable

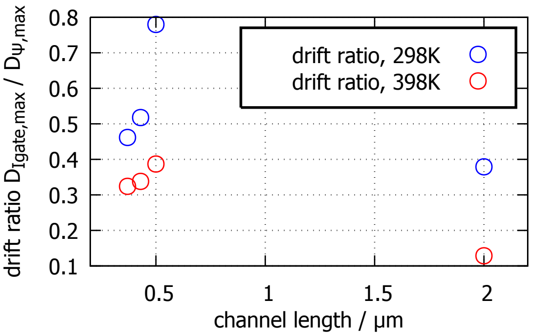


Figure 4.22.: The drift ratio increases with the channel length up to 0.5 µm. Between 0.5 µm and 2 µm a strong decrease is observable. A maximum of the ratio $D_{\mathrm{Igate,max}}$: $D_{\Psi,\mathrm{max}}$ is expected in-between but could not be determined experimentally due to the lack of other transistor geometries.

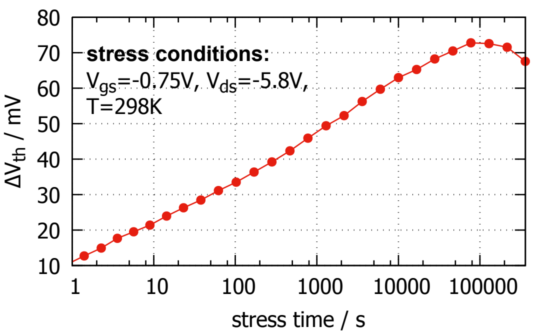


Figure 4.23.: Long term stress data show that even for low $V_{\mathrm{gs}} = -0.75\,\mathrm{V}$ the drift type $D_{\mathrm{Igate,max}}$ will be compensated by $D_{\Psi,\mathrm{max}}$. The turning point is located at 7.7×10^4 s.

Discussion of experimental results

In the literature several mechanisms for hot-carrier degradation have been reported. Single "hot" carriers which can gain high enough energies through field acceleration to break a Si-H bond [15, 169, 170], "hot" fractions of particle ensembles like for instance caused by impact ionization [171], Auger recombination [172], electron-phonon [173, 174] and electron-electron scattering [175, 176]. In addition, the superposition of single and multiple carrier effects has been discussed [69, 70].

Following the degration models of K. Hess et al. (see Section 1.1.2) and A. Bravaix (see Section 1.1.5) the first mechanism is assumed to be due to bond dissociation triggered by a solitary hot carrier. This process is typical for (relatively) high-voltage devices. In this case an indicator for the severity of hot-carrier degradation is the average (over the ensemble) carrier energy, and hence, as a related quantity, the gate current, is widely used for hot-carrier degradation assessment of p-MOSFETs. As a result, the worst-case conditions of hot-carrier degradation correspond to the gate current maximum ($D_{\mathrm{Igate,max}}$). In scaled devices, on the other hand, a series of bond bombardment events by "colder" carriers can induce multiple vibrational bond excitation [20, 56]. Thus, the multiple-carrier mechanism is dominating if device dimensions shrink. In this situation, the carrier flux rather than the particle energy is important and worst-case conditions correspond to the maximum flux ($D_{\Psi,\mathrm{max}}$). However, even in ultra-scaled devices the high-energy fraction of the carrier ensemble can be populated by scattering mechanisms such as impact ionization, Auger recombination, electron-phonon and electron-electron scattering.

In terms of hot-carrier degradation, the transition from long-channel to short-channel devices occurs at a channel length below 180nm for single gate oxide n-MOSFETs [175, 177]. In this work the investigated dual gate oxide p-MOSFETs of a 130nm technology for automotive and analog applications have a minimum gate length of 370nm and provide a unique possibility to study the interplay of competing single- and multiple-carrier effects in more detail. The shown results underline the importance of this approach and provide a benchmark for device degradation simulations due to the good separability of the observed effects.

The observed drift compensation takes place after 10^4s of stress at $V_{\mathrm{gs}} \approx -1.75$V and is independent of the transistor geometry. It was ascertained that the main influence of the geometry is an increase of the drift ratio with increasing length between 0.37 µm and 0.5 µm as well as a decrease from 0.5 µm to 2 µm for wide devices (Fig. 4.22). Therefore a maximum of the drift ratio is expected between 0.5 µm and 2 µm for the 10 µm wide channel stressed at room and elevated temperatures. This means that the single "hot" carrier effect where the particles have sufficiently high energies to degrade the device is only relevant for a specific channel length while degradation caused by multiple-carrier mechanisms can be found for shorter as well as for longer devices. The finding is valid for room and elevated temperatures and consistent with the hot-carrier degradation model presented in [69, 70].

In addition to that, it has been shown that even for low gate voltages the drift type $D_{\mathrm{Igate,max}}$ will be compensated by $D_{\Psi,\mathrm{max}}$ during long term stress (3.6×10^5s). The degradation turning point, which is depicted in Fig. 4.23, is located at a stress time of 77 ks.

4.2.2. Core device class

Contrary to the analog MOSFETs only stress conditions which do not cause BTI ($|V_{\mathrm{gs}}| \leq V_{\mathrm{nom}} = 1.5\,\mathrm{V}$) during the test time were chosen for degradation measurements with the standard equipment. In addition to that, the largest degradation is assumed for $V_{\mathrm{gs}} = V_{\mathrm{ds}}$ for short channel transistors (see Section 2.7). Therefore, the maximum parameter drift is expected at $|V_{\mathrm{gs}}| = 1.5\,\mathrm{V}$ for any accelerated drain-source voltage. The drift results of the 2-dimensional stress voltage matrix of the core MOSFETs are presented in Fig. 4.24 to Fig. 4.27.

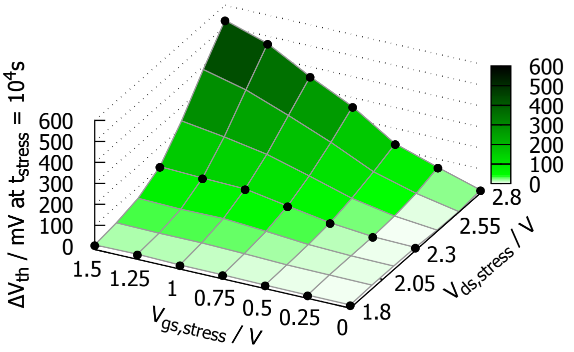


Figure 4.24.: Drift map of the core n-MOSFET with $L = 100\,\mathrm{nm}$, $W = 10\,\mu\mathrm{m}$ stressed at 298 K. The worst hot-carrier related drift is located at $V_{\mathrm{ds}} = 2.8\,\mathrm{V}$, $V_{\mathrm{gs}} = 1.5\,\mathrm{V}$.
Symbols: data points, surface: interpolated

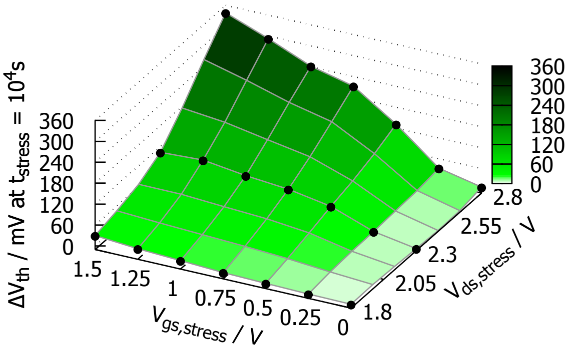


Figure 4.25.: Drift map of the core n-MOSFET with $L = 100\,\mathrm{nm}$, $W = 10\,\mu\mathrm{m}$ stressed at 398 K. The aging behavior is similar to the room temperature data but the maximum threshold voltage drift at the worst-case, $V_{\mathrm{ds}} = 2.8\,\mathrm{V}$, $V_{\mathrm{gs}} = 1.5\,\mathrm{V}$, is about 200 mV lower.
Symbols: data points, surface: interpolated

The maximum threshold voltage drift of the n-MOSFET occurs at $V_{\mathrm{ds}} = 2.8\,\mathrm{V}$, $V_{\mathrm{gs}} = 1.5\,\mathrm{V}$ and $T = 298\,\mathrm{K}$. For any lower gate- and drain-source voltage combination the degradation after 10^4 s is smaller. In addition to that the high temperature degradation is about 200 mV below the room temperature results at the worst case voltage condition (see Fig. 4.28). The detailed degradation information is presented in Fig. 4.30. From the drain current data it is obvious that the device was subjected to very strong hot-carrier stress because a saturation of the drift is observable.

The core p-MOSFET shows a little more hot-carrier degradation at 398 K than at room temperature (see Fig. 4.29). The maximum threshold voltage drift is observable at $V_{\mathrm{ds}} = -2.8\,\mathrm{V}$, $V_{\mathrm{gs}} = -1.5\,\mathrm{V}$ and becomes smaller for any lower drain- or gate-source voltage. Only at $V_{\mathrm{gs,stress}} = -1.25\,\mathrm{V}$ the drift of the room temperature data is higher. A strong change of the slope of the fitted threshold voltage degradation data indicates an NBTI activation for $V_{\mathrm{gs,stress}} < -1.25\,\mathrm{V}$ at elevated temperatures (with respect to the measurement delay). In Fig. 4.30 and 4.31 the worst case degradation of the threshold

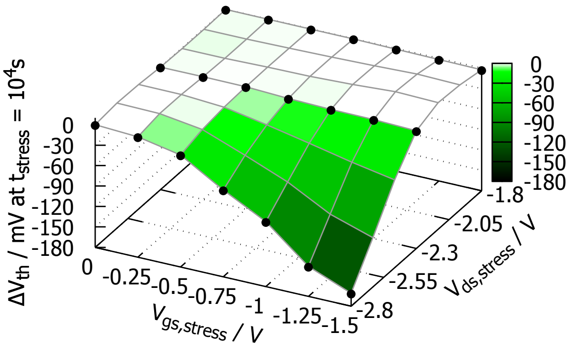


Figure 4.26.: Drift map of the core p-MOSFET with $L = 100\,\text{nm}$, $W = 10\,\mu\text{m}$ stressed at 298 K. The worst hot-carrier related degradation takes place at $V_{\text{ds}} = -2.8\,\text{V}$, $V_{\text{gs}} = -1.5\,\text{V}$.
Symbols: data points, surface: interpolated

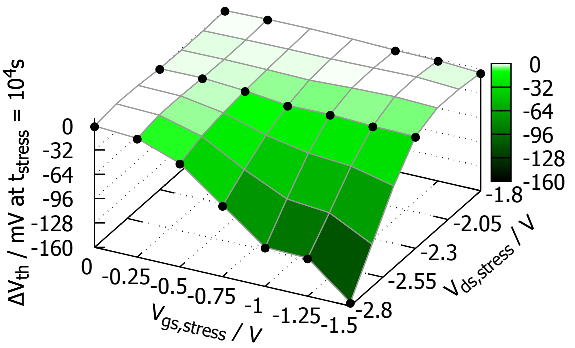


Figure 4.27.: Drift map of the core p-MOSFET with $L = 100\,\text{nm}$, $W = 10\,\mu\text{m}$ stressed at 398 K. The aging behavior is similar to the room temperature data only the threshold voltage drift at the worst-case is about 10 mV lower.
Symbols: data points, surface: interpolated

voltage and drain-source currents is presented in detail for the core n- and p-MOSFET. A comparison between the n-MOSFET and p-MOSFET drift data shows that (like for the analog device class) the p-MOSFET degrades less when the same absolute hot-carrier bias conditions are applied.

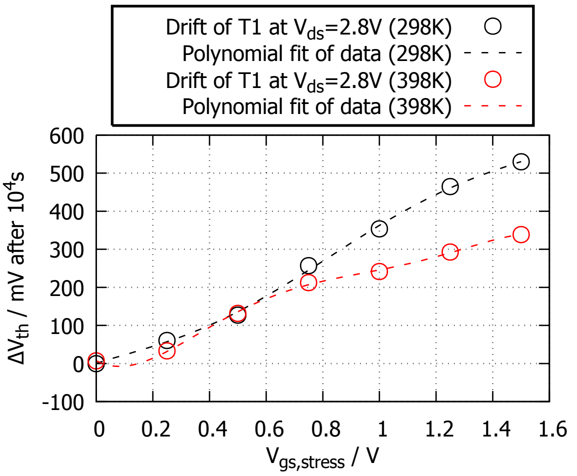


Figure 4.28.: The core n-MOSFET with $L = 100\,\text{nm}$, $W = 10\,\mu\text{m}$ shows a higher drift ($\Delta V_{\text{th,298K}} = 157\,\text{mV}$) at 298 K compared to the 398 K data for any $V_{\text{gs,stress}}$. In contrast to the analog device no actual PBTI activation is observable in the elevated temperature data.

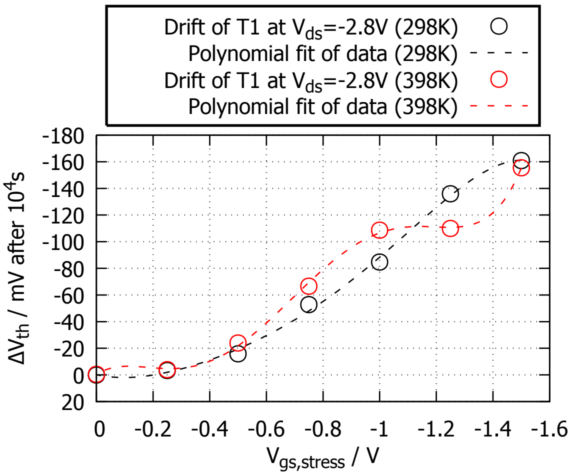


Figure 4.29.: The core p-MOSFET with $L = 100\,\text{nm}$, $W = 10\,\mu\text{m}$ degrades stronger at 398 K than at 298 K. Only at $V_{\text{gs,stress}} < -1.25\,\text{V}$ this picture is different. A strong NBTI activation ($t_{\text{recovery}} \approx 100\,\text{ms}$) takes place at $V_{\text{gs,stress}} < -1.25\,\text{V}$ for $T = 398\,\text{K}$ which can be concluded from the strong change of the slope of the threshold voltage drift.

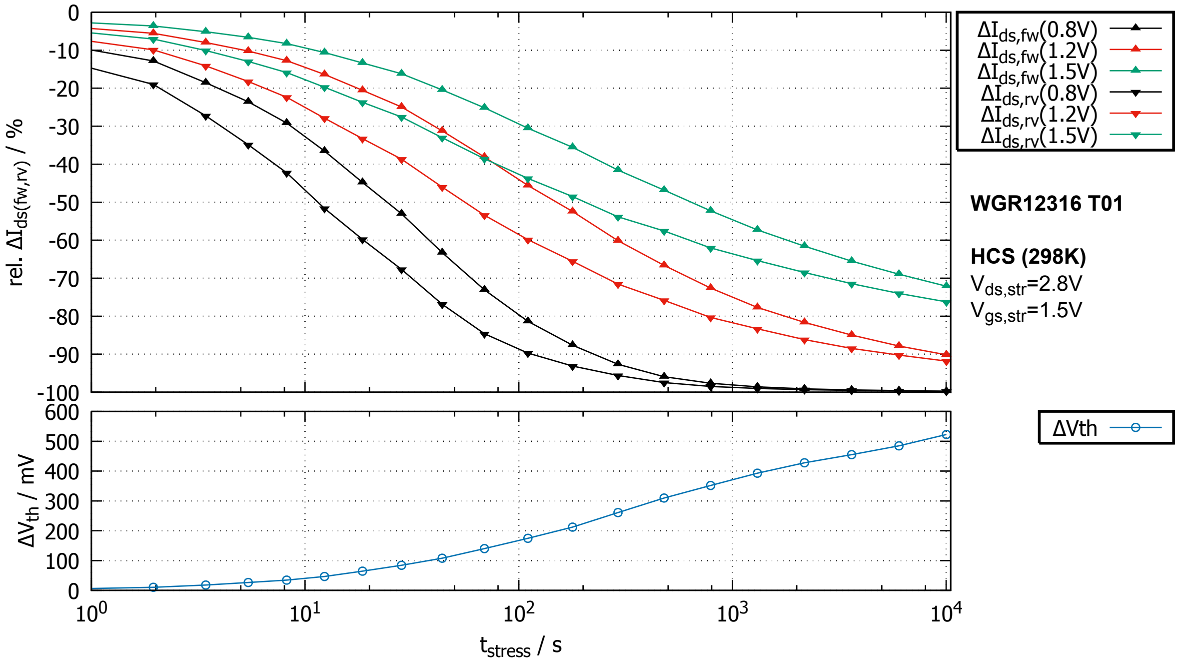


Figure 4.30.: Drift details of the core n-MOSFET with $L = 100\,\mathrm{nm}$, $W = 10\,\mu\mathrm{m}$ stressed at 298 K and worst-case hot-carrier voltage conditions. The maximum current degradation is $\Delta I_{\mathrm{ds,rv}}\,(V_{\mathrm{ds}} = V_{\mathrm{gs}} = 0.8\,\mathrm{V}) \approx 100\%$ and the final threshold voltage drift is about 530 mV. Because a very high stress bias was chosen also the saturation of the degradation is observable in the drain current data. Lines serve as guides to the eye.

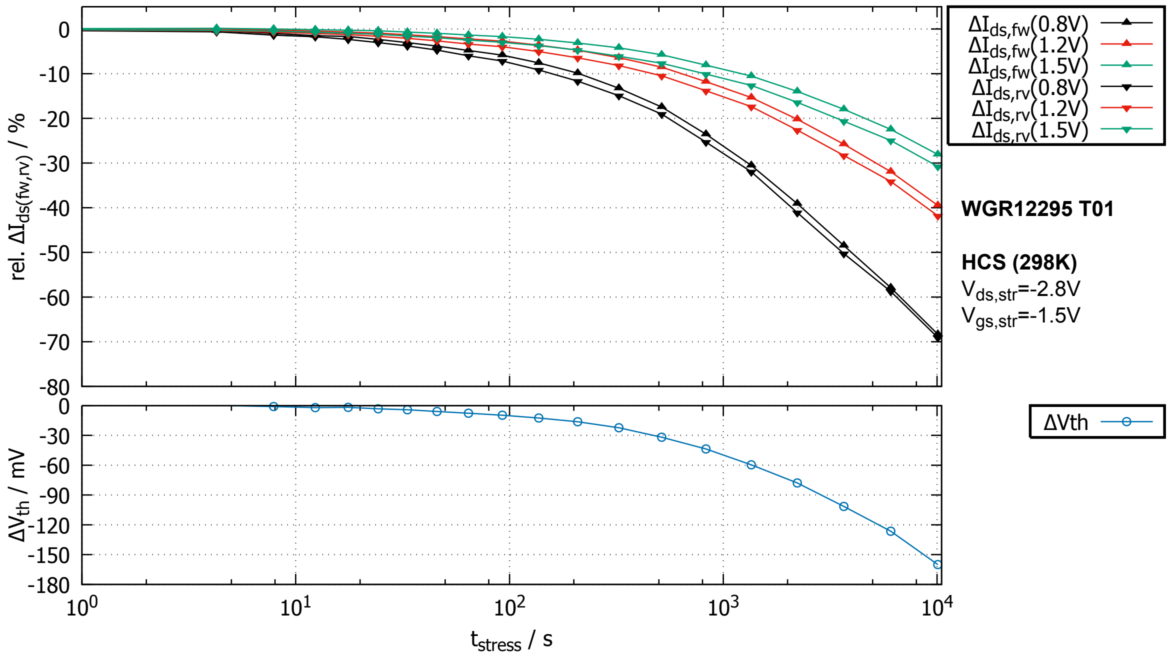


Figure 4.31.: Drift details of the core p-MOSFET with $L = 100\,\mathrm{nm}$, $W = 10\,\mu\mathrm{m}$ stressed at 298 K and worst-case hot-carrier voltage conditions. The maximum current degradation is $\Delta I_{\mathrm{ds,rv}}\,(V_{\mathrm{ds}} = V_{\mathrm{gs}} = 0.8\,\mathrm{V}) \approx 69\%$ and the final threshold voltage drift is about $-160\,\mathrm{mV}$. As for the analog device also the core p-MOSFET shows smaller degradation compared to the n-MOSFET (see Fig. 4.30) with respect to the chosen stress voltage combination. Lines serve as guides to the eye.

4.3. Degradation and recovery measurements with ultra-fast equipment

4.3.1. Analog device class

The very first NBTI degradation measurements of this work were performed with the analog p-MOSFET. A comparison between the data ($T = 443\,\mathrm{K}$, $V_{\mathrm{gs,stress}} = -5.3\,\mathrm{V}$) and the simulated stress and recovery behavior shows good agreement (Fig. 4.33). Only a few differences are observable like an underestimated threshold voltage shift for stress times between 1 s and 10^3 s. The relaxation time, where this error begins, becomes larger with increasing stress time. The calculated CET map depicts that the recovery time always exceeds the stress time (see Fig. 4.32).

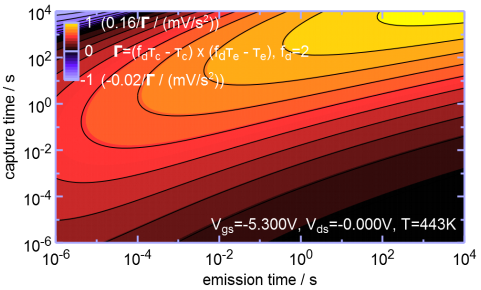
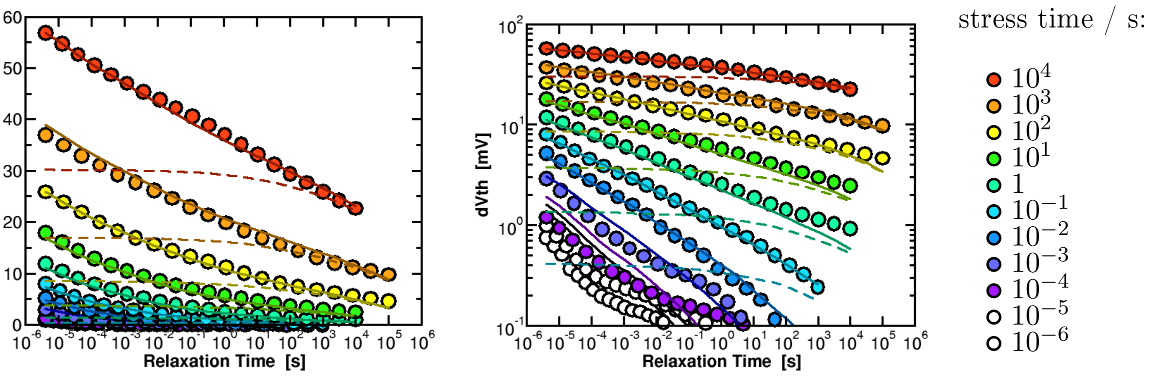


Figure 4.32.: CET map of the analog p-MOSFET calculated from the data shown in Fig. 4.33 and limited to the measurement range. The maximum recovery probability density is colored yellow.



linear scaled ΔV_{th} logarithmic scaled ΔV_{th}

Figure 4.33.: Comparison between measurement data (symbols) and a trap based simulation performed after [166] (solid line, permanent part: dashed line) of the threshold voltage shift of the core p-MOSFET stressed at 443 K. The simulation is based on the model of [178] and shows good agreement to the measured data. Only small differences can be identified, e.g. the slope of the recovery traces for $t_{\mathrm{str}} = 1\,\mathrm{s}..10^3\,\mathrm{s}$ is leading to an underestimated ΔV_{th} for long recovery times. The relaxation time onset of this deviation increases with stress time.

4.3.2. Core device class

The core p-MOSFET is the main focus of this work because it is used in many circuits and known to have recognizable NBTI degradation including recovery effects. Typical degradation and recovery data are shown in Fig. 4.34 for homogeneous NBTI at $V_{\mathrm{gs,stress}} = -2.8\,\mathrm{V}$ and $T = 398\,\mathrm{K}$. The influence of the measurement delay on the power law exponent of the threshold voltage drift is very strong and changes it by a factor of 2 from $t_{\mathrm{rec}} = 1.8 \times 10^{-6}\,\mathrm{s}$ to $t_{\mathrm{rec}} = 1.0\,\mathrm{s}$. The analyzed transistor has a medium sized gate area which on the one hand is big enough to show smooth recovery traces and on the other hand has a large W/L ratio allowing ultra-short measurement delays. Only low noise is recognizable in the recovery data of this p-MOSFET with $1\,\mu\mathrm{m}^2$ gate area.

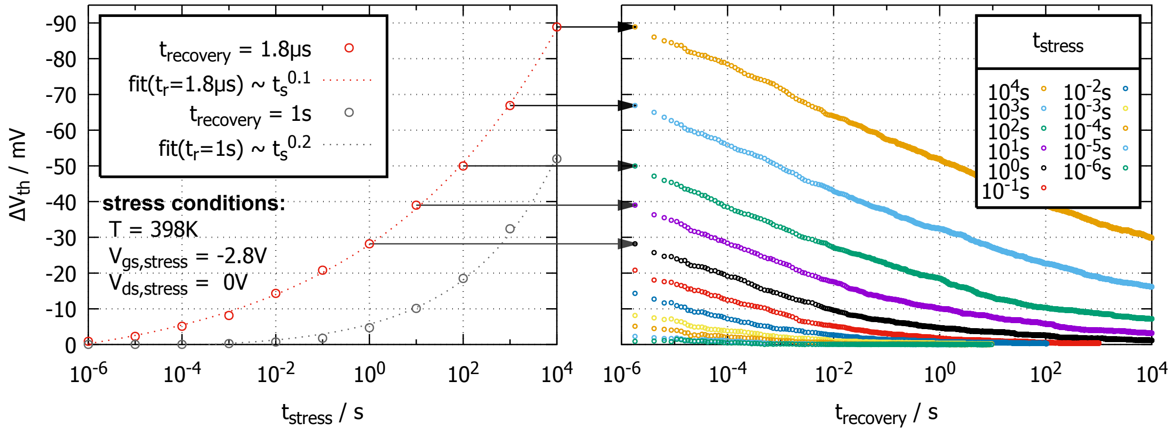


Figure 4.34.: Left: After each stress time a recovery trace is measured (arrows). The first readout of the recovery trace is shown here with a measurement delay of $1.8 \times 10^{-6}\,\mathrm{s}$. The depicted drift of the core p-MOSFET follows a power law of the stress time. Dependent on the recovery time the power law exponent changes from $n = 0.1$ for $t_{\mathrm{rec}} = 1.8 \times 10^{-6}\,\mathrm{s}$ (red) to $n = 0.2$ for $t_{\mathrm{rec}} = 1.0\,\mathrm{s}$ (gray). **Right:** The complete recovery traces after each stress time show only little noise which is caused by the transistor due to its rather small gate area of $1\,\mu\mathrm{m}^2$. The data of stress times above 1 s contain a threshold voltage shift offset from the previous stress decade.

The influence of the chosen setpoint of the operational amplifier (i.e. drain-source current of the transistor) on the threshold voltage drift needs to be taken into account because it changes the measured operating point of the device under test. As shown in Fig. 4.35 the relative difference between the measured threshold voltage shift at $I_{\mathrm{ds}} = -2.5 \times 10^{-7}\,\mathrm{A}$ and $-5 \times 10^{-5}\,\mathrm{A}$ is about 35%. In addition to that also the readout voltage $V_{\mathrm{ds,read}}$ impacts the threshold voltage drift assessment. The relative difference between the linear ($V_{\mathrm{ds,read}} = -0.1\,\mathrm{V}$) and the saturated threshold voltage ($V_{\mathrm{ds,read}} = -1.5\,\mathrm{V}$) ranges from 25% to 35% (see Fig. 4.36).

The CET maps of the linear and saturated threshold voltage data illustrate that for $V_{\mathrm{ds,read}} = -0.1\,\mathrm{V}$ the maximum of the probability density of the recoverable defect

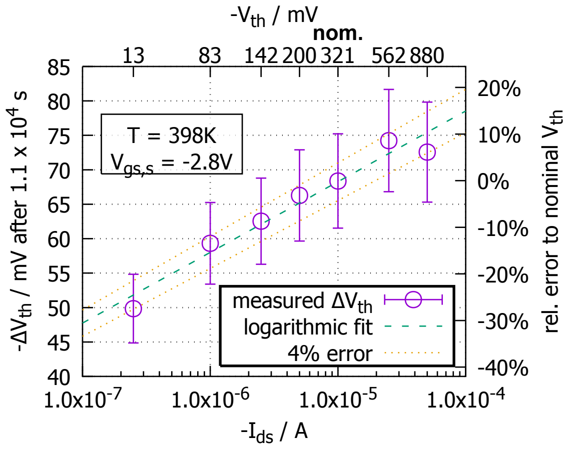


Figure 4.35.: The threshold voltage drift after 1.1×10^4 s is dependent on the set value of the drain-source current of the transistor. This dependence can be fitted by a logarithmic function within an error of 4%. The shown error bars denote the typical device to device variation of 10%.

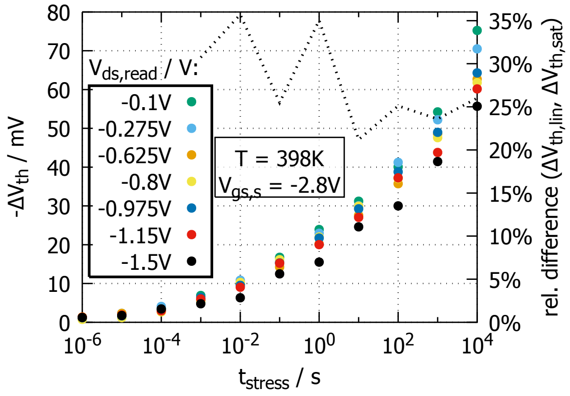
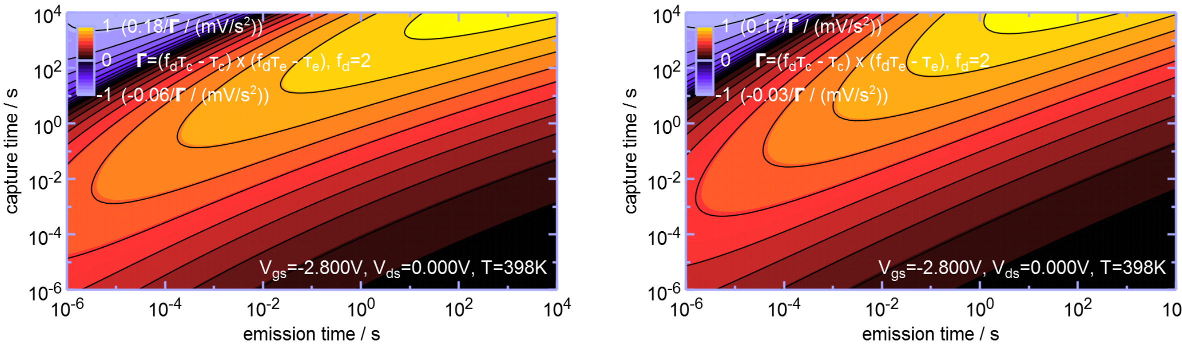


Figure 4.36.: The chosen readout voltage $V_{\mathrm{ds,read}}$ has a strong impact on the resulting threshold voltage drift after 1.1×10^4 s. The relative difference (dashed line) between the measured linear ($V_{\mathrm{ds,read}} = -0.1$ V) and saturated ($V_{\mathrm{ds,read}} = -1.5$ V) threshold voltage drift is between 25% and 35%.

classes has a steeper slope between $t_{\mathrm{emission}} = 10^{-6}$ s and 10^{-1} s compared to $V_{\mathrm{ds,read}} = -1.5$ V. In addition to that the integrated probability density is also higher for the linear case ($\Delta V_{\mathrm{th,lin}} > \Delta V_{\mathrm{th,sat}}$). This implies that the linear threshold voltage is the better parameter to address more defects by the measurement.

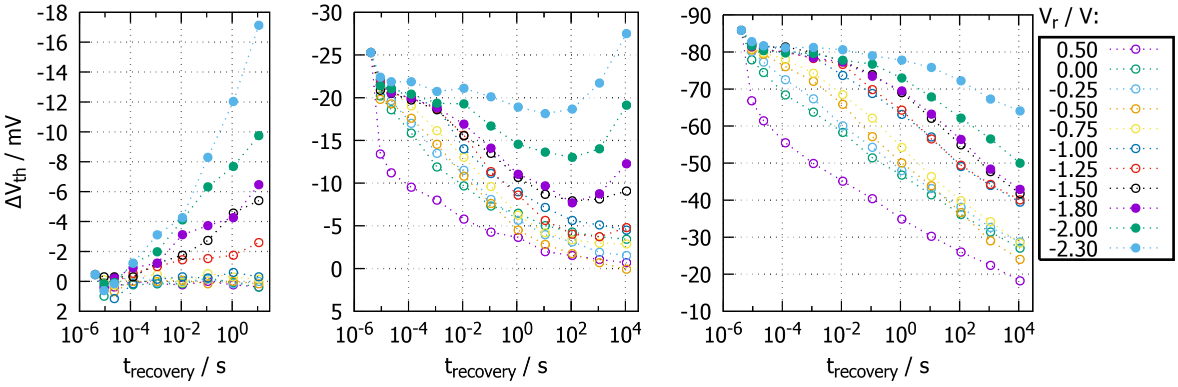


$V_{\mathrm{ds,read}} = -0.1$ V, $V_{\mathrm{th,0h}} = -538.9$ mV.

$V_{\mathrm{ds,read}} = -1.5$ V, $V_{\mathrm{th,0h}} = -398.2$ mV.

Figure 4.37.: The CET maps of the linear (left) and saturated (right) threshold voltage show a different position and slope of the maximum probability density of the recoverable defect classes. In general the measurement of the linear threshold voltage yields more recoverable defects of one class. For $V_{\mathrm{ds,read}} = -0.1$ V the slope is shallower and shorter capture times are needed compared to the case at $V_{\mathrm{ds,read}} = -1.5$ V to get a comparable amount of recoverable defects within the same defect class. The maximum threshold voltage drift is influenced by the drain-source readout voltage as shown in Fig. 4.36 and is $\Delta V_{\mathrm{th}} = -75.3$ mV for the linear and $\Delta V_{\mathrm{th}} = -55.7$ mV for the saturation case.

During the recovery measurement it is possible to apply a constant voltage V_r to the gate terminal between the readouts. Dependent on the stress time the recovery signal with this additional bias behaves very differently as shown in Fig. 4.38. After a short stress time any $V_\mathrm{r} < -1\,\mathrm{V}$ leads to an additional degradation of the threshold voltage. For $t_\mathrm{str} = 1\,\mathrm{s}$ this is still valid but after a first improvement of the threshold voltage the subsequent onset of the degradation depends on the chosen V_r and varies from $t_\mathrm{rec} = 10^{-4}\,\mathrm{s}$ ($V_\mathrm{r} = -2.80\,\mathrm{V}$) up to $t_\mathrm{rec} = 10^{3}\,\mathrm{s}$ ($V_\mathrm{r} = -1.25\,\mathrm{V}$). After $10^4\,\mathrm{s}$ of stress only a bias of $V_\mathrm{r} = -2.8\,\mathrm{V}$ can cause a further increase of the threshold voltage degradation. The accumulative signal ($V_\mathrm{r} = 0.5\,\mathrm{V}$) always accelerates the recovery and thereby strongly changes the curvature of the recovery trace at $t_\mathrm{str} = 10^4\,\mathrm{s}$. The influence of V_r on the recovery time is of high interest for compact aging models. Therefore all recovery traces with $t_\mathrm{str} = 1 \times 10^4\,\mathrm{s}$ which show no post-stress aging were scaled to the recovery trace at $V_\mathrm{r} = -0.5\,\mathrm{V}$ by multiplying their t_r component with an acceleration factor a_{t_r}. The result is depicted in Fig. 4.39 together with the fit of the acceleration factor of the recovery time which follows an exponential law of the chosen V_r. With this information the recovery of the stressed transistors can be formulated as a function of the recovery time t_r and the recovery voltage V_r, $\Delta V_\mathrm{th}\left(t_\mathrm{r}, V_\mathrm{r}\right) = \Delta V_\mathrm{th}\left(a_{t_\mathrm{r}}\left(V_\mathrm{r}\right) t_\mathrm{r}, -0.5\,\mathrm{V}\right)$ in a compact model used for circuit simulations (see Chapter 5).



$t_\mathrm{str} = 1 \times 10^{-6}\,\mathrm{s}$ $t_\mathrm{str} = 1\,\mathrm{s}$ $t_\mathrm{str} = 1 \times 10^{4}\,\mathrm{s}$

Figure 4.38.: V_r dependence of the recovery behavior. **Left:** The absolute threshold voltage drift further increases with the recovery time for $V_\mathrm{r} \leq -1.25\,\mathrm{V}$. **Middle:** After a first recovery of the threshold voltage the degradation increases again for $2.8\,\mathrm{V} > |V_\mathrm{r}| > V_\mathrm{nom}$ and $t_\mathrm{rec} \geq 10\,\mathrm{s}$. For $V_\mathrm{r} = -2.8\,\mathrm{V}$ the onset of this effect is located at about $10^{-4}\,\mathrm{s}$. **Right:** The readout delay leads to a recovery of the threshold voltage for the longest stress time for a few conditions with $2.8\,\mathrm{V} > |V_\mathrm{r}| > V_\mathrm{nom}$. Only for $V_\mathrm{r} = -2.8\,\mathrm{V}$ a further degradation is observable. The recovery traces of each V_r condition (always a fresh transistor) were scaled to the same starting point to eliminate device to device variations.

To discharge traps with short emission times a positive (*accumulative*) pulse can be applied to the gate terminal before the measurement of the recovery trace. This has been performed for three different pulse lengths ($t_\mathrm{acc.pulse} = 10^{-1}\,\mathrm{s}, 1\,\mathrm{s}, 10\,\mathrm{s}$) after homogeneous

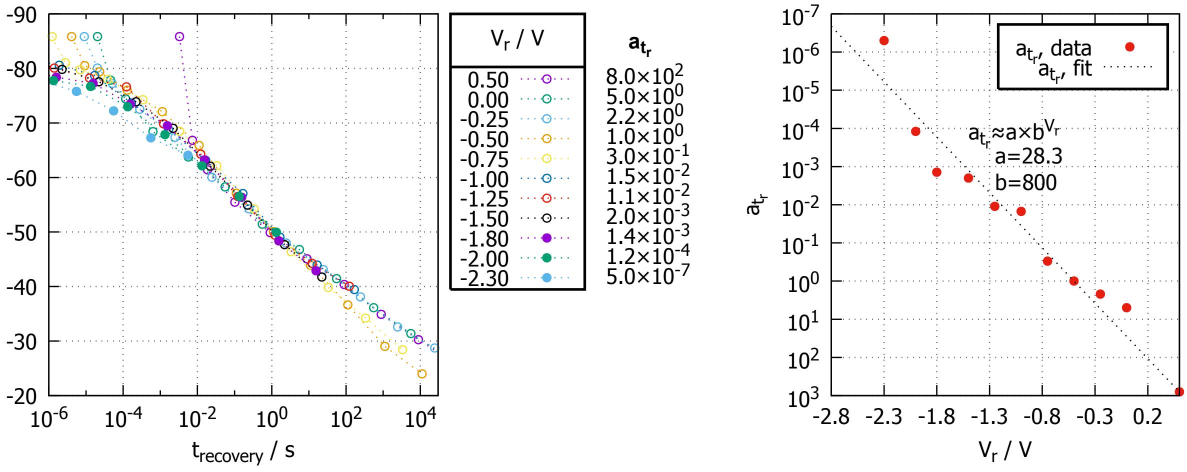


Figure 4.39.: The acceleration factor a_{t_r} of the recovery time is strongly dependent on the chosen V_r and can be fitted with good agreement to the measurement data via an exponential law.

NBTI stress of 10 s at $V_{\text{gs,stress}} = -2.8\,\text{V}$. The results presented in Fig. 4.40 indicate that the number of traps which can be discharged increases with the pulse length leading to a threshold voltage drift ranging between $-39\,\text{mV}$ and $-32\,\text{mV}$. After a plateau phase the start of the recovery ($t_{\text{onset}} = 3 \times 10^{-1}\,\text{s}...6\,\text{s}$) also depends on the applied length of the accumulative pulse ($t_{\text{acc.pulse}} = 10^{-1}\,\text{s}...10\,\text{s}$).

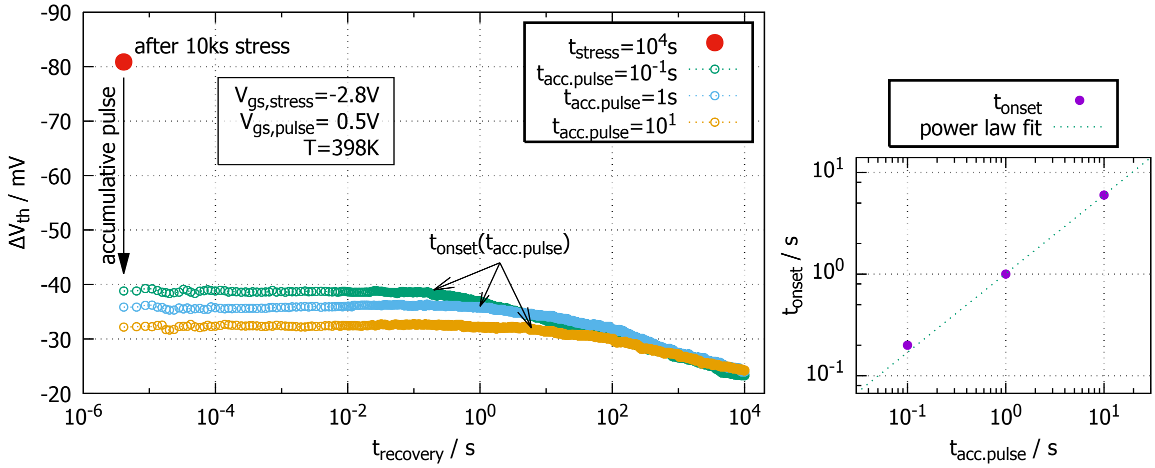


Figure 4.40.: **Left:** After the p-MOSFET has been stressed at $V_{\text{gs}} = -2.8\,\text{V}$ and $T = 398\,\text{K}$ for $10^4\,\text{s}$ the threshold voltage drift is determined (red). Subsequently, a positive pulse with $V_{\text{gs}} = 0.5\,\text{V}$ and different durations $t_{\text{acc.pulse}}$ for each tested transistor is applied. The subsequent measurement (green/blue/orange) indicates that the number of traps with short recovery times which can be discharged by the pulse increases with the pulse length, thus leading to a smaller threshold voltage drift. Also the recovery after the plateau phase begins at different recovery times, $t_{\text{onset}}\left(t_{\text{acc.pulse}} = 10^{-1}\,\text{s}\right) \approx 3 \times 10^{-1}\,\text{s}$ (green), $t_{\text{onset}}\left(t_{\text{acc.pulse}} = 1\,\text{s}\right) \approx 1\,\text{s}$ (blue), $t_{\text{onset}}\left(t_{\text{acc.pulse}} = 10\,\text{s}\right) \approx 6\,\text{s}$ (orange). **Right:** The onset of the recovery can be fitted well by a power law of the accumulative pulse duration.

The measurement of operating points can either be realized as described in Section 3.1.2 or without the record of the threshold voltage. The latter allows a determination of the operating points, which is less influenced by the recovery of the degradation. Single measurements of the linear and saturation mode drain-source current following the time recipe of CET map recording (see Table B.1) were performed and are shown in Fig. 4.41 and Fig. 4.42. As expected the current recovers towards its initial value during the monitoring phase. The applied homogeneous NBTI stress with $V_{\mathrm{gs,stress}} = -2.8\,\mathrm{V}$ at $T = 398\,\mathrm{K}$ leads to a larger degradation of the saturation mode operating point compared to the linear mode case. Also the slopes of the recovery traces are different (compare with Fig. 4.37). At $t_{\mathrm{str}} = 10^2\,\mathrm{s}$ (light blue) and $t_{\mathrm{rec}} = 6 \times 10^{-5}\,\mathrm{s}$ the relative drift of $I_{\mathrm{ds,lin}}$ is about 1.6% less compared to the one of $I_{\mathrm{ds,sat}}$ whereas at $t_{\mathrm{rec}} = 10^1\,\mathrm{s}$ it is about 0.5% more (the device to device variation is below those values).

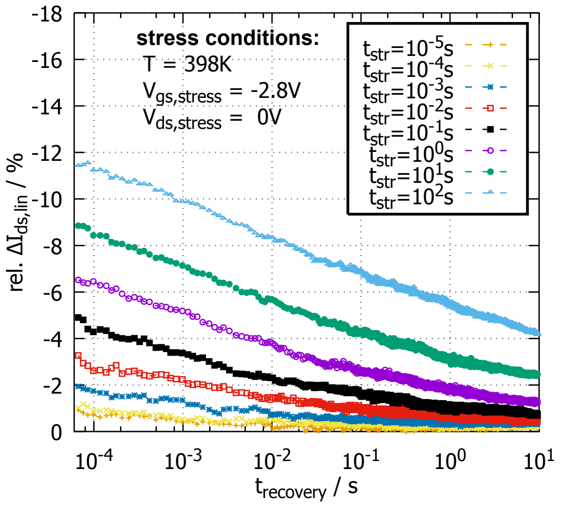


Figure 4.41.: Recovery traces of the linear mode drain-source current ($V_{\mathrm{gs}} = -0.1\,\mathrm{V}$, $V_{\mathrm{ds}} = -1.5\,\mathrm{V}$).

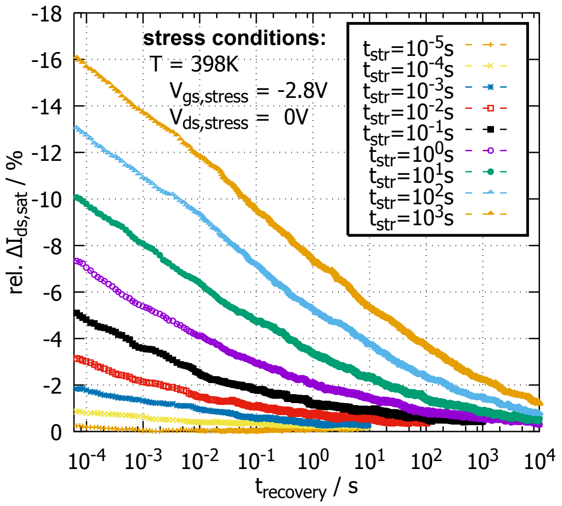


Figure 4.42.: Recovery traces of the saturation mode drain-source current ($V_{\mathrm{gs}} = V_{\mathrm{ds}} = -1.2\,\mathrm{V}$).

Influence of the stress temperature on CET maps of homogeneous NBTI

CET map data were recorded for the whole 2-dimensional stress voltage matrix with varying $V_{\mathrm{ds,stress}}$ and $V_{\mathrm{gs,stress}}$ at 398 K and 443 K. Once the temperature is increased the homogeneous NBTI stress data (Fig. 4.43) indicate a shift to shorter capture and emission times of the probability density maximum (yellow area). In addition, the amount of recoverable defects is higher, which can be seen by the values of the color mapping $(\max(g, 398\,\mathrm{K}) = 0.18/\Gamma\mathrm{mV\,s^{-2}} < 0.27/\Gamma\mathrm{mV\,s^{-2}} = \max(g, 443\,\mathrm{K}))$.

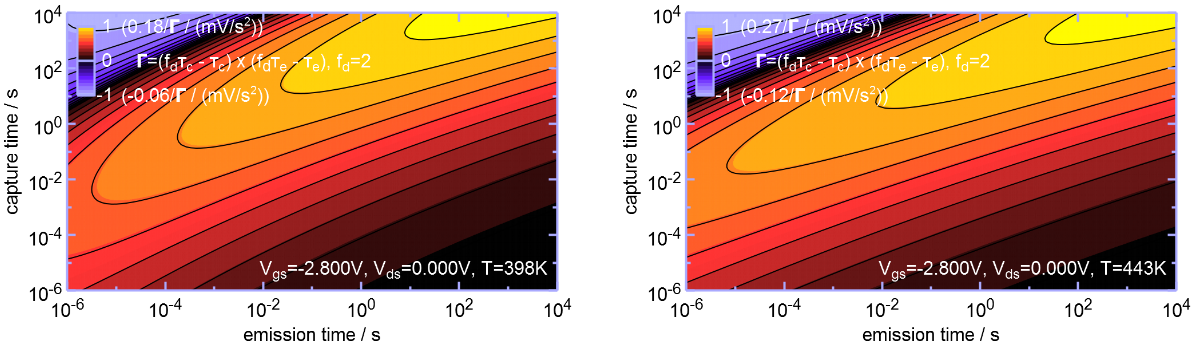


Figure 4.43.: CET maps of homogenous NBTI stress, $V_{\mathrm{gs,stress}} = -2.8\,\mathrm{V}$, $V_{\mathrm{ds,stress}} = 0\,\mathrm{V}$ at 398 K and 443 K.

The temperature activation of defect capture and emission time constants can be described by an Arrhenius law [153, 179],

$$\tau_{\mathrm{c/e}} = \tau_{0,\mathrm{c/e}} \mathrm{e}^{\left(\beta E_{\mathrm{A,c/e}}\right)} \Leftrightarrow E_{\mathrm{A,c/e}} = \beta^{-1} \ln \frac{\tau}{\tau_{0,\mathrm{c/e}}} \tag{4.1}$$

with the activation energies $E_{\mathrm{A,c/e}}$ for hole capture and emission and the thermodynamic beta $\beta = 1/(k_{\mathrm{B}}T)$. For a large ensemble of defects the prefactors $\tau_{0,\mathrm{c}}$ and $\tau_{0,\mathrm{e}}$ are expected to have the same effective average value for all defects [166].

Taking two different temperatures, T_1 and T_2, into account it follows:

$$\begin{aligned}
&\left| \begin{array}{l} E_{\mathrm{A,c/e}} = \beta^{-1}(T_1) \ln \frac{\tau(T_1)}{\tau_{0,\mathrm{c/e}}} \\ E_{\mathrm{A,c/e}} = \beta^{-1}(T_2) \ln \frac{\tau(T_2)}{\tau_{0,\mathrm{c/e}}} \end{array} \right| \Rightarrow \beta^{-1}(T_1) \ln \frac{\tau(T_1)}{\tau_{0,\mathrm{c/e}}} = \beta^{-1}(T_2) \ln \frac{\tau(T_2)}{\tau_{0,\mathrm{c/e}}} \\
&\Leftrightarrow \left(\mathrm{e}^{\ln \frac{\tau(T_1)}{\tau_{0,\mathrm{c/e}}}} \right)^{\beta^{-1}(T_1)} = \left(\mathrm{e}^{\ln \frac{\tau(T_2)}{\tau_{0,\mathrm{c/e}}}} \right)^{\beta^{-1}(T_2)} \Leftrightarrow \left(\frac{\tau(T_1)}{\tau_{0,\mathrm{c/e}}} \right)^{\beta^{-1}(T_1)} = \left(\frac{\tau(T_2)}{\tau_{0,\mathrm{c/e}}} \right)^{\beta^{-1}(T_2)} \\
&\Leftrightarrow \frac{\tau(T_2)}{\tau_{0,\mathrm{c/e}}} = \left(\frac{\tau(T_1)}{\tau_{0,\mathrm{c/e}}} \right)^{\frac{\beta^{-1}(T_1)}{\beta^{-1}(T_2)}} \Leftrightarrow \tau(T_2) = \tau_{0,\mathrm{c/e}} \left(\frac{\tau(T_1)}{\tau_{0,\mathrm{c/e}}} \right)^{\frac{T_1}{T_2}}
\end{aligned} \tag{4.2}$$

This thermal equivalent time $\tau(T_2)$ is also referred to as *temperature time* by [180].

The capture and emission times of the defect classes measured at $T_1 = 443\,\mathrm{K}$ can be converted to $T_2 = 398\,\mathrm{K}$ using (4.2) with the extracted $\tau_{0,\mathrm{c}} = 1.1 \times 10^{-9}\,\mathrm{s}$ and $\tau_{0,\mathrm{e}} = 4.1 \times 10^{-15}\,\mathrm{s}$. At this, a fit of the minimum difference between the two CET maps is used for the extraction of $\tau_{0,\mathrm{c}}$ and $\tau_{0,\mathrm{e}}$. Following Eq. (1.119) and (4.2) this yields for the differential CET map of two temperatures:

$$\mathcal{D}\left(\mathrm{MAP1}\left(T_1\right), \mathrm{MAP2}\left(T_2\right)\right) = \frac{\delta^2 \Delta V_{\mathrm{th}}\left(\mathrm{MAP1}\left(T_1\right)\right)}{\delta t_{\mathrm{s}} \delta t_{\mathrm{r}}} - \frac{\delta^2 \Delta V_{\mathrm{th}}\left(\mathrm{MAP2}\left(T_2\right)\right)}{\delta t_{\mathrm{s}} \delta t_{\mathrm{r}}} \tag{4.3}$$

$$\begin{aligned}
= \; & \left(\Delta V_{\mathrm{th,MAP1}}\left(\vartheta\left(t_{\mathrm{s,i}}\right), \vartheta\left(t_{\mathrm{r,j}}\right)\right) - \Delta V_{\mathrm{th,MAP1}}\left(\vartheta\left(t_{\mathrm{s,i-1}}\right), \vartheta\left(t_{\mathrm{r,j}}\right)\right)\right) \\
& - \left(\Delta V_{\mathrm{th,MAP1}}\left(\vartheta\left(t_{\mathrm{s,i}}\right), \vartheta\left(t_{\mathrm{r,j-1}}\right)\right) - \Delta V_{\mathrm{th,MAP1}}\left(\vartheta\left(t_{\mathrm{s,i-1}}\right), \vartheta\left(t_{\mathrm{r,j-1}}\right)\right)\right) \\
& - \left(\Delta V_{\mathrm{th,MAP2}}\left(t_{\mathrm{s,i}}, t_{\mathrm{r,j}}\right) - \Delta V_{\mathrm{th,MAP2}}\left(t_{\mathrm{s,i-1}}, t_{\mathrm{r,j}}\right)\right) \\
& - \left(\Delta V_{\mathrm{th,MAP2}}\left(t_{\mathrm{s,i}}, t_{\mathrm{r,j-1}}\right) - \Delta V_{\mathrm{th,MAP2}}\left(t_{\mathrm{s,i-1}}, t_{\mathrm{r,j-1}}\right)\right) \\
& \forall \left(t_{\mathrm{s,i-1}} < t_{\mathrm{s,i}}, t_{\mathrm{r,j-1}} < t_{\mathrm{r,j}}\right), \\
\text{with} \quad & \vartheta\left(t_{\mathrm{s,i}}\right) = \tau_{0,\mathrm{c}} \left(\frac{t_{\mathrm{s,i}}}{\tau_{0,\mathrm{c}}}\right)^{\frac{T_2}{T_1}}, \vartheta\left(t_{\mathrm{r,j}}\right) = \tau_{0,\mathrm{e}} \left(\frac{t_{\mathrm{r,j}}}{\tau_{0,\mathrm{e}}}\right)^{\frac{T_2}{T_1}}
\end{aligned}$$

The calculated differential CET map of the two CET maps shown in Fig. 4.43 – the 443 K data was converted to 398 K prior the subtraction – is presented in Fig. 4.44 and reveals that by increasing the stress temperature additional defects occur while others vanish. Contrary to the results of single recovery traces shown in [181] the differential picture of both CET maps does not result in $\delta^2 \Delta V_{\mathrm{th}}\left(\mathrm{MAP1}\left(T_1\right)\right) / \left(\delta t_{\mathrm{s}} \delta t_{\mathrm{r}}\right) - \delta^2 \Delta V_{\mathrm{th}}\left(\mathrm{MAP2}\left(T_2\right)\right) / \left(\delta t_{\mathrm{s}} \delta t_{\mathrm{r}}\right) = 0$ for all combinations of capture and emission time constants but shows a distinct area of additionally activated defects (yellow) at $\tau_{\mathrm{e}} = 100\,\mathrm{s}$, $\tau_{\mathrm{c}} = 100\,\mathrm{s}$ and an area of defect disappearance at the upper left corner (blue). This implies that the barrier energy (4.1) of the ensemble may change with temperature leading to an increased (yellow) or decreased (blue) probability density of defect classes for a given capture and emission time. Because of that the simplification of the time constants $\tau_{0,\mathrm{c}}$ and $\tau_{0,\mathrm{e}}$ to a single value representing the defect ensemble [166] is questionable.

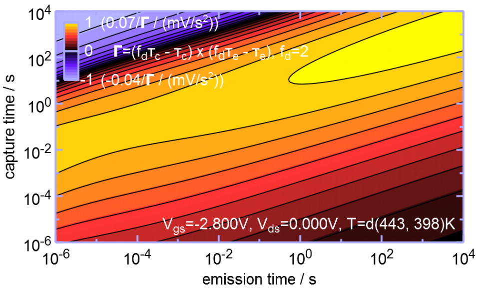


Figure 4.44.: Differential CET map of homogenous NBTI stress, $V_{\mathrm{gs,stress}} = -2.8\,\mathrm{V}$, $V_{\mathrm{ds,stress}} = 0\,\mathrm{V}$ from 398 K to 443 K. The color scale represents the change of presence of defect clusters, yellow signifies the occurrence of new defects, blue their disappearance.

Comparison to large gate area devices and variability of NBTI (I)

In order to reduce the individual noise of the single transistor for the stress time independent fit method the averaged data of 10 adjacent p-MOSFETs with individual gate areas of $1\,\mu\mathrm{m}^2$ was calculated and is presented in Fig. 4.45 together with the CET map of a transistor of the same type and a large gate area of 100 µm^2. A low positive drain-source voltage was applied to the device under test to allow a reasonable settling time as it drives the operational amplifier to the correct output sign during the stress phase (see Fig. 2.8). Compared to e.g. Fig. 2.19 (a) it becomes obvious that the noise of the individual transistor especially for short recovery times vanishes for the mean signal.

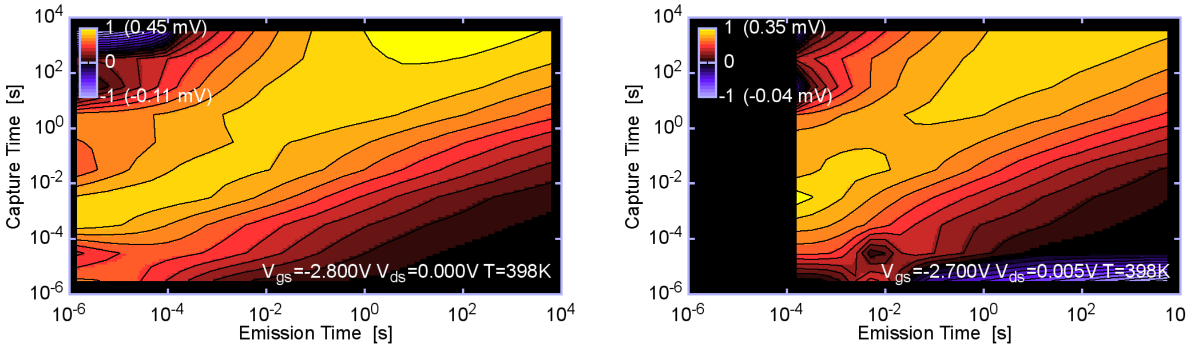


Figure 4.45.: Left: Averaged CET maps (fit method as described in Section 2.6.1) of 10 core p-MOSFETs stressed with homogenous NBTI ($V_{\text{gs,stress}} = -2.8\,\text{V}$, $V_{\text{ds,stress}} = 0\,\text{V}$) at 398 K. **Right:** CET map of a similar stress ($V_{\text{gs,stress}} = -2.7\,\text{V}$, $V_{\text{ds,stress}} = 0.005\,\text{V}$) applied to a large p-MOSFET with a gate area of 100 µm^2. Due to the lower current of the large transistor the settling time is longer ($\approx 1.1 \times 10^{-4}\,\text{s}$) and therefore all data points with recovery times shorter than $1.1 \times 10^{-4}\,\text{s}$ were excluded during the fit and the plot of the CET map.

Furthermore, the standard deviation of the recovery traces of each stress time of the ten transistors shown in Fig. 4.46 reveals three interesting statistical properties. First, a noticeable decrease of $\sigma\left(\Delta V_{\text{th}}\right)$ is observable if the recovery time increases. Second, if recovery times longer than 60 s are recorded a strong increase of $\sigma\left(\Delta V_{\text{th}}\right)$ establishes for long stress times. This finding reveals that for long-term stress the variability of the permanent component exceeds the variability of the recoverable component. And last, the relative standard deviation decreases with a power law of the stress time, $\sigma_{\text{relative}} \propto t_{\text{str}}^n$ with $n \approx -0.18$, including an offset of about 3%. For short stress times the small amount of activated defects dominates the relative variability. For long stress times σ_{relative} saturates. Here, the variability caused by an additionally charged defect has no effect on the relative ensemble statistics.

The detailed analysis of the relative standard deviation presented in Fig. 4.47 illustrates that the fit parameters a and n of the relative variability can nicely be estimated with power laws of the recovery time. This directly enables a prediction of the variability of the threshold voltage degradation in circuit simulations, i.e.

$$\sigma_{\text{relative}} \approx \left(0.12 \times t_{\text{r}}^{0.16}\right) \times t_{\text{s}}^{-0.03\times t_{\text{r}}^{0.11}-0.18} + 0.027. \tag{4.4}$$

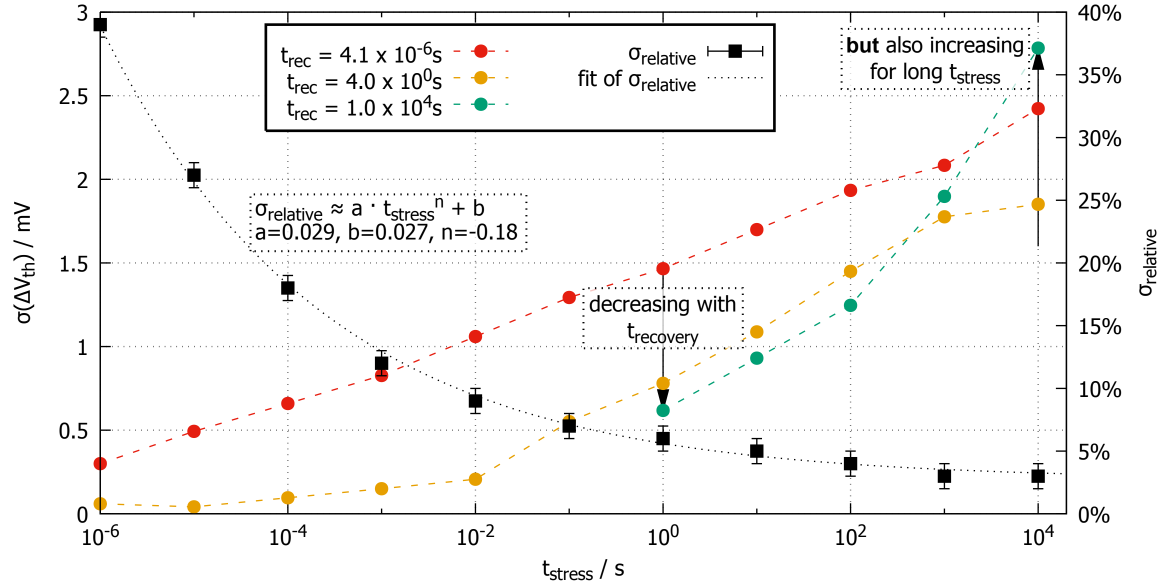


Figure 4.46.: The standard deviation of the threshold voltage shift after homogeneous NBTI stress of 10 adjacent core class p-MOSFETs shows three interesting properties.
First, with increasing recovery time (yellow vs. red data) $\sigma\left(\Delta V_{\mathrm{th}}\right)$ decreases, **second**, once long recovery times ($\geq 60\,\mathrm{s}$) are reached $\sigma\left(\Delta V_{\mathrm{th}}\right)$ shows a strong increase for long stress times (green data) and **third**, $\sigma_{\mathrm{relative}} = \sigma\left(\Delta V_{\mathrm{th}}\left(t_{\mathrm{rec}} = 4.1 \times 10^{-6}\,\mathrm{s}\right)\right) / \left|\Delta V_{\mathrm{th}}\left(t_{\mathrm{rec}} = 4.1 \times 10^{-6}\,\mathrm{s}\right)\right|$ decreases with a power law of the stress time including an offset of about 3%.

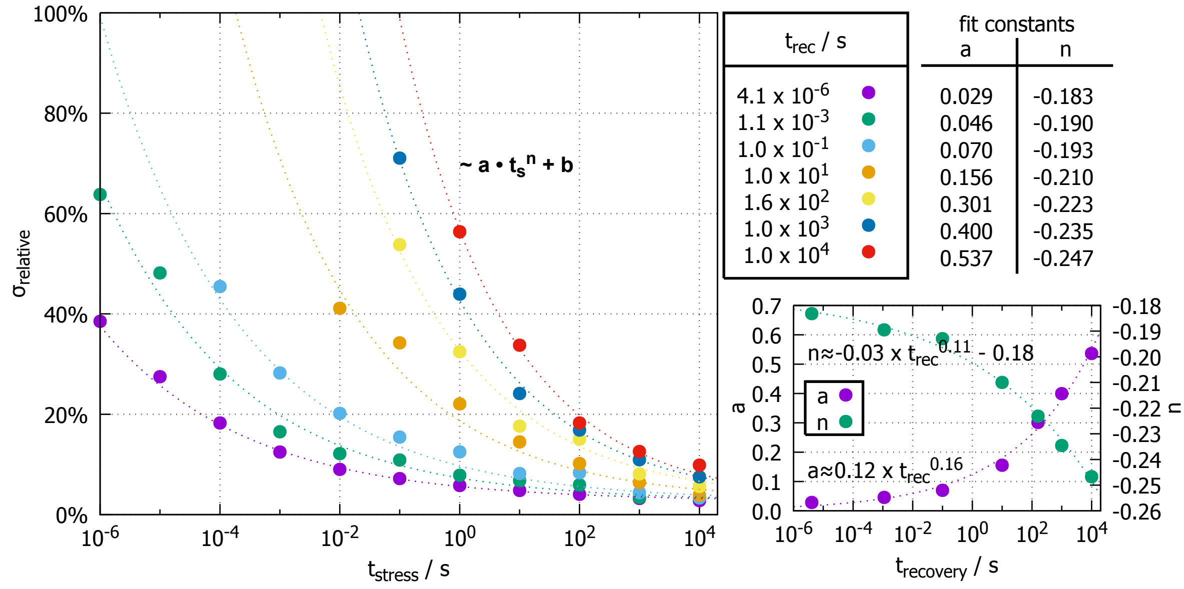


Figure 4.47.: **Left:** The relative variability $\sigma_{\mathrm{relative}} = \sigma\left(\Delta V_{\mathrm{th}}\left(t_{\mathrm{rec}}\right)\right) / \left|\Delta V_{\mathrm{th}}\left(t_{\mathrm{rec}}\right)\right|$ decreases with a power law of the stress time ($b = 0.027$), while (**right**) the prefactor a and the stress time exponent n change with power laws of the recovery time.

4.3.3. Arbitrary signal NBTI stress

Different periodical gate stress signals (left graph of Fig. 4.49) have been applied to core p-MOSFETs in order to obtain the stress pattern dependent aging and recovery response. All measurements were performed on transistors of adjacent chips to keep the variation of the fresh and drift threshold voltage data as low as possible. To enhance NBTI the stress temperature was set to 398 K. The drift over stress time is presented in Fig. 4.48 and indicates that the AC stress pattern with offset (dAC) is the most critical one for long term stress, followed by the sine and sawtooth patterns. Please note that the stress voltage was chosen to produce lifetime relevant drifts. That is why also the crossings are of high interest for a circuit design as they reveal that for an early operation phase the sawtooth signal causes the maximum drift and the dAC pattern (fit) the smallest one.

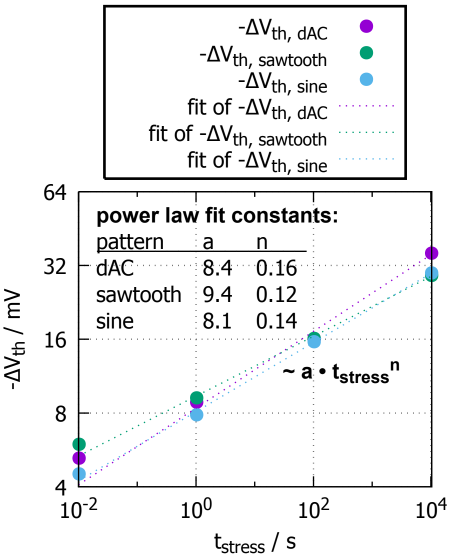


Figure 4.48.: The threshold voltage drift after NBTI stress with arbitrary gate signals shows that for long stress times AC stress with an offset is most critical ($n \approx 0.16$) followed by sine ($n \approx 0.14$) and sawtooth patterns ($n \approx 0.12$). Vice versa (fit), at short stress times the most critical drift is found for the sawtooth pattern.

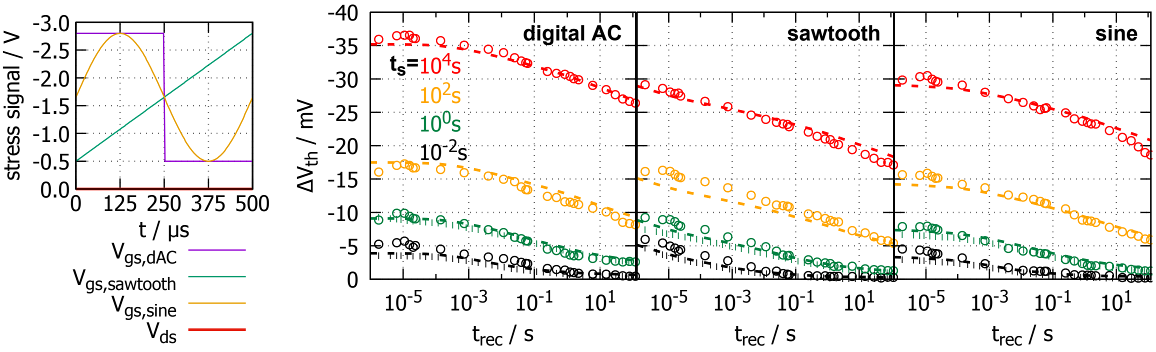


Figure 4.49.: Left: The analog stress MSM experiments periodically apply one of the following stress patterns: sine, sawtooth, digital AC. The stress voltage oscillates between the values $V_\mathrm{l} \leq -V_\mathrm{gs}(t) \leq V_\mathrm{h}$. **Right (three plots):** The MSM ΔV_th recovery curves after four stress durations 10^{-2} s, 1 s, 10^2 s and 10^4 s (bottom-up in the plots) show a good agreement of experimental data (symbols, reduced data density), TCAD results (thick dashes) and compact model results (solid line). Due to the numerical complexity, TCAD data so far are only available for the two shorter stress times. The difference between TCAD and compact model results is often hardly noticeable. Most importantly, the theoretical predictions and experimental measurements are in good agreement. The consistency of experimental and compact model data at large stress times validates the compact model's extrapolation method to long stress times also from a practical point of view. The present setup applies a stress frequency of 2 kHz, a temperature of 398 K and stress voltages between $V_\mathrm{l} = 0.5\,\mathrm{V} \approx -V_\mathrm{th}$ and $V_\mathrm{h} = 2.8\,\mathrm{V}$.

The right three plots of Fig. 4.49 show the comparison between the measurement data, the compact model [168] and the TCAD simulation results from the microscopic model (see Section 1.2.3). It is obvious that all three methods reproduce the data in the same way and only minor differences exist. Because the TCAD and the compact model simulation are based on the same input trap database their difference is minimal. The beginning of the measured recovery traces sometimes, but reproducible, shows a slight increase of the threshold voltage shift which cannot be reproduced by either model.

4.3.4. Duty factor dependence

While applying square wave stress signals which oscillate between $V_{\mathrm{l}} = 0\,\mathrm{V}$ and V_{h} (here $= -2.8\,\mathrm{V}$) the duty factor has a strong influence on the resulting drift. Following [157, 182] a so-called S-curve is expected as a result. Defect classes which have a much longer capture time compared to the emission time lead to the sharp rise between a duty cycle of 0% and 5%, whereas the opposite defect time configuration causes the high slope between 95% and 100%. The measured threshold voltage shift after $10^4\,\mathrm{s}$ NBTI stress is shown in Fig. 4.50 and is in accordance with the findings reported in [157, 182].

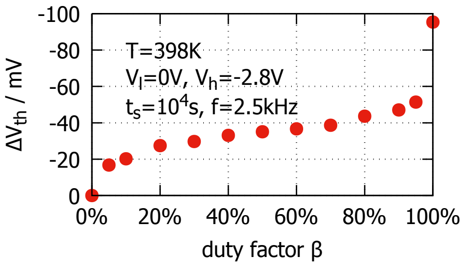


Figure 4.50.: The threshold voltage drift ($t_{\mathrm{str}} = 10^4\,\mathrm{s}$) depends strongly on the chosen duty factor β of the AC NBTI stress and follows an S-curve. Following [157, 182] this can be explained by a simplified picture of three defect classes, $\tau_e \ll \tau_c$ ($0\% < \beta < 5\%$), $\tau_e = \tau_c$ and $\tau_e \gg \tau_c$ ($95\% < \beta < 100\%$).

4.4. Comparison of used equipment

To compare the standard with the ultra-fast equipment the output characteristics at the saturation condition ($V_{\mathrm{ds}} = -1.5\,\mathrm{V}$) were measured for the core p-MOSFET with each instrument. As shown in Fig. 4.51 both show a high level of agreement for a wide range of typical threshold voltages. It should be noted that for the measurement with the standard equipment the “parasitic” source-bulk voltage caused by the chosen setpoint of the ultra-fast measurement equipment had to be applied directly to the device. This ensures a correct source-bulk offset correction and allows a quantitative comparison between both data sets.

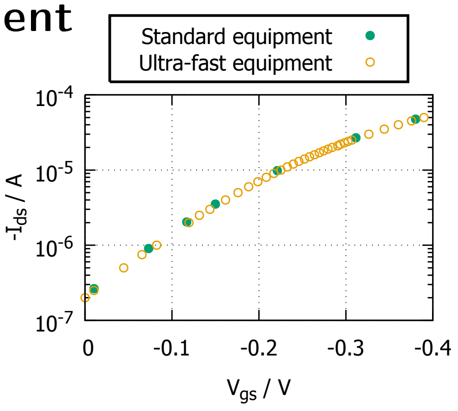


Figure 4.51.: The comparison between the data recorded by the standard and the ultra-fast equipment clearly shows that both are in a very good agreement. Shown is the drain-source current ($V_{\mathrm{ds}} = -1.5\,\mathrm{V}$) as a function of the gate-source voltage.

4.5. Recovery of core p-MOSFET HCD

4.5.1. “Pure” hot-carrier degradation condition

If the core p-MOSFET is stressed with HCD conditions for which the gate voltage is kept small enough to suppress NBTI, which is the case for $V_{\mathrm{gs}} \geq -0.525\,\mathrm{V}$, even long term recovery measurements show no HCD recovery for $T_{\mathrm{hotchuck}} \leq 463\,\mathrm{K}$ (see Fig. 4.52).

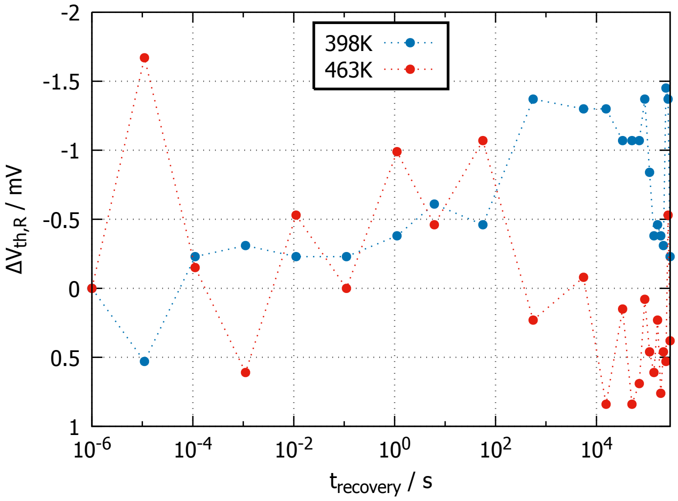


Figure 4.52.: Long term threshold voltage recovery signal ($\Delta V_{\mathrm{th,R}} = \Delta V_{\mathrm{th}}\,(t_{\mathrm{recovery}}) - \Delta V_{\mathrm{th}}\,(t_{\mathrm{recovery}} = 10^{-6}\,\mathrm{s})$) of the core p-MOSFET after $5 \times 10^{4}\,\mathrm{s}$ HCD stress, $V_{\mathrm{ds}} = -2.8\,\mathrm{V}$, $V_{\mathrm{gs}} = -0.525\,\mathrm{V}$. $\Delta V_{\mathrm{th}}\,(t_{\mathrm{recovery}} = 10^{-6}\,\mathrm{s}) = -42\,\mathrm{mV}$. Even at the maximum chuck temperature no recovery is observable for this device type.

4.5.2. The onset of NBTI

To examine the onset of NBTI during HCD additional transistors were stressed at 11 stress times ($10^{-6}\,\mathrm{s} \ldots 10^{4}\,\mathrm{s}$) and various gate- and drain-source voltage combinations. The recovery response to each condition is depicted in Fig. 4.53. Three different post-stress behaviors can be distinguished. Besides the group which shows no recovery, the change of the threshold voltage is below $\pm 0.5\,\mathrm{mV}$, there are two different types possible: At the highest $V_{\mathrm{gs}} = -1.5\,\mathrm{V}$ a recovery of the threshold voltage shift is observable while for the highest hot-carrier acceleration $V_{\mathrm{ds}} = -2.8\,\mathrm{V}$ and $V_{\mathrm{gs}} \geq -1.0\,\mathrm{V}$ post-stress aging up to $-2.5\,\mathrm{mV}$ is noticeable. Especially the CET maps of the data in Fig. 4.54 reveal that the higher the absolute drain-source voltage is chosen, compared to the gate-source voltage, the probability density of recoverable traps of one $\tau_{\mathrm{c}}, \tau_{\mathrm{e}}$ class decreases. From the color mapping of Fig. 4.54(d) one can also see that the post-stress aging effect starts to occur for stress times longer than 1 s. The observed post-stress aging has also been reported by [183] for n-MOSFETs. There, a rather small V_{gs} compared to V_{ds} led to the injection of holes into the oxide and the number of trapped holes was found to be related to the post-stress generated interface trap density. Because the measured threshold voltage change is much smaller compared to the data presented in [183] the creation of interface traps which is described there to increase logarithmically with time is difficult to observe in the present data and the experiment remains inconclusive.

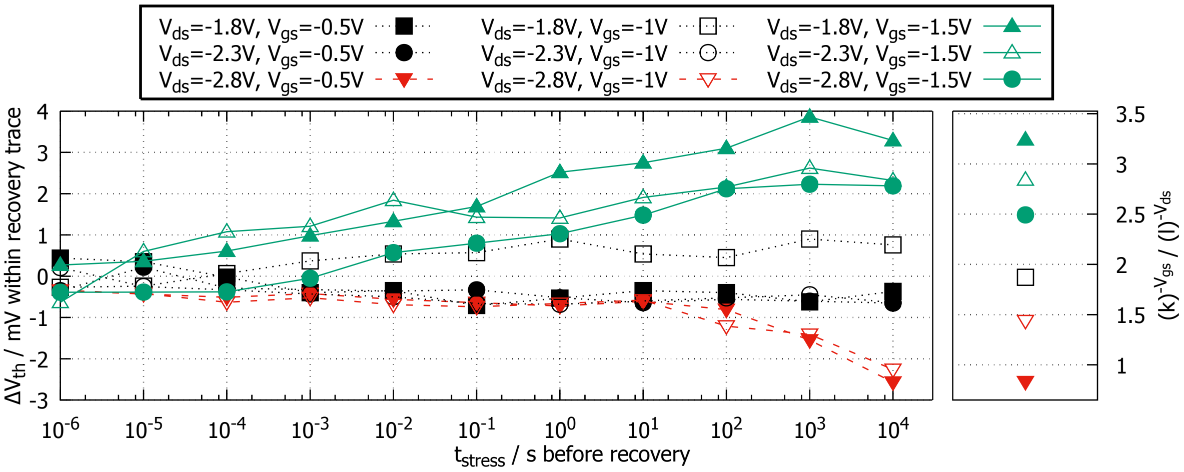
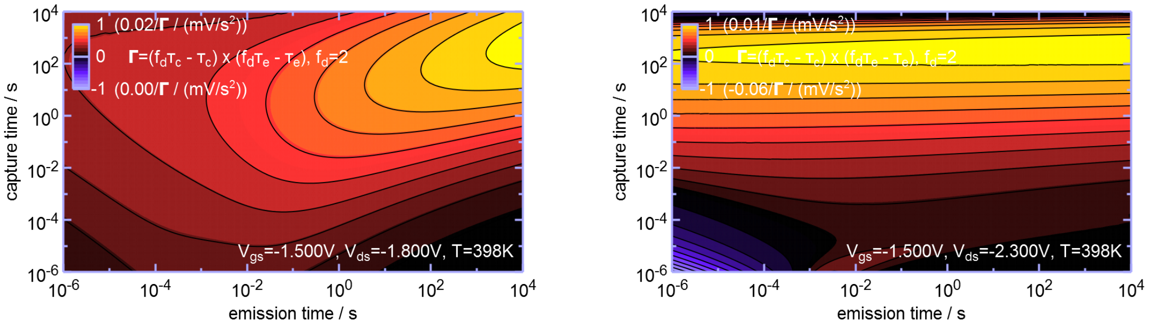
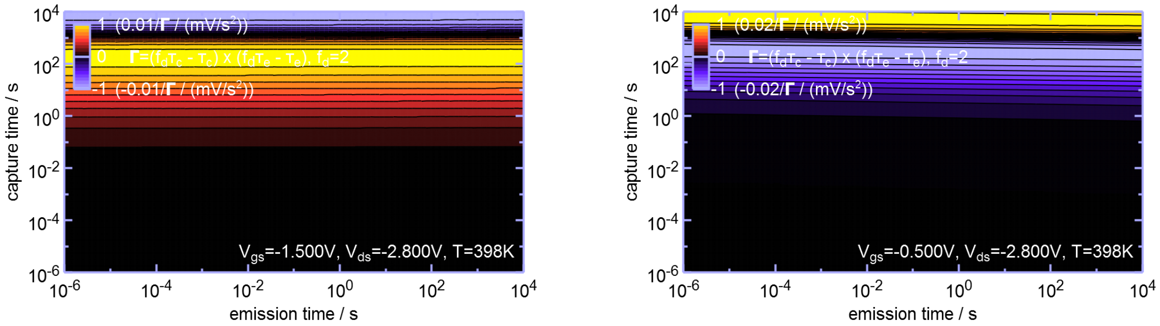


Figure 4.53.: Left: The recovery signal within a single threshold voltage trace after stress shows that there are three distinguishable groups of the post stress behavior. The first group which shows no recovery after stress (black) is stable within ±0.5 mV, the second group (red) tends to post-stress aging for long recovery times and the third group (green) with the highest $V_{\mathrm{gs}} = -1.5\,\mathrm{V}$ features a recognizable recovery signal. Although the dataset with $V_{\mathrm{ds}} = -1.8\,\mathrm{V}, V_{\mathrm{gs}} = -1\,\mathrm{V}$ recovers a little it cannot clearly be associated with the second or third group.
Right: The fit function $f = (k)^{-V_{\mathrm{gs}}} / (l)^{-V_{\mathrm{ds}}}$ with $k = 3.0$, $l = 1.3$, was fitted to the last recovery data points in order to weight the influence of HCD and NBTI (respectively the chosen stress voltages) on the behavior after stress.



(a) $V_{\mathrm{gs}} = -1.5\,\mathrm{V}$, $V_{\mathrm{ds}} = -1.8\,\mathrm{V}$. The maximum threshold voltage shift of a single recovery trace compared to the other recovery traces was $\Delta V_{\mathrm{th,max}} = 3.9\,\mathrm{mV}$.

(b) $V_{\mathrm{gs}} = -1.5\,\mathrm{V}$, $V_{\mathrm{ds}} = -2.3\,\mathrm{V}$, $\Delta V_{\mathrm{th,max}} = 2.6\,\mathrm{mV}$. The influence of HCD increases and less recovery is observable compared to (a).



(c) $V_{\mathrm{gs}} = -1.5\,\mathrm{V}$, $V_{\mathrm{ds}} = -2.8\,\mathrm{V}$, $\Delta V_{\mathrm{th,max}} = 2.2\,\mathrm{mV}$. The high drain-source voltage leads to strong HCD and almost no NBTI-like recovery is observable.

(d) $V_{\mathrm{gs}} = -0.5\,\mathrm{V}, V_{\mathrm{ds}} = -2.8\,\mathrm{V}$, $\Delta V_{\mathrm{th,max}} = -2.6\,\mathrm{mV}$. A high drain-source voltage in combination with a relatively low gate-source voltage leads to post-stress degradation (occupation of additional traps is colored blue).

Figure 4.54.: (a) The combination of a relatively low drain-source voltage compared to the gate-source voltage reduces HCD and emphasizes NBTI. Consequently, NBTI-like recovery is observable. (b) and (c) With increasing drain-source voltage HCD dominates aging and less recovery is observable. (d) For the pure HCD condition no recovery is observable but additional aging is taking place after stress. All fits were performed with the stress time dependent fit method (see Section 2.6.2).

4.6. Mixture of HCD and BTI

Only few publications about the mixture of HCD and NBTI exist (e.g. [184, 185]), although transistors very often face the combination of both aging mechanisms in the application. Therefore, in this section data of the mixture of NBTI and HCD will shown and discussed. The measurement results are in form of threshold voltage drift plots, recovery traces and CET maps including NBTI and HCD. All measurements were performed on core device class p-MOSFETs which were stressed using a 2-dimensional parameter space of gate and drain voltage combinations at elevated temperature (398 K). The chosen stress conditions include the homogeneous ($V_{\mathrm{ds}} = 0$) and inhomogeneous ($V_{\mathrm{ds}} \ll V_{\mathrm{nom}}$) NBTI case, the pure HCD ($V_{\mathrm{gs}} \geq -0.525\,\mathrm{V}$) case as well as the mixture of NBTI and HCD. The results clearly show that for increasing $V_{\mathrm{ds}} > V_{\mathrm{nom}}$, NBTI recovery becomes less severe and mainly the permanent degradation due to HCD remains after the end of stress. Furthermore, there is a drift minimum of NBTI observable for a specific V_{ds}. Using CET maps it is quite evident that for high stress times the probability density of emission becomes very small, whereas for shorter stress times there is a recoverable component noticeable.

Typically, devices are either tested under NBTI or HCD critical stress conditions and their interplay is not analyzed in detail. On the other hand for analog circuits this knowledge is urgently needed to estimate the degradation of the device during its life time. For modern technologies with small channel lengths it was shown that the drift caused by HCD can only be described correctly by considering both the influence of the drain and gate voltage [186] because different drift effects can either be facilitated by single or by multiple particle mechanisms. The results of this work underline that also for NBTI there are several drift effects possible and therefore also the influence of the gate and drain voltage needs to be taken into account. Because of that the analysis of only the homogeneous NBTI case ($V_{\mathrm{ds}} = 0$) is insufficient, as will be presented in the following, in addition to the homogeneous case, several inhomogeneous NBTI ($V_{\mathrm{ds}} \ll V_{\mathrm{nom}}$) conditions, the mixture of NBTI and HCD ($V_{\mathrm{gs}} > V_{\mathrm{nom}}$, $V_{\mathrm{ds}} > V_{\mathrm{nom}}$) and pure HCD ($V_{\mathrm{gs}} < V_{\mathrm{nom}}$) need to be taken into account. As well the findings can be directly implemented into a voltage based transistor degradation model to allow more realistic circuit simulations where a full V_{ds} and V_{gs} dependence and a detailed picture of the time behavior [187] are needed. This covers not only digital circuits where e.g. the hot-carrier degradation of inverters is dependent on the rise and fall times of the used gates but also analog circuits where comparators, current mirrors or I/O devices are implemented. Therefore only the analysis of the results of a 2-dimensional stress voltage matrix gives a complete picture of the drift behavior under various stress modes.

The threshold voltage drift (after a measurement delay of $10^{-6}\,\mathrm{s}$) for a given V_{gs}, V_{ds}

combination after 1.1×10^4 s at 398 K is presented in Fig. 4.55. A comparison of the drift results for different stress voltage regimes is shown in Fig. 4.56.

For homogeneous NBTI stress ($V_{ds} = 0$) the threshold voltage drift shows an increase with the absolute applied gate stress voltage. The maximum drift at $V_{gs} = -2.8$ V is 33% of the mixed HCD and NBTI stress condition ($V_{gs} = V_{ds} = -2.8$ V). For a gate voltage of -2.3 V the fraction is 14% and for $V_{gs} = -1.8$ V 7%.

The degradation results for inhomogeneous NBTI ($V_{ds} < V_{nom}$) show a drift minimum at ($V_{ds} = -0.5$ V) (Fig. 4.57). Interestingly, the drift minimum is always located at the same drain-source stress voltage and independent of the stress time.

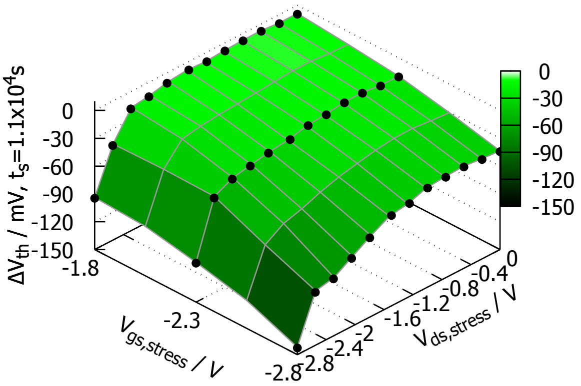


Figure 4.55.: Stress voltage matrix results showing the threshold voltage drift after 1.1×10^4 s at 398 K. With increasing V_{gs} NBTI becomes more severe and with increasing V_{ds} HCD becomes more severe. The maximum drift for each series of fixed V_{gs} is dominated by HCD. Symbols: measurement data, surface: interpolated

For the mixed stress condition $V_{gs} = V_{ds}$ the drift also shows an increase with the absolute applied gate stress voltage. The drift amount compared to $V_{gs} = V_{ds} = -2.8$ V

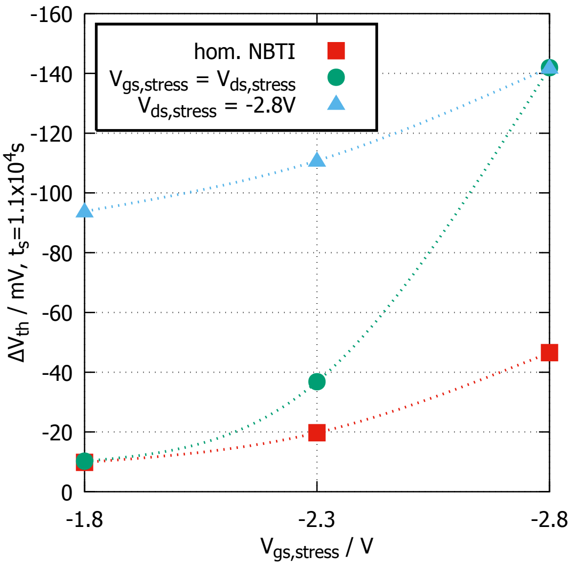


Figure 4.56.: Drift of different stress voltage regimes, homogeneous NBTI ($V_{ds} = 0$), NBTI mixed with HCD where $V_{ds} = V_{gs}$ and where $V_{ds} = -2.8$ V. The lowest degradation is observable for homogeneous NBTI (blue), drain stress voltages lead to higher degradation (green, red). Lines are guides to the eye.

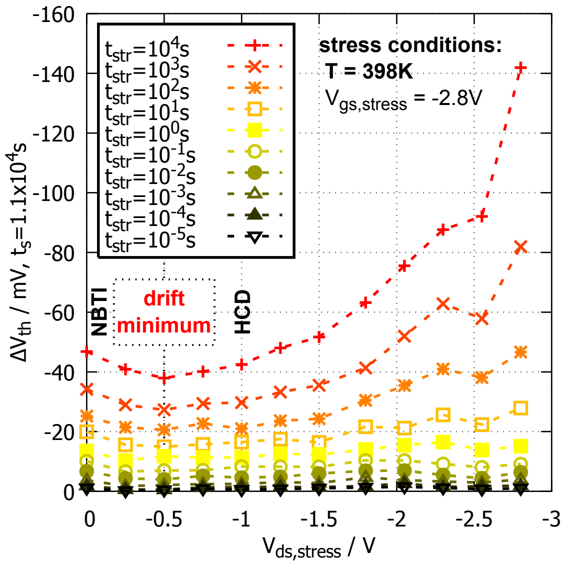
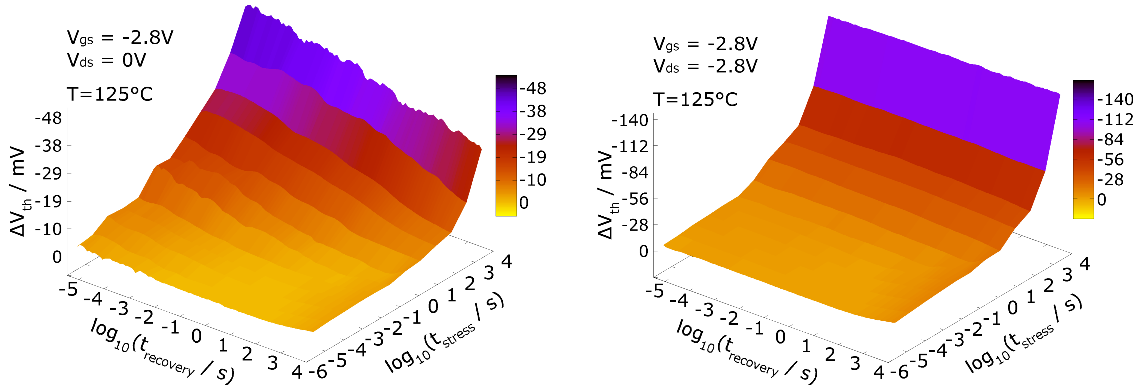


Figure 4.57.: Drift dependence for fixed $V_{gs} = -2.8$ V and varied $V_{ds} = (0 \ldots -2.8)$V. At $V_{gs} = -2.8$ V, $V_{ds} = -0.5$ V the interplay of NBTI and HCD leads to a drift minimum for all stress times. The measurements were performed on adjacent chips to eliminate the global drift variation.

is 26% at $V_{\mathrm{gs}} = V_{\mathrm{ds}} = -2.3\,\mathrm{V}$ and about 7% at $V_{\mathrm{gs}} = V_{\mathrm{ds}} = -1.8\,\mathrm{V}$.

The results for $V_{\mathrm{gs}} = -1.8 \ldots -2.8\,\mathrm{V}, V_{\mathrm{ds}} = -2.8\,\mathrm{V}$ show the highest degradation of the threshold voltage. Here the drift at $V_{\mathrm{gs}} = -1.8\,\mathrm{V}, V_{\mathrm{ds}} = -2.8\,\mathrm{V}$ is 66% and at $V_{\mathrm{gs}} = -2.3\,\mathrm{V}, V_{\mathrm{ds}} = -2.8\,\mathrm{V}$ 78% of the maximum drift which is about 142 mV.



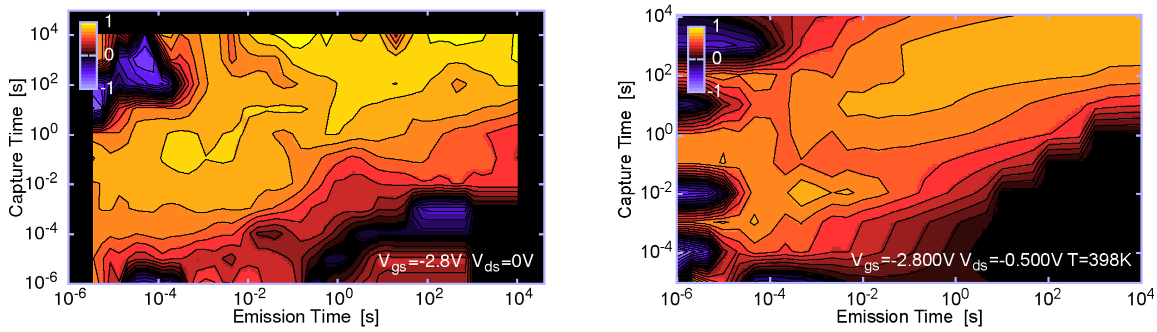
(a) homogeneous NBTI,$V_{\mathrm{gs}} = -2.8\,\mathrm{V}$, $V_{\mathrm{ds}} = 0\,\mathrm{V}$ (b) mix. NBTI and HCD,$V_{\mathrm{gs}} = V_{\mathrm{ds}} = -2.8\,\mathrm{V}$

Figure 4.58.: Three-dimensional plots of recovery traces for given stress times showing the strong influence of HCD on the drift and recovery behavior. (a) for homogeneous NBTI ($V_{\mathrm{ds}} = 0$) the recovery of the threshold voltage drift is clearly recognizable whereas (b) for high drain stress voltages the recovery is almost negligible.

The recovery traces of homogeneous NBTI show the well-known recovery behavior (Fig. 4.58(a)). The longer the stress time the longer the recovery of the threshold voltage degradation can be observed. At this stress condition the maximum change of the threshold voltage within one recovery trace can be found after the longest stress time of 10^4s which is about 32mV (Fig. 4.60). In addition to that the CET maps for homogeneous NBTI clearly show that the threshold voltage recovers from its degradation for all stress decades (Fig. 4.59(a)). The CET maps shown in this section were generated using the stress-time independent fit method (see Section 2.6.1).

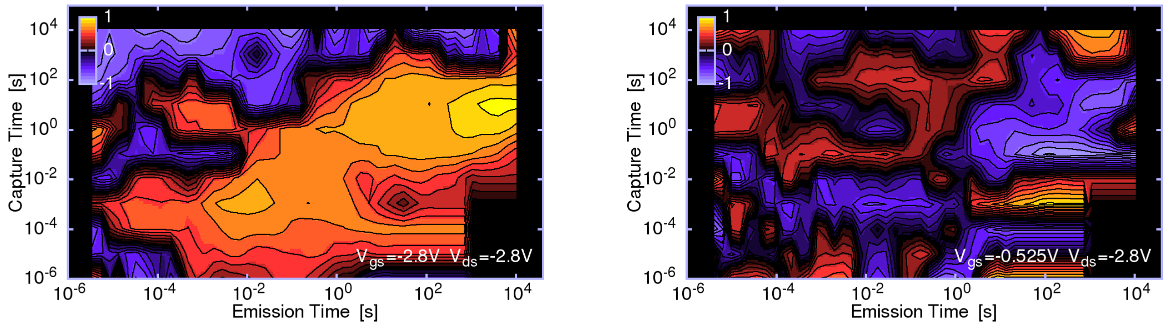
Interestingly, at the inhomogeneous NBTI stress voltage condition causing the observed drift minimum, the recoverable component is slightly shifted to longer emission and shorter capture times 4.59(b). This means that even at this condition the low drain-source voltage influences the recovery of NBTI. Taking the results of higher drain-source voltages into account this can be considered as the onset of HCD.

A further increase of the drain stress voltages leads to the mixture of NBTI and HCD. At this condition the amount of recoverable degradation diminishes with increasing stress time (Fig. 4.58(b)). For the last decades of stress time the recoverable component can be neglected. The details of the stress condition $V_{\mathrm{ds}} = V_{\mathrm{gs}}$ can be found in the corresponding CET map (Fig. 4.59(c)). Here the recoverable component is present for short stress times but vanishes for long stress times. In the case of pure HCD (Fig. 4.59(d)) the recoverable



(a) Homogeneous NBTI, $V_{\mathrm{gs}} = -2.8\,\mathrm{V}$, $V_{\mathrm{ds}} = 0\,\mathrm{V}$. The maximum threshold voltage shift of a single recovery trace compared to the other recovery traces was $\Delta V_{\mathrm{th,max}} = 32\mathrm{mV}$.

(b) Inhomogeneous NBTI, $V_{\mathrm{gs}} = -2.8\,\mathrm{V}$, $V_{\mathrm{ds}} = -0.5\,\mathrm{V}$, at drift minimum condition (Fig. 4.57), $\Delta V_{\mathrm{th,max}} = 24\mathrm{mV}$.



(c) Mixture of NBTI and HCD, $V_{\mathrm{gs}} = V_{\mathrm{ds}} = -2.8\,\mathrm{V}$, $\Delta V_{\mathrm{th,max}} = 17\mathrm{mV}$. Compared to homogeneous NBTI the recoverable component is shifted to shorter capture times and longer emission times.

(d) Pure HCD, $V_{\mathrm{gs}} = -0.525\,\mathrm{V}, V_{\mathrm{ds}} = -2.8\,\mathrm{V}$, $\Delta V_{\mathrm{th,max}} = 2\mathrm{mV}$ is very small compared to homogeneous NBTI and the mixture of NBTI and HCD.

Figure 4.59.: (a), (b), (c) continuous capture emission time maps for different stress conditions showing the influence of V_{ds} on the recoverable component and (d) the absence of a recoverable component for a pure HCD stress condition. The fits were performed following the stress time independent fit method (see Section 2.6.1).

component is absent with respect to the very low change of the threshold voltage over time.

Because the ultra-fast measurement equipment allows a very short measurement delay for the determination of the threshold voltage after stress, the drift of this parameter was used for the comparison among the presented stress conditions. The device-to-device variation of the measured chips is below 5% regarding the drift at the same stress voltage condition. Additional data including the whole gate-source voltage range have been recorded to provide a comparison between the threshold voltage drift and the relative linear and saturation drain-source current drift which is depicted in Fig. 4.61. As these measurement results show an apparent degradation of the threshold voltage not only for NBTI but also for a wide gate-source voltage range of HCD conditions this is a perfect key parameter allowing a comparison of the different stress voltage regimes.

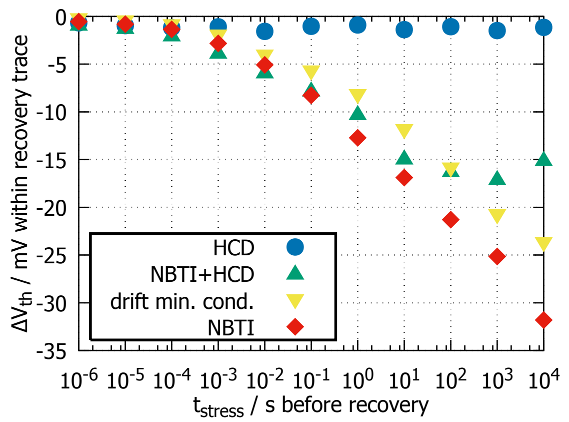


Figure 4.60.: Comparison of the recovery signal after a given stress time for HCD (blue), NBTI (red), their mixture (green) and the drift minimum condition (yellow, Fig. 4.57). The change of the threshold voltage within a recovery trace after stress is strongly dependent on the drain voltage.

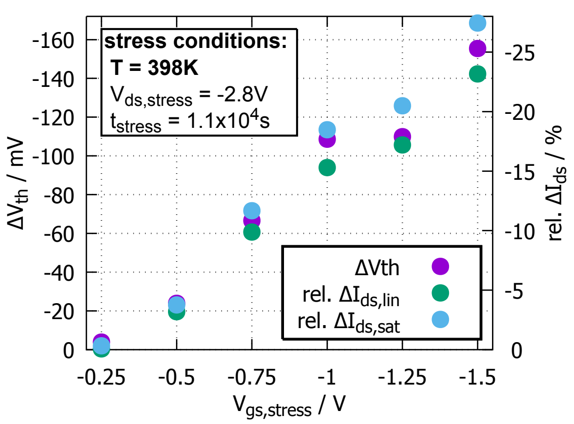


Figure 4.61.: Comparison of the drift of important device parameters after 1.1 × 10^4s hot-carrier degradation. All parameters (purple: ΔV_{th}, green: $\Delta I_{\text{ds,lin}}$, blue: $\Delta I_{\text{ds,sat}}$) feature an apparent drift over a wide gate-source voltage range thus showing their relevance for HCD.

Discussion of experimental results

The observed recovery of the threshold voltage as well as the non-recoverable degradation component induced by HCD of the aged transistors are in good agreement with the literature [163, 186, 188]. As one compares the recovery traces shown in Fig. 4.58(a) and 4.58(b) with the corresponding CET maps (Fig. 4.59(a), 4.59(c)) the recoverable component of the threshold voltage for the mixture of NBTI and HCD ($V_{\text{gs}} = V_{\text{ds}} = -2.8$V) can nicely be observed in the CET map. For all stress decades above 1s one has to consider the offset of the following recovery traces which is caused by non-discharged oxide traps and interface traps with recovery times above 10^4s, contributing to a quasi-permanent degradation component [189].

By applying low drain voltages to the device under test the oxide field becomes inhomogeneous. The maximum electric field ($E_{\text{max}} \approx V_{\text{gs}}/t_{\text{ox}}$) over the oxide can be found at the source side whereas the minimum ($E_{\text{min}} \approx (V_{\text{gs}} - V_{\text{ds}})/t_{\text{ox}}$) is located near the drain. Therefore the maximum NBTI-related degradation and recovery is expected at the source side.

The drain voltage dependence of the threshold voltage drift shows a minimum for -0.5V and is caused by the interplay of two degradation mechanisms. On the one hand the NBTI voltage condition causes holes to interact with the defects in the gate oxide and its interface [189]. On the other hand the HCD voltage condition enhances multiple particle mechanisms for the investigated short channel (l = 100nm) device. Impact ionization [171], Auger recombination [172], electron-phonon [173, 174] and electron-electron scattering [175, 176] lead to the creation of interface states [40] causing a positive charge near the drain, which perturbs the electrostatics of the transistor, and resulting in a drift of e.g. the threshold voltage. The minimum of the threshold voltage degradation which is found for $V_{\mathrm{gs}} = -2.8\mathrm{V}, V_{\mathrm{ds}} = -0.5\mathrm{V}$ is related to the energy distribution of the carriers in the channel which can cause either NBTI or HCD (Fig. 4.57). This effect has already been reported in the literature for older technologies and, when the different channel lengths are considered, is in good agreement with those results [190, 191].

For high drain stress voltages HCD has a considerable impact on the permanent component of the threshold voltage shift and NBTI as well as its recovery become less severe. This can nicely be seen in the drift data after $1.1 \times 10^4\mathrm{s}$ (Fig. 4.57), the recovery traces for each stress decade (Fig. 4.58(b)) and the CET map (Fig. 4.59(c)) for the case of mixed hot-carrier and NBTI degradation. This observation has also been reported for single oxide defects in [185]. For $V_{\mathrm{ds}} < -1.5\,\mathrm{V}$ a group of defects captures and emits their charge during stress (prior recovery, $\tau_{\mathrm{e}} \ll \tau_{\mathrm{c}}$) and therefore the resulting threshold voltage and its recovery signal will be unaffected by such defects [185]. In addition to that a shift of the recoverable component to about 100× shorter stress and 100× longer recovery times is observable if one compares the CET maps of homogeneous NBTI (Fig. 4.59(a)) with mixed NBTI and HCD (Fig. 4.59(c)).

If one chooses a low gate voltage to suppress the NBTI degradation but applies a high drain voltage to emphasize pure hot-carrier induced degradation no recoverable component can be found in the CET map (Fig. 4.59(d)). The absence of a recoverable component underlines the fact that for the investigated thin-oxide transistor (in contrast to a thick-oxide power device [192]) a recovery of the threshold voltage has only to be considered for NBTI degradation.

To provide a comparison between the threshold voltage and other important device parameters (e.g. $I_{\mathrm{ds,lin}}$, $I_{\mathrm{ds,sat}}$) additional data have been recorded (Fig. 4.61). These data show that for the investigated technology the threshold voltage shift is a perfect key parameter to allow a comparison among the different stress voltage regimes as it shows an apparent degradation not only for NBTI but also for a wide range of gate voltages under HCD stress conditions.

As discussed in [193] the interaction between HCD and NBTI has to be determined to get a more realistic reliability assessment. Along these lines the additional information about the recovery behavior of a device obtained from CET maps can be used to simulate the aging of circuits more accurately compared to the use of a superposition of compact models for NBTI and hot-carrier degradation. Following this approach it is ensured that the drift is neither over- nor underestimated by the simulation because all relevant voltage conditions are covered by the measurement data.

4.6.1. Influence of the stress temperature on CET maps of inhomogeneous NBTI

In addition to the results depicted in Fig. 4.44 where the differential CET map between 398 K to 443 K was calculated for homogeneous NBTI at $V_{\mathrm{gs,stress}} = -2.8\,\mathrm{V}$, $V_{\mathrm{ds,stress}} = 0\,\mathrm{V}$ also CET maps of inhomogeneous NBTI ($V_{\mathrm{ds,stress}} < 0\,\mathrm{V}$) and the mixture of NBTI and HCD ($V_{\mathrm{ds,stress}} = -1.5\,\mathrm{V}$) where recorded at 398 K and 443 K. At the drift minimum condition ($V_{\mathrm{ds,stress}} = -0.5\,\mathrm{V}$) where the recoverable component is slightly shifted to longer emission and shorter capture times (4.59(b)) an increase of the temperature further intensifies this effect. The capture time of the maximum recovery probability density (yellow) is decreased by a factor of about 10 while the emission time is increased by a factor of about 100 (Fig. 4.62) compared to homogeneous NBTI (Fig. 4.43 (left)). Once the drain-source voltage equals the nominal device voltage of $-1.5\,\mathrm{V}$ a strong decrease of the capture time as well as a prominent broadening at 443 K of the maximum of the recoverable component over the whole recovery time scale is observable (Fig. 4.63). An increase of the drain-source voltage to the condition $V_{\mathrm{gs,stress}} = V_{\mathrm{ds,stress}} = -2.8\,\mathrm{V}$ where NBTI and HCD are emphasized together leads to a decrease of the capture time of the maximum probability density to about 10 s at 398 K (compared to 10^4 s for homogeneous NBTI). An increase of the temperature reduces the probability density for traps with short capture times and decreases the slope $s = \Delta\tau_{\mathrm{c}}/\Delta\tau_{\mathrm{e}}$ of the maximum probability density. In general, the slope has its maximum for homogeneous NBTI and reduces with increasing $V_{\mathrm{ds,stress}}$.

Note that all CET maps shown in this section were recorded with a readout voltage of $V_{\mathrm{ds,read}} = -0.1\,\mathrm{V}$. As presented in Fig. 4.37 the influence of the chosen $V_{\mathrm{ds,read}}$ has to be taken into account if the results are compared with Fig. 2.20 where the same stress conditions were chosen but $V_{\mathrm{ds,read}}$ was set to $-1.5\,\mathrm{V}$ resulting in a different appearance of both CET maps.

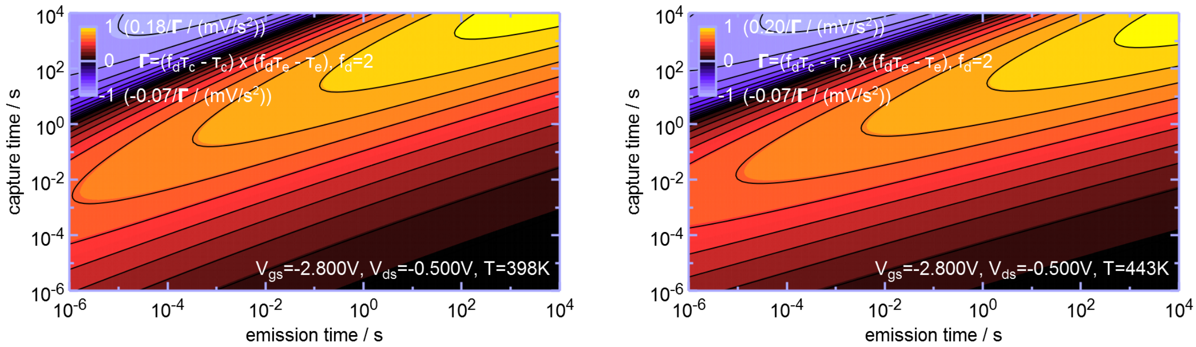


Figure 4.62.: CET maps of inhomogenous NBTI stress, $V_{\mathrm{gs,stress}} = -2.8\,\mathrm{V}$, $V_{\mathrm{ds,stress}} = -0.5\,\mathrm{V}$ at 398 K and 443 K.

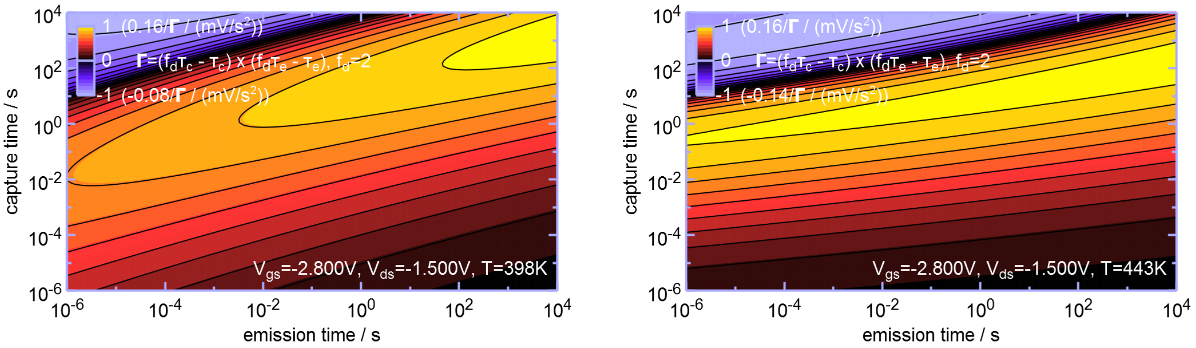


Figure 4.63.: CET maps of inhomogenous NBTI stress, $V_{\mathrm{gs,stress}} = -2.8\,\mathrm{V}$, $V_{\mathrm{ds,stress}} = -1.5\,\mathrm{V}$ at 398 K and 443 K.

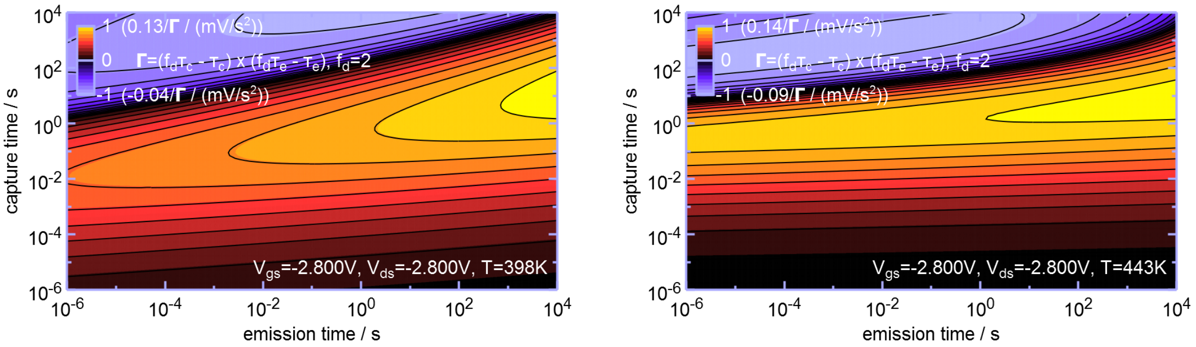


Figure 4.64.: CET maps of combined NBTI and HCD stress, $V_{\mathrm{gs,stress}} = -2.8\,\mathrm{V}$, $V_{\mathrm{ds,stress}} = -2.8\,\mathrm{V}$ at 398 K and 443 K.

4.6.2. Interaction of HCD and NBTI

In order do investigate the interaction of HCD and NBTI 16 devices of adjacent chips were stressed and measured at 443 K following the sequence depicted in Fig. 4.65. During a long-term phase readouts of the threshold voltage of each transistor were performed at various times to monitor the recovery behavior. Finally, the whole wafer was stored for 19 days in a furnace at 573 K which is equivalent to a recovery time of $1.76 \times 10^{12}\,\mathrm{s}$ (following Eq. 4.2) at 443 K and subsequently the threshold voltage values of all devices were measured. Unfortunately one p-MOSFET of each group had a poor contact between the pads and the probe card thus leaving 7 transistors for the statistics of either HCD followed by NBTI or vice versa.

After $10^4\,\mathrm{s}$ HCD stress the first group experiences a threshold voltage shift of $(-40.1 \pm 6.8)\,\mathrm{mV}$. The following NBTI stress leads to a maximum drift of $(-169.5 \pm 5.8)\,\mathrm{mV}$ and recovers within $10^4\,\mathrm{s}$ to $\Delta V_{\mathrm{th}} = (-97.8 \pm 4.7)\,\mathrm{mV}$. Interestingly, the slope of the recovery after $10^4\,\mathrm{s}$ changes to a higher value although the devices were not contacted to the measurement equipment in-between the single readouts ($V_{\mathrm{gs}} = 0\,\mathrm{V} > V_{\mathrm{gs,read}} = V_{\mathrm{th}} \approx -0.5\,\mathrm{V}$). This means that for a more positive gate bias the recovery is accelerated which can be explained by the results of the V_{r} dependence presented in Fig. 4.38 and Fig. 4.39.

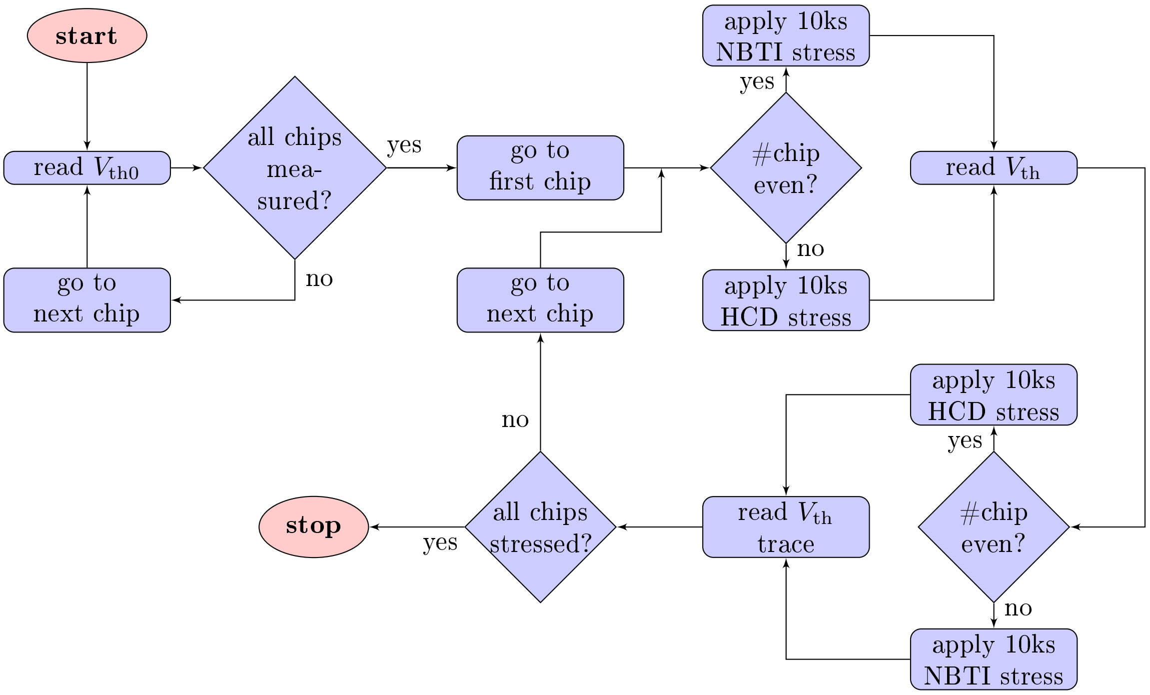


Figure 4.65.: The measure-stress-measure sequence for the determination of the interaction of NBTI and HCD including both stress and recovery. After the readout of the fresh V_{th} each device is stressed under NBTI(/HCD) conditions, an intermediate readout is performed and then HCD(/NBTI) is applied to the transistor. Finally a 10^4 s long recovery trace is recorded.

The second group shows a drift of $(-139.9 \pm 5.9)\,\mathrm{mV}$ after 10^4 s NBTI stress which is decreased by $-64.9\,\mathrm{mV}$ to $(-75.0 \pm 5.9)\,\mathrm{mV}$ after 10^4 s HCD stress. This reveals that during HCD a proportion of the via NBTI trapped charges can get de-trapped which leads to a recovery of NBTI. The de-trapping of the remaining charges attributed to NBTI is observable after 10^4 s and shows a smaller recovery slope compared to the first group.

Comparing both experiments it is obvious that NBTI can solely cause a threshold voltage drift of about $-140\,\mathrm{mV}$, with preceding HCD this reduces to $-170\,\mathrm{mV}+40\,\mathrm{mV} \approx -130\,\mathrm{mV}$. Because HCD is most severe at the drain side of the gate oxide (see Fig. 1.10, Fig. 1.20) only a small part of the whole oxide area is impacted by that degradation mechanism. In contrast to that NBTI affects the whole channel-oxide interface area. Consequently, only the defects (hydrogen bridges and hydroxyl E' centers) and Si-H bonds which are located in the region of influence of NBTI as well as HCD can participate in the interaction of both degradation mechanisms. Once HCD causes the breakage of a Si-H bond in such an area the remaining silicon dangling bond can trap a charge thus contributing to HCD. The HCD caused Si-H bond breakage and the following charge capture increase the threshold voltage ($\Delta V_{\mathrm{th}}\,(10^4\,\mathrm{s}) = -40\,\mathrm{mV}$) of the device. As a result of that fewer locations for NBTI-related charge trapping compared to the unstressed transistor will be present after HCD. Therefore, subsequent NBTI stress will show lower NBTI-related degradation $\Delta V_{\mathrm{th,NBTI*}}$. In fact, the experimental data clearly show that this is the case and the additional NBTI degradation is 10 mV lower. As a result of that one has to contemplate that the HCD pre-conditioned charge trapping sites $((\Delta V_{\mathrm{th,NBTI*}} - \Delta V_{\mathrm{th,NBTI}}) / \Delta V_{\mathrm{th,NBTI}} \approx 7\%)$ are not accessible to NBTI related degradation anymore.

If NBTI is applied prior to HCD (NBTI-HCD) to the device under test the recovery during HCD is about 7 mV lower compared to the recovery after the sequence HCD, NBTI (HCD-NBTI). For comparison, the recovery of a 10^4 s NBTI ($V_{\mathrm{gs}} = -2.8\,\mathrm{V}, T = 443\,\mathrm{K}$) stressed p-MOSFET after 10^4 s is about 85 mV. In summary, not only the degradation but also the recovery of NBTI is reduced by the influence of HCD.

For two and a half months the threshold voltage recovery was monitored at 443 K. During that phase the difference of the mean threshold voltage drift of the two groups reduced to about 5 mV and only little NBTI related drift remained (dashed-dotted line at $V_{\mathrm{th,HCD}} = -40\,\mathrm{mV}$). To further recover the transistors the wafer was stored in a furnace at 573 K for 19 days. Following Eq. 4.2 this is equivalent to a recovery time of 1.76×10^{12} s at 443 K. After that bake phase both groups had the same mean threshold voltage drift while the standard deviation of the HCD-NBTI group is only about 28% of that of the NBTI-HCD group. The remaining difference of 0.06 mV is negligible with respect to the standard deviation of 3.7 mV.

It is assumed that during the bake the recovery of HCD (similar to [192]) and the permanent component of NBTI take place. Following [155] (see also Section 1.2.5) hydrogen is released during NBTI stress at the gate side and quickly migrates to the channel where it can be trapped by new sites which were previously not available at $V_{\mathrm{gs}} = V_{\mathrm{th}}$. These trapped H^0 build up the permanent component of NBTI. Due to the high diffusivity of hydrogen [194] the process is reaction-limited[156]. The bake at $V_{\mathrm{gs}} = 0\,\mathrm{V}$ has two effects on the trapped hydrogen. On the one hand the high temperature will recover some part of the permanent component but on the other hand at the apparent gate bias also hydrogen originating from the gate sided reservoir (barrier of 2.5 eV, e.g. Si-H bonds) will get trapped by the recently vacated H-trapping sites. This leads to a recovery voltage dependent value of the permanent component. If a positive gate bias was chosen during the bake phase a larger amount of the permanent component had recovered resulting in an even smaller threshold voltage drift. The remaining threshold voltage drift of 19 mV is comparable to the results presented in [155].

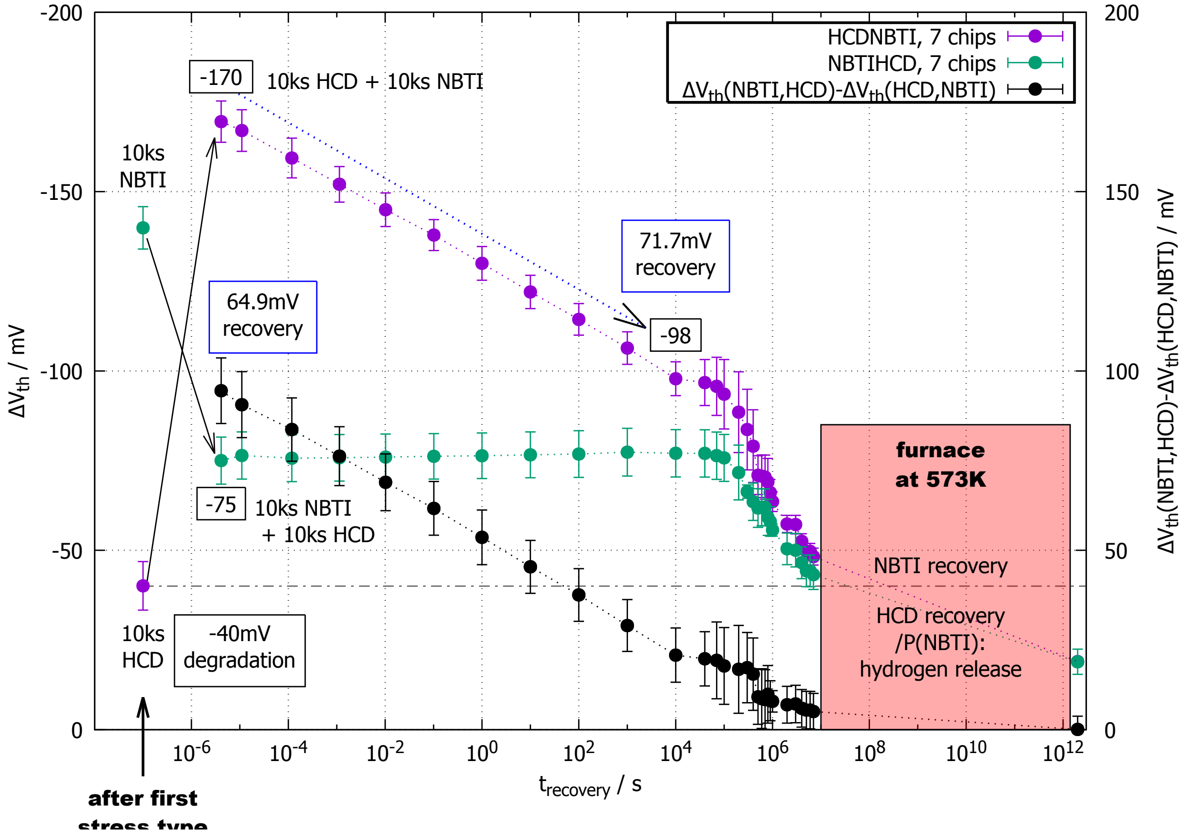


Figure 4.66.: This graph shows the degradation and recovery information of 14 devices including their statistics. Seven devices were first stressed at HCD conditions ($V_{\mathrm{ds}} = -2.8\,\mathrm{V}, V_{\mathrm{gs}} = -0.525\,\mathrm{V}$) and after an intermediate readout at NBTI conditions ($V_{\mathrm{gs}} = -2.8\,\mathrm{V}, V_{\mathrm{ds}} = 0\,\mathrm{V}$). The other seven devices were treated at the same conditions but with HCD and NBTI sequences reversed. After that the recovery trace was recorded over 7×10^6 s. NBTI causes about -140 mV degradation. After HCD this offset is reduced to $-170\,\mathrm{mV} + 40\,\mathrm{mV} \approx -130\,\mathrm{mV}$. About 1 µs after both stress types the difference between the recovery traces (black) is about 90 mV and decreases with the recovery time. At 7×10^6 s this difference reduces to 5 mV. To accelerate the recovery after 7×10^6 s the wafer was stored in a furnace at 573K for almost 19 days. The calculated effective recovery time is about 1.76×10^{12} s (following Eq. 4.2) and used for the last two data points of the plot. In addition to the expected NBTI recovery HCD seems to recover at such high temperatures as well as hydrogen might be released at the gate side (see Section 1.2.5). The final difference between both experiments (0.06 mV) is negligible with respect to the standard deviation of 3.7 mV.

4.7. Variability of NBTI (II)

The distribution of the fresh threshold voltage as well as the recoverable (R) and permanent (P) part of the threshold voltage drift are needed for the statistical information of mean threshold voltage drift models. The variation of the threshold voltage of unstressed transistors of a full wafer is depicted in Fig. 4.67. The mean threshold voltage is $\overline{V_{\mathrm{th0}}} = -524.7\,\mathrm{mV}$ and has a standard deviation of $\sigma = 8.6\,\mathrm{mV}$ which is a typical value of this production quality hardware. In order to extract the standard deviation of R and P 40 chips of two wafers were stressed with homogeneous NBTI conditions ($T = 398\,\mathrm{K}, V_{\mathrm{gs}} = -2.8\,\mathrm{V}$). 10 CET map data sets from the transistors of the first wafer were recorded (see also Fig. 4.45). 30 chips of the second wafer were stressed for $10^3\,\mathrm{s}$, a single readout was performed to get the information about R, then an accumulative pulse was applied for 10 s (red and green chips) and 100 s (shaded chips) to discharge the quickly recoverable traps followed by a recovery measurement of 10 s. Finally the threshold voltage recovery was monitored between $9 \times 10^3\,\mathrm{s}$ and $10^4\,\mathrm{s}$. Fig. 4.68 shows the location of the measured chips.

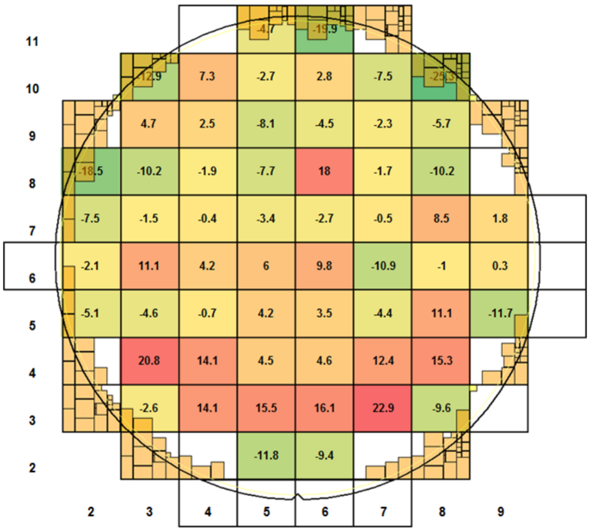


Figure 4.67.: The production quality wafers used for the experiments in this thesis show a very narrow distribution of the used threshold voltage of the unstressed p-MOSFETs with $\overline{V_{\mathrm{th0}}} = -524.7\,\mathrm{mV}$, $\sigma = 8.6\,\mathrm{mV}$.

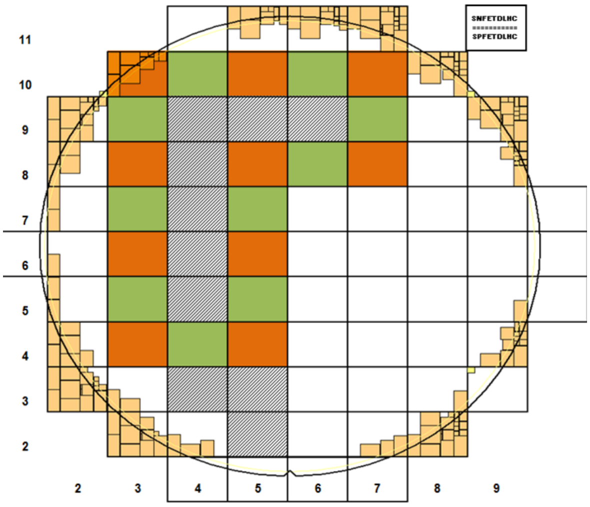


Figure 4.68.: The chosen chip subsets (red, green, shaded) for the statistical analysis of the recoverable ("R") and permanent ("P") component of NBTI. Both $\overline{V_{\mathrm{th0}}} = -524.0\,\mathrm{mV}$ and $\sigma = 8.8\,\mathrm{mV}$ of the chosen subset are comparable to the statistics of the whole wafer. **Red:** For 10 chips V_{th} was measured directly after NBTI stress. For 20 chips V_{th} was measured after stress and after an accumulative pulse ($V_{\mathrm{gs}} = 1\,\mathrm{V}$) of either 10 s (**green**) or 100 s (**shaded**) was applied to the gate terminal to discharge the quickly recoverable traps.

The data of the 10 samples without an accumulative pulse after a stress time of $10^3\,\mathrm{s}$ is presented in Fig. 4.69 and shows the typical recovery behavior after NBTI stress. The variability of R ($\sigma_\mathrm{R} = 2.9\,\mathrm{mV}$) is higher than that of P ($\sigma_\mathrm{P} = 1.3\,\mathrm{mV}$) which is in accordance with the results depicted in Fig. 4.46 where the variability of P ($t_\mathrm{r} = 10^4\,\mathrm{s}$, green data set) shows an increase with longer stress times. For $t_\mathrm{s} = 10^3\,\mathrm{s}$ σ_P is smaller than σ_R but for $t_\mathrm{s} = 10^4\,\mathrm{s}$ it superimposes the variability of R ($t_\mathrm{r} = 4.1 \times 10^{-6}\,\mathrm{s}$, red data set). The steeper slope of the recovery between $t_\mathrm{r} = 10\,\mathrm{s}$ and $10^4\,\mathrm{s}$ is caused by the discharge ($V_\mathrm{gs} = 0$) of the gate terminal before the floating phase during which the transistor was without any terminal contact. Following Fig. 4.39 the recovery at $V_\mathrm{gs} = 0\,\mathrm{V}$ is accelerated by a factor of five compared to the typical recovery voltage which is $V_\mathrm{gs} = V_\mathrm{th} \leq -0.5\,\mathrm{V}$. A comparable acceleration of the threshold voltage recovery can also be identified here.

The single information of the initial and degraded threshold voltage of each measured transistor is depicted in Fig. 4.70. To visualize the different spread of both quantities their average values superimpose and the scaling of both ordinates ($\Delta V_\mathrm{th}\,(R)$, $\Delta V_\mathrm{th}\,(P)$) is set to an equal range.

As shown in Fig. 4.71 the permanent component of the threshold voltage drift which is left after an accumulative pulse of 10 s and 100 s first increases slightly after stress until a recovery time of about 1 s and then decreases a little again. Interestingly, the variability of the two groups with different pulse lengths shows a different evolution of the variability. For a pulse length of 10 s the variability of P decreases between $t_\mathrm{r} = 10^{-1}\,\mathrm{s}$ and 10 s. In contrast to that the 100 s pulse data exhibits a slight increase about 0.4 mV of the variability of P over the recorded recovery period until $t_\mathrm{r} = 10\,\mathrm{s}$. After the $10^4\,\mathrm{s}$ long floating phase an additional recovery of P of about 3.5 mV is observable and the remaining variability of the permanent component at $t_\mathrm{r} = 10^4\,\mathrm{s}$ is about 1.3 mV.

Compared to Fig. 4.40 it is obvious that the amplitude of the accumulative pulse has a strong impact on the discharge of the quickly recoverable traps. A 10 s long pulse at 0.5 V after $10^4\,\mathrm{s}$ NBTI stress ($\Delta V_\mathrm{th} = -81\,\mathrm{mV}$) leads to a threshold voltage shift of about $-32\,\mathrm{mV}$ and recovers $(-81\,\mathrm{mV} + 32\,\mathrm{mV})\,/{-81}\,\mathrm{mV} \approx 60\%$ of the traps. A pulse of 10 s at 1 V results after $10^3\,\mathrm{s}$ NBTI stress ($\Delta V_\mathrm{th} \approx -75\,\mathrm{mV}$) in a threshold voltage shift of about $-15\,\mathrm{mV}$ and a recovery of $(-75\,\mathrm{mV} + 15\,\mathrm{mV})\,/{-75}\,\mathrm{mV} = 80\%$ of the charged traps.

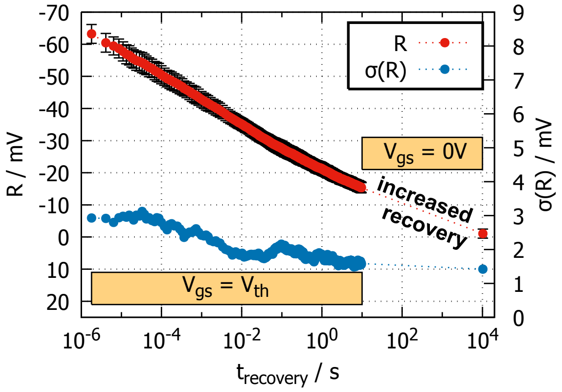


Figure 4.69.: The recoverable component of NBTI decreases with the recovery time, P ($\approx -14.5\,\text{mV}$) subtracted. During the non-contact phase ($V_{\text{gs}} = 0\,\text{V}$) which was applied for 10^4 s the slope of the recovery steepens.

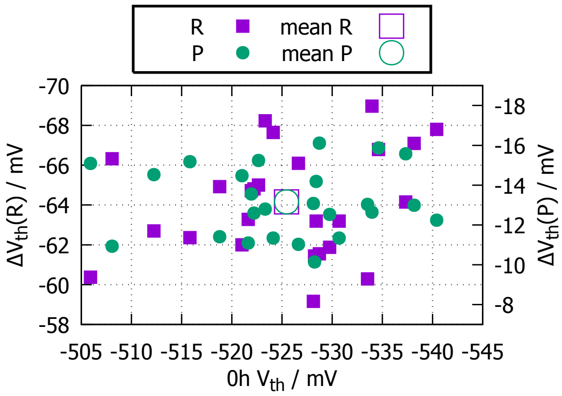


Figure 4.70.: The variability of the permanent threshold voltage drift is smaller ($\sigma_{\text{P}} = 1.3\,\text{mV}$) compared to that of the recoverable threshold voltage drift ($\sigma_{\text{R}} = 2.9\,\text{mV}$) for $t_{\text{s}} = 10^3$ s.

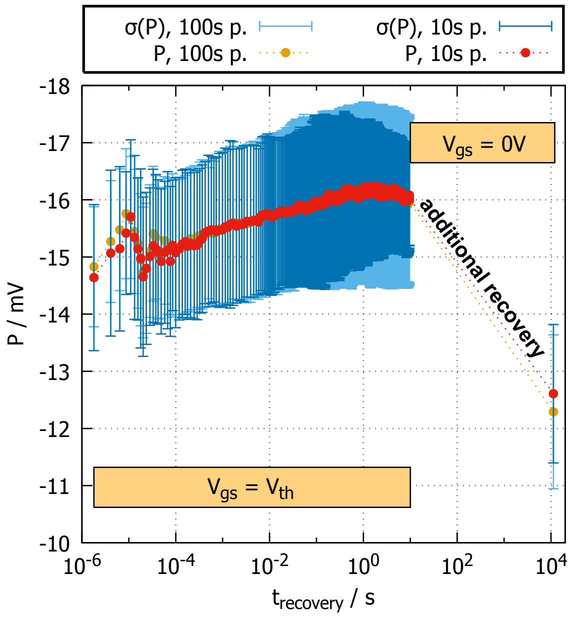


Figure 4.71.: The permanent component (red and orange symbols) of NBTI increases with the recovery time between 10^{-6} s and 3 s. After that it settles to a value of about $-16\,\text{mV}$. After the transistor was not contacted for 10^4 s an additional recovery of the permanent component of about 3.5 mV is observable. The data of the 10 s and 100 s long accumulation pulse are similar. Only their standard deviation shows a different behavior over the recovery time. Compared to the 100 s pulse the distribution of the 10 s narrows between 10^{-1} s and 10 s recovery. At the last readout they are quite similar again.

4.8. Evaluation of CET map parameters

The stress time dependent fit method of CET maps (see Section 2.6.2) allows a qualitative and quantitative interpretation of the measured stress and recovery data after NBTI, HCD and their combination. Because the calculated CET maps show many interesting features the influence of the experimental stress parameters like the gate-source and drain-source voltage combination on the fitted parameters can be considered.

The impact of the gate-source voltage for homogeneous NBTI on the eight fit parameters (e_1 to f_4) and the resulting S-shape curve parameters ($\Delta V_{\mathrm{th},\infty}$, $\Delta V_{\mathrm{th},0}$, c, d) is depicted in Fig. 4.72. The fit parameters e_1, e_2 and e_4 depend strongly on the set gate-source stress, f_1 and f_2 seem to have a weak correlation with $V_{\mathrm{gs,s}}$ which is mainly based on their rather small variation range while e_3 and f_3 appear to be constant.

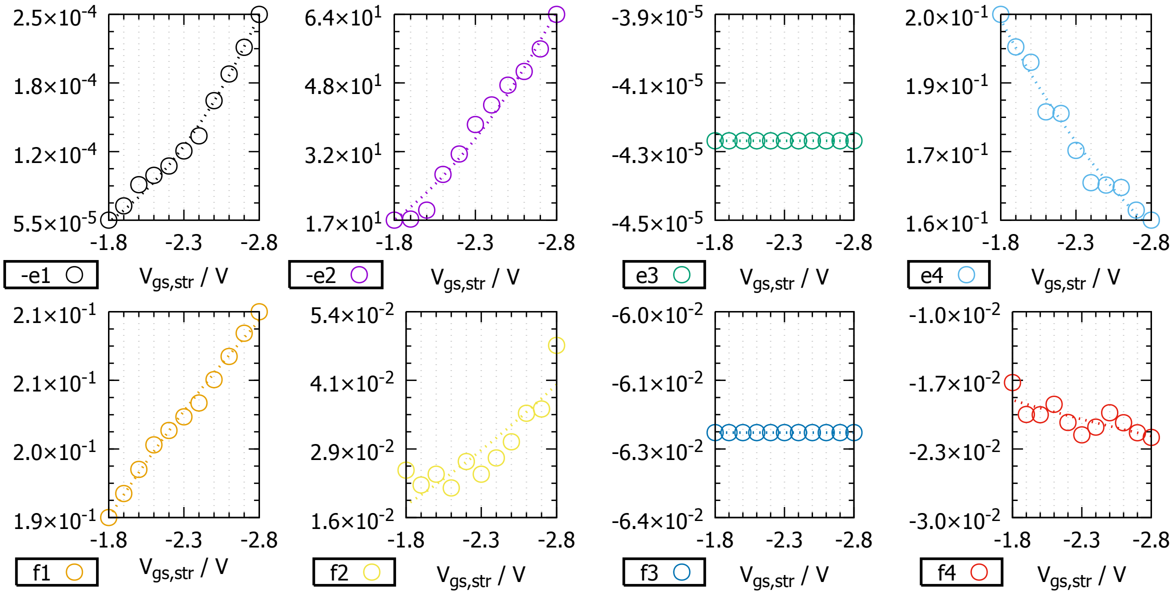


Figure 4.72.: The gate-source voltage dependence of the eight fit parameters (symbols) of the CET maps presented in this work can be approximated with power laws (dashed lines) as shown in Eq. 4.5.

For homogeneous NBTI the threshold voltage drift can be expressed via the parameters $e_{\mathrm{p}}\left(V_{\mathrm{gs}}\right)$ and $f_{\mathrm{p}}\left(V_{\mathrm{gs}}\right)$ by:

$$\Delta V_{\mathrm{th}}\left(V_{\mathrm{gs,s}}, t_{\mathrm{r}}, t_{\mathrm{s}}\right) \approx \frac{\Delta V_{\mathrm{th},\infty}}{1+\left(\frac{\Delta V_{\mathrm{th},\infty}}{\Delta V_{\mathrm{th},0}}-1\right)e^{c t_{\mathrm{r}}^{d}}} = \frac{e_1 t_{\mathrm{s}}^{f_1}}{1+\left(\frac{e_1 t_{\mathrm{s}}^{f_1}}{e_2 t_{\mathrm{s}}^{f_2}}-1\right)e^{\left(e_3 t_{\mathrm{s}}^{f_3}\right)t_{\mathrm{r}}^{\left(e_4 t_{\mathrm{s}}^{f_4}\right)}}}$$

with

$$e_{\mathrm{p}} \approx g_{\mathrm{p}1}\left|V_{\mathrm{gs,s}}\right|^{h_{\mathrm{p}1}}, f_{\mathrm{p}} \approx g_{\mathrm{p}2}\left|V_{\mathrm{gs,s}}\right|^{h_{\mathrm{p}2}}, \mathrm{p} \in \{1\ldots4\}, \tag{4.5}$$

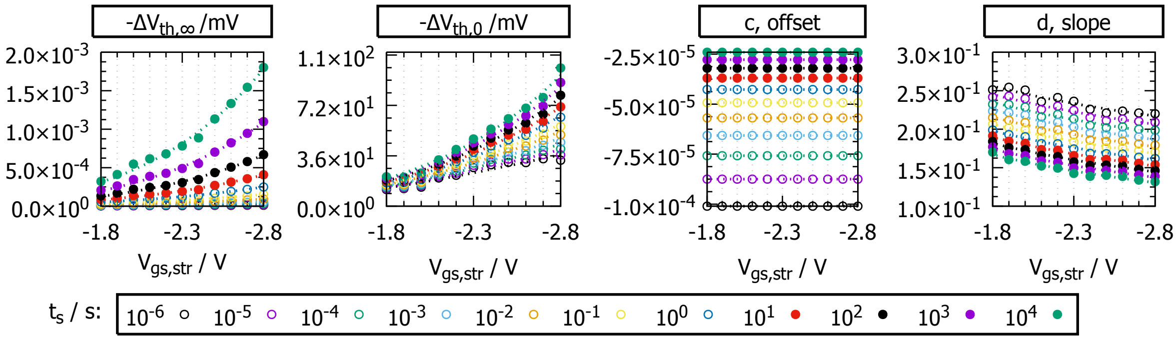


Figure 4.73.: The four recalculated CET map parameters (dashed lines) are in very good agreement with the original values (symbols).

ab	11	12	21	22	31	32	41	42
g_{ab}	6.8×10^{-6}	0.168	2.778	0.006	-4.3×10^{-5}	-0.062	0.263	-0.015
h_{ab}	3.500	0.236	3.053	1.800	1.00×10^{-5}	-4.7×10^{-13}	-0.481	0.390

Table 4.2.: The fitted values of the 16 parameters to calculate the gate-source voltage, stress and recovery time dependent threshold voltage drift of the core p-MOSFET.

The additional HCD contribution ($V_{\mathrm{gs,s}} = -2.8\,\mathrm{V}$) can be approximated by:

$$\begin{aligned} e_{\mathrm{p}} &\approx g_{\mathrm{p1}}\left|V_{\mathrm{gs,s}}\right|^{h_{\mathrm{p1}}} + i_{\mathrm{p1}}\left|V_{\mathrm{ds,s}}\right|^{k_{\mathrm{p1}}},\, f_{\mathrm{p}} \approx g_{\mathrm{p2}}\left|V_{\mathrm{gs,s}}\right|^{h_{\mathrm{p2}}} + i_{\mathrm{p2}}\left|V_{\mathrm{ds,s}}\right|^{k_{\mathrm{p2}}},\, \mathrm{p} \in \{1,3,4\} \\ e_2 &\approx g_{21}\left|V_{\mathrm{gs,s}}\right|^{h_{21}} + i_{21}\left|V_{\mathrm{ds,s}}\right|^{k_{21}} + i_{23}\left|V_{\mathrm{ds,s}}\right|^{k_{23}},\, f_2 \approx g_{22}\left|V_{\mathrm{gs,s}}\right|^{h_{22}} + i_{22}\left|V_{\mathrm{ds,s}}\right|^{k_{22}} \end{aligned} \tag{4.6}$$

which yields fit results presented in Fig. 4.74 and Fig. 4.75. The parameters listed in Table 4.2 and 4.3 allow a good quantitative reproduction of the recorded data and the extracted four fit parameters of each CET map (see Fig. 4.73).

The influence of the temperature on the fitted parameters has not been presented here because most of the CET maps for this work were recorded at only two temperatures, 398 K and 443 K. But in the future this additional dependence will be analyzed together

ab	11	12	21	22	23
i_{ab}	1.36×10^{-5}	7.54×10^{-4}	-1.15×10^{3}	9.02×10^{-3}	1.14×10^{3}
k_{ab}	4.892	5.228	0.845	0.899	0.877

ab	31	32	41	42
i_{ab}	1.29×10^{-10}	1.01×10^{-13}	-0.020	0.020
k_{ab}	0.400	0.997	1.712	1.238

Table 4.3.: The fitted values of the 18 parameters to calculate the additional drain-source voltage related component of the threshold voltage drift of the core p-MOSFET with respect to the stress and recovery time including inhomogeneous NBTI and the mixture of NBTI and HCD. The drift minimum condition depicted in Fig. 4.57 is realized by an extension of e_2 via i_{23} and k_{23}.

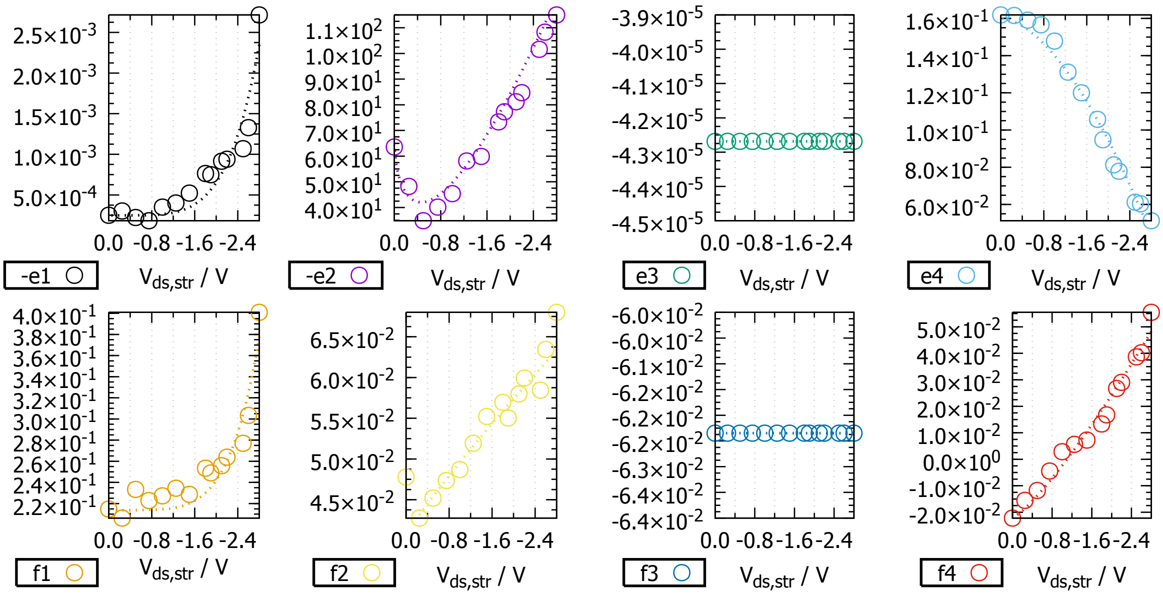


Figure 4.74.: The drain-source extension for the additional HCD contribution of the eight fit parameters (symbols) of the CET maps presented in this work can be approximated with power laws (dashed lines) as shown in Eq. 4.6.

with the thermal equivalent time (see Eq. 4.2). Therefore, at the moment one may simply substitute t_r and t_s with their converted thermal equivalent times which will reproduce the real recovery trace data with a reasonably small error. Only the second order differentiated data which is used as input for the CET maps will show noticeable differences to the real data as shown in Fig. 4.44 and discussed on page 101.

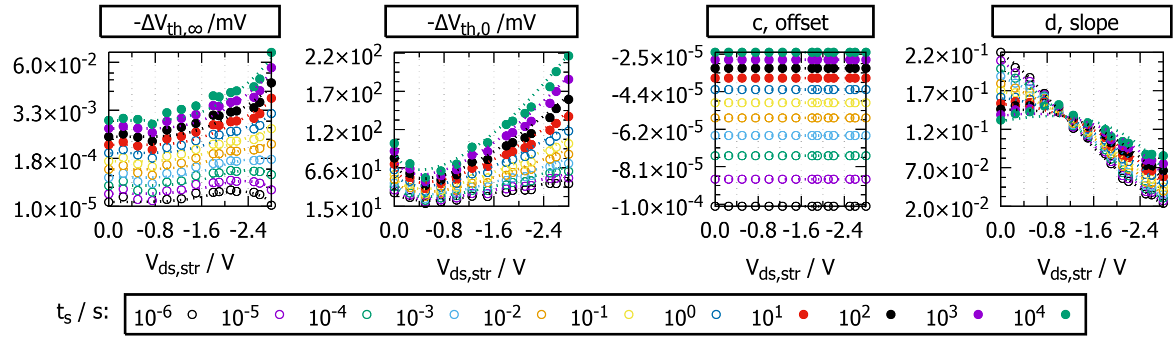


Figure 4.75.: The four recalculated CET map parameters (dashed lines) reflect the original values (symbols) very well. In addition, the drift minimum condition at $V_\mathrm{ds} = -0.5\,\mathrm{V}$ which mainly originates from $\Delta V_\mathrm{th,0}$ can be reproduced by the proposed model using the parameter e_2 which is fitted by a superposition of two power laws.

4.9. Trap class activation in dependence of the gate-source stress voltage

A trap class is defined as a boundary of two capture and two emission times in-between which defects can be identified in a CET map (see Section 1.2.6 and Fig. 1.43). In order to understand the voltage dependence of the activation of the individual trap classes with different τ_c and τ_e ranges the fit model presented in Section 4.8 (Eq. 4.5) is used to calculate differential CET maps. By that method the additional probability density (g_+) gained via an increasing gate-source stress voltage can be represented as shown in Fig. 4.76. For a fixed $\delta V_\mathrm{gs,s} = 0.4\,\mathrm{V}$ the slope of the maximum of the differential probability density shows a strong dependence on the selected voltage domain. At a rather low gate-source voltage (difference between $V_\mathrm{gs,s,h} = -1.2\,\mathrm{V}$ and $V_\mathrm{gs,s,l} = -0.8\,\mathrm{V}$) the slope is much steeper compared to the high stress domain (difference between $V_\mathrm{gs,s,h} = -2.8\,\mathrm{V}$ and $V_\mathrm{gs,s,l} = -2.4\,\mathrm{V}$). To illustrate the stress and recovery time dependence of g_+, slices of the differential CET maps with $V_\mathrm{gs,s,h} = -2.8\,\mathrm{V}$, $-2.4\,\mathrm{V}$, $-2\,\mathrm{V}$, $-1.6\,\mathrm{V}$ and $-1.2\,\mathrm{V}$ were calculated . Compared to the typical CET map representation in this work g_+ is not normalized in Fig. 4.77. Because of that the influence of the stress time on the recovery behavior of the additional traps of each class can be observed more easily. For a short stress time the differential signal reveals that the additional activated defects of one class of any voltage level (shown is always g_+ of $V_\mathrm{gs,s,h} - V_\mathrm{gs,s,l} = \delta V_\mathrm{gs,s} = 0.4\,\mathrm{V}$) simply discharge monotonously during the recovery phase leading to a value of $g_+ = 0$ (Fig. 4.77, left). By increasing the stress time this picture changes. The longer the stress is applied the shallower the slope of the recovery signal becomes (see e.g. Fig. 2.20). This leads to a negative value of g in the second order derivative. As the total value of g increases with $V_\mathrm{gs,s}$ the differential picture of two CET maps calculated for $V_\mathrm{gs,s,h}$ and $V_\mathrm{gs,s,l}$ also follows this behavior. Therefore, long stress times lead to a negative value of g_+ and the findings are more pronounced for an increasing $|V_\mathrm{gs,s,h}|$ (Fig. 4.77, right). After that, the negative value of g_+ increases to a maximum and then decreases to 0.

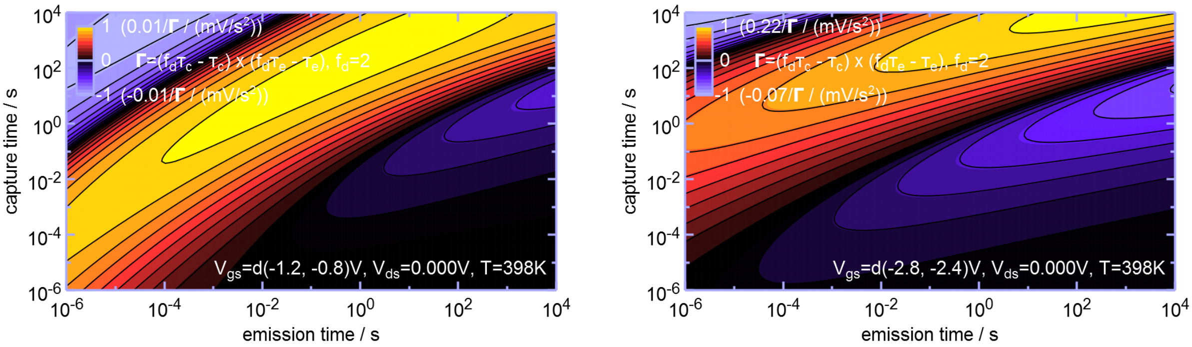


Figure 4.76.: Differential CET maps of homogenous NBTI showing the amount of the additional trap density gained via an increasing gate-source stress voltage. It is obvious that the slope of the maximum of the gained probability density strongly depends on the chosen $V_{\mathrm{gs,s}}$ domain. The slope at the difference between $V_{\mathrm{gs,s,h}} = -1.2\,\mathrm{V}$ and $V_{\mathrm{gs,s,l}} = -0.8\,\mathrm{V}$ is much steeper compared to the slope at $V_{\mathrm{gs,s,h}} = -2.8\,\mathrm{V}$ and $V_{\mathrm{gs,s,l}} = -2.4\,\mathrm{V}$.

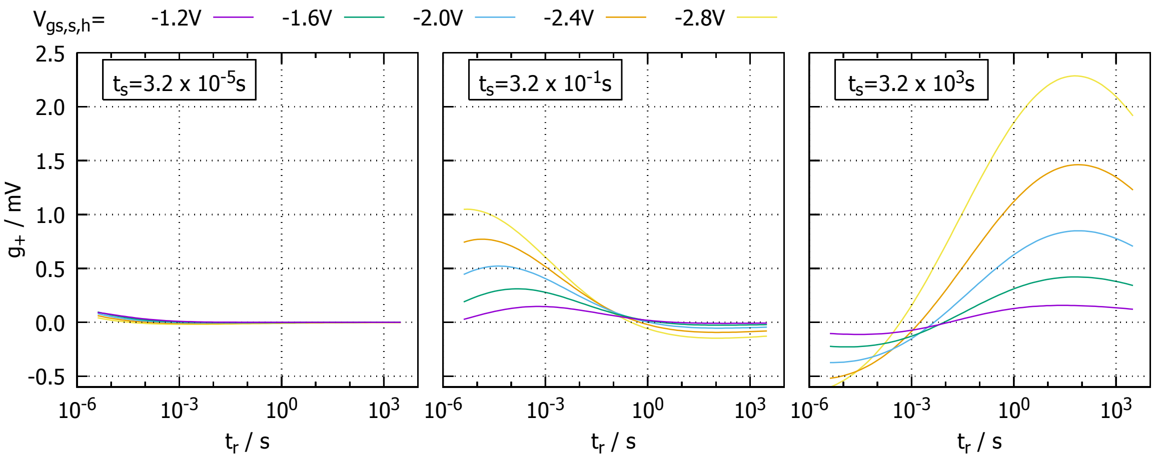


Figure 4.77.: The stress time (left: $t_{\mathrm{s}} = 3.2 \times 10^{-5}\,\mathrm{s}$, middle: $t_{\mathrm{s}} = 3.2 \times 10^{-1}\,\mathrm{s}$, right: $t_{\mathrm{s}} = 3.2 \times 10^{3}\,\mathrm{s}$) has a strong influence on the recovery behavior of the additionally charged traps. For short stress times the stress voltage dependence may be approximated by an amplitude factor (left), but for longer stress times also the maximum of the differential signal changes its recovery time position. The binning of the emission and capture time is set to integer powers of ten ($f_{\mathrm{d}} = 10$).

5. Circuit simulation

5.1. Compact models for circuit simulation

The drift data of transistors stressed at the worst-case condition for one aging mechanism allow the extraction of a reliability model for each measured device parameter. This model can be used to calculate either the drift of the electrical parameters (e.g. threshold voltage, linear operating point) or the lifetime of the transistor in the designated application, employing the widely used simple but effective power law of the stress time for the parameter drift dependence ($\Delta P \propto t^n_{\text{operation}}$) i.e.

$$\Delta P = \left(t_{\text{operation}} D\right)^n f\left(V_{\text{ds}}, V_{\text{gs}}, T\right) \tag{5.1}$$

$$\frac{t_{\text{lifetime}}}{\text{s}} = \left(\frac{\Delta P(t_{\text{lifetime}})}{\Delta P_{\text{target}}}\right)^{\frac{1}{n}} \frac{1}{D} f'\left(V_{\text{ds}}, V_{\text{gs}}, T\right), \tag{5.2}$$

$$f' = f^{-\frac{1}{n}} = \mathcal{A}\mathcal{B}\varGamma, \quad D = \frac{t_{\text{stress}}}{t_{\text{operation}}} \tag{5.3}$$

- Arrhenius-like temperature activation:

$$\mathcal{A} = \mathrm{e}^{\frac{E_a}{k_B}\left(\frac{1}{T_{\text{stress}}} - \frac{1}{T_{\text{target}}}\right)} \tag{5.4}$$

- typical examples for $\mathcal{B}$ and $\varGamma$:

$$\begin{aligned}
\mathcal{B}_{\text{exponential}} &= \mathrm{e}^{\beta\left(\frac{1}{L_{\text{design}}} - \frac{1}{L_{\text{target}}}\right)}, \quad \mathcal{B}_{\text{power}} = \left(L_{\text{design}}\right)^{\beta} \\
\text{HCD}: \quad \varGamma_{\text{exponential}} &= \mathrm{e}^{\gamma\left(\frac{1}{V_{\text{ds,stress}}} - \frac{1}{V_{\text{ds,target}}}\right)}, \quad \varGamma_{\text{power}} = \left(V_{\text{ds,stress}}\right)^{\gamma} \\
\text{BTI}: \quad \varGamma_{\text{exponential}} &= \mathrm{e}^{\gamma(V_{\text{gs,stress}} - V_{\text{gs,target}})}, \quad \varGamma_{\text{power}} = \left(V_{\text{gs,stress}}\right)^{\gamma}
\end{aligned}$$

with parameter drift ΔP, fixed stress conditions for operation time $t_{\text{operation}}$, power-law coefficient n, the duty factor D, the activation energy E_a of the tested degradation mechanism for the Arrhenius term as well as fit functions for the degradation dependence of the length ($\mathcal{B}$) and the voltage ($\varGamma$) at the terminal which is most relevant for the assessed degradation mechanism (e.g. $V_{\text{ds,stress}}$ for HCD or $V_{\text{gs,stress}}$ for BTI). ΔP_{target} is the maximum allowed drift for which the device parameter is qualified (e.g. 100 mV for the threshold voltage) at temperature T_{target}, channel length L_{target} and applied drain-source voltage $V_{\text{ds,target}}$.

Eq. 5.1 is used to calculate the drift caused by the main degradation mechanism. In order to include, for instance, the second order gate voltage dependence for hot-carrier degradation this can be extended by the experimental data of the 2-dimensional stress voltage matrices presented in Section 4.2 to:

$$\Delta P = \left(t_{\text{operation}} D\right)^n f\left(V_{\text{ds}}, T, ...\right) \mathcal{C}\left(V_{\text{gs}}\right) \tag{5.5}$$

with a fit function $\mathcal{C}$ which depends on the gate-source voltage. An example for this function is a polynomial like

$$\mathcal{C} = a_0 + a_1 \frac{V_{\text{gs,stress}}}{V} + a_2 \left(\frac{V_{\text{gs,stress}}}{V}\right)^2 + ... + a_6 \left(\frac{V_{\text{gs,stress}}}{V}\right)^6 \tag{5.6}$$

and depicted in Fig. 5.1.

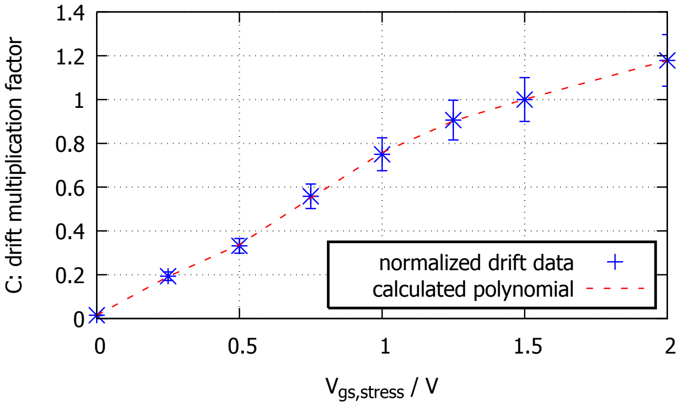


Figure 5.1.: A polynomial to describe the influence of the gate-source voltage on the parameter drift for hot-carrier degradation of the examined core device class n-MOSFET. The error bars indicate a device to device variation of the derived parameter drift of 10%. Here C is set to 1 for the worst-case HCD operation condition ($V_{\text{dd}} = 1.5\,\text{V}$) at which no PBTI is expected.

The drift formula (Eq. 5.1) presented above is used to calculate the drift of a device parameter with fixed stress conditions for a specific operation time. To enable the assessment of arbitrary stress signals (V_{stress}), time-dependent stress conditions need to be considered. This can be realized with an effective duty factor D_{deg} for the degradation under arbitrary voltages. The equivalent time ($t_{\text{step,eff}}$) which is needed to reach the same amount of degradation for two different stress voltages (V_1 and V_2) for the simulated time step with duration t_{step} is derived as follows:

$$\Delta P\left(V_1, t_{\text{step}}\right) \overset{def.}{=} \Delta P\left(V_2, t_{\text{step}} D_{\text{deg}}\right) \tag{5.7}$$

$$\overset{eq.\ 5.1}{\Leftrightarrow} Cf\left(V_1\right)\left(t_{\text{step}}\right)^n = Cf\left(V_2\right)\left(t_{\text{step}} D_{\text{deg}}\right)^n \tag{5.8}$$

$$\Leftrightarrow D_{\text{deg}} = \left[\frac{f\left(V_1\right)}{f\left(V_2\right)}\right]^{\frac{1}{n}} \tag{5.9}$$

$$\Rightarrow t_{\text{step,eff}} = t_{\text{step}} \left[\frac{f\left(V_1\right)}{f\left(V_2\right)}\right]^{\frac{1}{n}} \tag{5.10}$$

This equivalent time can now be inserted into the equation with V_2 and the result will be the degradation after t_{step} by applying the stress voltage V_1.

To include all other time-dependent functions (e.g. Arrhenius law for a temperature profile) this method can be generalized to:

$$t_{\text{eff}} = t_{\text{step}} \left[\frac{\Delta P\left(V_1, T_1, ..., t_{\text{step}}\right)}{\Delta P\left(V_2, T_2, ..., t_{\text{step}}\right)}\right]^{\frac{1}{n}} = t_{\text{step}} \left[\frac{f\left(V_1, T_1, ..., t_{\text{step}}\right)}{f\left(V_2, T_2, ..., t_{\text{step}}\right)}\right]^{\frac{1}{n}} \tag{5.11}$$

The derived equations are valid for one simulation step with length t_{step}. In order to calculate the effective duty factor for the whole simulation period with duration $t_{\text{period}} = t_{\text{end}} - t_{\text{start}} = \sum_i t_{\text{step,i}}$, the arbitrary (time-dependent) stress condition being represented by index "1" and a fixed (time-independent) stress condition like in the original drift Eq. 5.1 being represented by index "2", the following time step weighted average is introduced:

$$D_{\text{deg}} = \frac{\sum_{\text{t}_{\text{start}}}^{t_{\text{end}}} \left[\frac{f(V_1, T_1, ..., t_{\text{step}})}{f(V_2, T_2, ..., t_{\text{step}})}\right]^{\frac{1}{n}} t_{\text{step}}}{t_{\text{period}}} \tag{5.12}$$

Because the original drift Eq. 5.1 is time-independent the term indexed with "2" is a constant of the sum over t_{step} and can be factorized:

$$D_{\text{deg}} = \frac{\sum_{\text{t}_{\text{start}}}^{t_{\text{end}}} \left[f\left(V_1, T_1, ..., t_{\text{step}}\right)\right]^{\frac{1}{n}} t_{\text{step}}}{\left[f\left(V_2, T_2, ..., t_{\text{step}}\right)\right]^{\frac{1}{n}} t_{\text{period}}} = \frac{\left\langle \left[f\left(V_1, T_1, ..., t_{\text{step}}\right)\right]^{\frac{1}{n}} \right\rangle_t}{\left[f\left(V_2, T_2, ..., t_{\text{step}}\right)\right]^{\frac{1}{n}}} \tag{5.13}$$

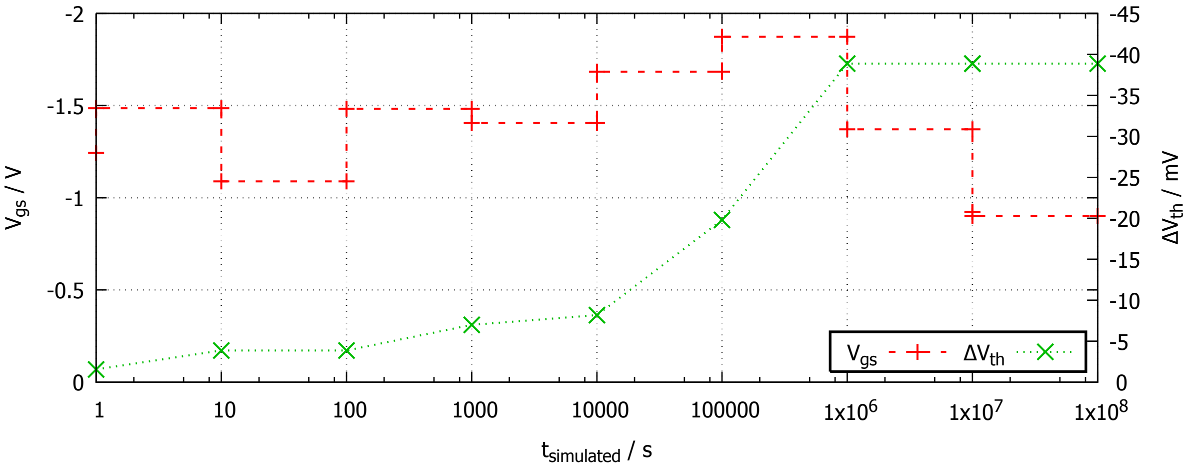


Figure 5.2.: The simulated NBTI degradation (neglecting recovery effects) of the examined 130 nm technology p-MOSFET with $W/L = 10/0.1$ under an arbitrary signal at the gate terminal. Red symbols: stress voltage at simulation time $t_{\text{simulated}}$. Green symbols: accumulated degradation of the threshold voltage of the stressed transistor. Lines are guides to the eye.

This leads to the final expression to calculate the degradation caused by arbitrary stress conditions over one simulation period:

$$\begin{aligned}\Delta P\left(t_{\text{period}}\right) &= Cf\left(V_2, T_2, ..., t_{\text{step}}\right)\left(t_{\text{step}} D_{\text{deg}}\right)^n &(5.14)\\ &= Cf\left(V_2, T_2, ..., t_{\text{step}}\right)\left(t_{\text{step}} \frac{\left\langle\left[f\left(V_1, T_1, ..., t_{\text{step}}\right)\right]^{\frac{1}{n}}\right\rangle_t}{\left[f\left(V_2, T_2, ..., t_{\text{step}}\right)\right]^{\frac{1}{n}}}\right)^n &(5.15)\\ &= C\left(t_{\text{step}}\left\langle\left[f\left(V_1, T_1, ..., t_{\text{step}}\right)\right]^{\frac{1}{n}}\right\rangle_t\right)^n &(5.16)\end{aligned}$$

An example of this methodology is shown in Fig. 5.2. Two implementations of the drift formula were realized for this work, a RelXpert URI [195] and a SPICE assertion based setup. These setups simulate the aging of one single operation period of the circuit which is typically in the range of several microseconds and extrapolate the calculated result of the parameter drift to the desired period of use in the application (e.g. 10 years). This method also ensures that the simulated drift matches the experimental data at the expected lifetime because the aging models with a power law dependence of the drift on the stress time are optimized for that time region (see Fig. 5.11).

To include recovery effects in aging simulations of the examined 130 nm technology the compact model from [168] should be employed because it is based on the measurement data presented in this work. Here, the four-state NMP model (see Fig. 1.29 on page 41) is projected onto a Markov two-state model as shown in Fig. 5.3. The two possible states describe either a neutral or a charged defect while the latter contributes to the threshold voltage shift caused by NBTI degradation.

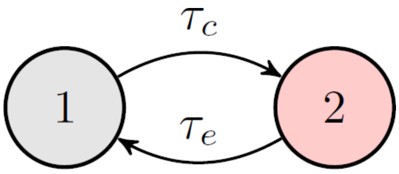


Figure 5.3.: The defects of the Markov two-state model can transform from state "1" (neutral) to state "2" (charged) during the needed mean capture time τ_c and the retransformation from state "2" to state "1" within the emission time τ_e, [196]. Of course, the capture and emission time are gate bias and temperature dependent.

A first-order ordinary differential equation (ODE) is needed to describe the probability ω of the defect being charged at time t:

$$\dot{\omega}\left(t\right) = a\left(t\right)\omega\left(t\right) + b\left(t\right), \omega\left(t_0\right) = \omega_0 \tag{5.17}$$

$$a\left(t\right) = -\left(\tau_\mathrm{e}^{-1}\left(t\right) + \tau_\mathrm{c}^{-1}\left(t\right)\right) < 0 \tag{5.18}$$

$$b\left(t\right) = \tau_\mathrm{c}^{-1}\left(t\right) \geq 0 \tag{5.19}$$

with functions a and b representing time-dependent capture and emission time constants. Contrary to digital stress with only two biases, $V_\mathrm{gs,low}$ and $V_\mathrm{gs,high}$, and a single pair of capture and emission times, here, this time-dependence is needed to model analog stress signals (i.e. the oxide field dependence of NBTI).
The solution of Eq. 5.17 is:

$$\omega\left(t\right) = P_1\left(t, t_0\right)\omega\left(t_0\right) + P_2\left(t, t_0\right) \tag{5.20}$$

with the two time integrals

$$P_1\left(t_2, t_1\right) = \exp\left[\int_{t_1}^{t_2} \mathrm{d}s\, a\left(s\right)\right] \tag{5.21}$$

$$P_2\left(t_2, t_1\right) = \int_{t_1}^{t_2} \mathrm{d}s\, b\left(s\right) \exp\left[\int_{s}^{t_2} \mathrm{d}r\, a\left(r\right)\right] = \int_{t_1}^{t_2} \mathrm{d}s\, b\left(s\right) P_1\left(t_2, s\right) \tag{5.22}$$

and $0 < P_1\left(t_2, t_1\right) \leq 1$, $0 \leq P_2\left(t_2, t_1\right)$ for $t_2 \geq t_1$ and $P_2\left(t, t\right) = 0$. These can be arranged in a 2×2 matrix

$$P\left(t_2, t_1\right) = \begin{pmatrix} P_1\left(t_2, t_1\right) & P_2\left(t_2, t_1\right) \\ 0 & 1 \end{pmatrix} \tag{5.23}$$

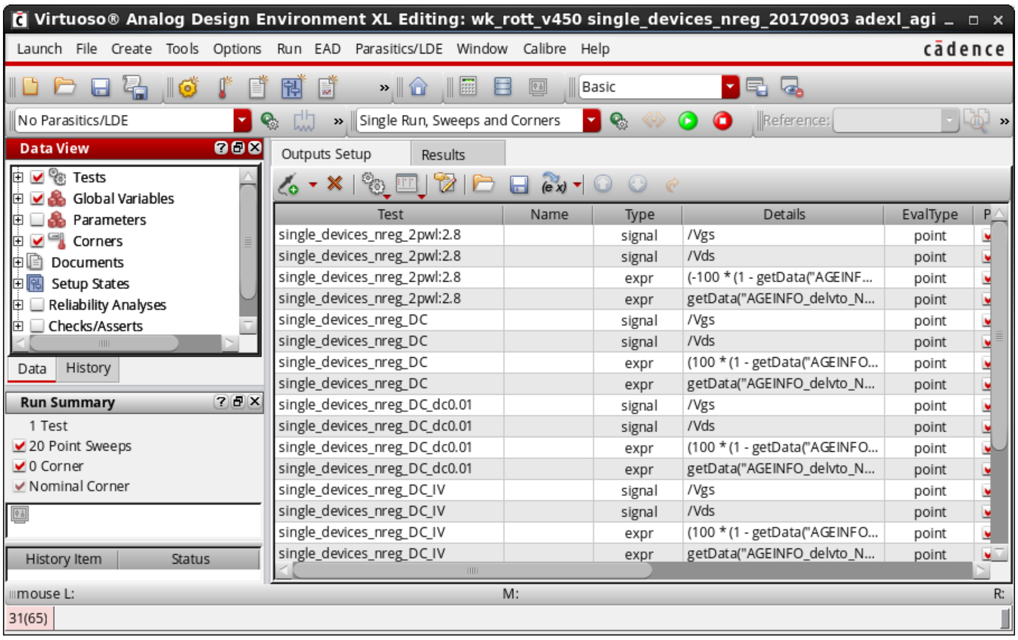
which propagates the solution of the ODE from the initial (t_1) to the solution time (t_2) with $t_2 > t_1$. In addition, the group property

$$P\left(t_2, t_0\right) = P\left(t_2, t_1\right) P\left(t_1, t_0\right) \tag{5.24}$$

is satisfied, showing that the solution $\omega\left(t_2\right)$ can be obtained by propagation via the intermediate time t_1. That is why, the numerical but exact solution can be calculated for any time period $T > t_2 - t_1 > 0$ of V_gs.

The presented compact model from [168] is in a very good agreement with the measured data of this work as well as to microscopic NBTI simulations (see Fig. 4.49 in Section 4.3 and [196]).

5.2. SPICE simulations with SPECTRE



Test	Name	Type	Details	EvalType	P
single_devices_nreg_2pwl:2.8		signal	/Vgs	point	☑
single_devices_nreg_2pwl:2.8		signal	/Vds	point	☑
single_devices_nreg_2pwl:2.8		expr	(-100 * (1 - getData("AGEINF...	point	☑
single_devices_nreg_2pwl:2.8		expr	getData("AGEINFO_delvto_N...	point	☑
single_devices_nreg_DC		signal	/Vgs	point	☑
single_devices_nreg_DC		signal	/Vds	point	☑
single_devices_nreg_DC		expr	(100 * (1 - getData("AGEINFO...	point	☑
single_devices_nreg_DC		expr	getData("AGEINFO_delvto_N...	point	☑
single_devices_nreg_DC_dc0.01		signal	/Vgs	point	☑
single_devices_nreg_DC_dc0.01		signal	/Vds	point	☑
single_devices_nreg_DC_dc0.01		expr	(100 * (1 - getData("AGEINFO...	point	☑
single_devices_nreg_DC_dc0.01		expr	getData("AGEINFO_delvto_N...	point	☑
single_devices_nreg_DC_IV		signal	/Vgs	point	☑
single_devices_nreg_DC_IV		signal	/Vds	point	☑
single_devices_nreg_DC_IV		expr	(100 * (1 - getData("AGEINFO...	point	☑
single_devices_nreg_DC_IV		expr	getData("AGEINFO_delvto_N...	point	☑

Figure 5.4.: Cadence Virtuoso ADE XL was used to set up all SPICE simulations presented in this chapter. The SPICE simulations were performed employing the Cadence simulator SPECTRE.

All SPICE simulations presented in this chapter were performed with the help of the Cadence simulator SPECTRE. The schematics which are used in the simulation flow of the Cadence Virtuoso Analog Design Environment (ADE, Fig. 5.4) consist of transistors, resistors, capacitors and other basic device types as well as the wiring and the terminals of the circuit. To simulate the circuit, so-called *stimuli* are applied to the terminals of the circuit and other environmental variables are set (e.g. the ambient temperature). The final setup is referred to as a *test bench* which can be executed by a SPICE simulator. Several test benches can be grouped in Virtuoso as a *view* of type "adexl" together with global variables for e.g. parameter sweeps and corner setups which allow a parametrization of the stimuli and environment variables.

5.3. Aging simulations

5.3.1. Single transistor degradation

Constant stress bias

A test bench has been set up in Cadence Virtuoso ADE XL to demonstrate the hot-carrier related drift behavior of the core n- and p-MOSFET at 298 K (see Fig. 5.5). The comparison of measured and simulated drift data depicted in Fig. 5.6 shows the results for different drain- and gate-source stress conditions after a constant stress bias of 10^4 s of the core n- and p-MOSFET.

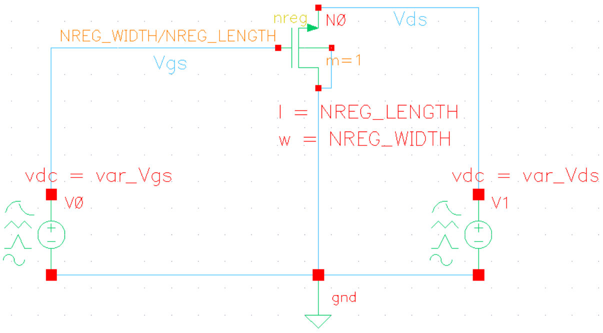


Figure 5.5.: The single device aging circuit has been equipped with individual multi-purpose voltage sources between the drain and source as well as between the gate and source terminal. This setup allows a simulation of the lab conditions and a direct comparison with the measurement data.

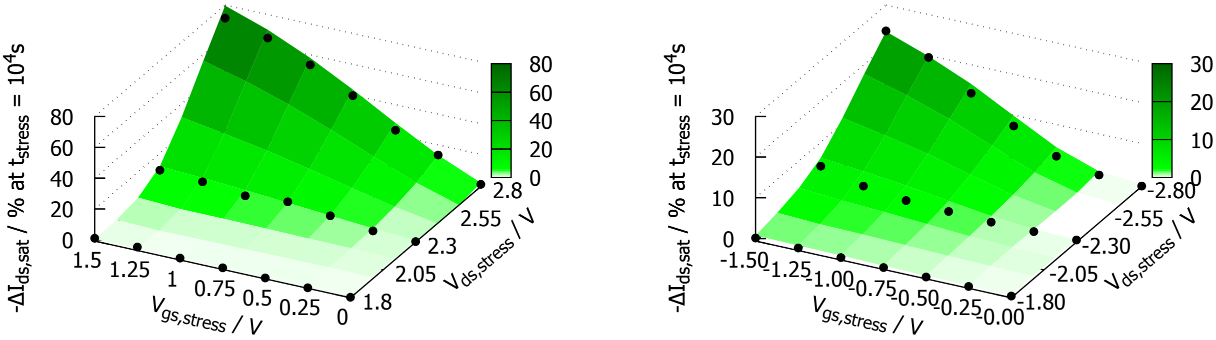


Figure 5.6.: Comparison of the results of the 2-dimensional HCD stress voltage matrix of the core n-MOSFET (left) and p-MOSFET (right), symbols: measurement data, surface: aging model.

A comparison of the simulation results to the measurement data presented in Fig. 5.7 shows good agreement between both. Of course, the chosen fit model with the simplification $\Delta P \propto t_s^n$ overestimates the drift of the transistor during the early aging and wear-out phase but is used due to its compatibility to the device qualification drift model because it reproduces the measurement data around the drift limit of the application (e.g. $\Delta I_{\mathrm{ds,sat}} = 10\%$ or $\Delta V_{\mathrm{th}} = 100\,\mathrm{mV}$) very well.

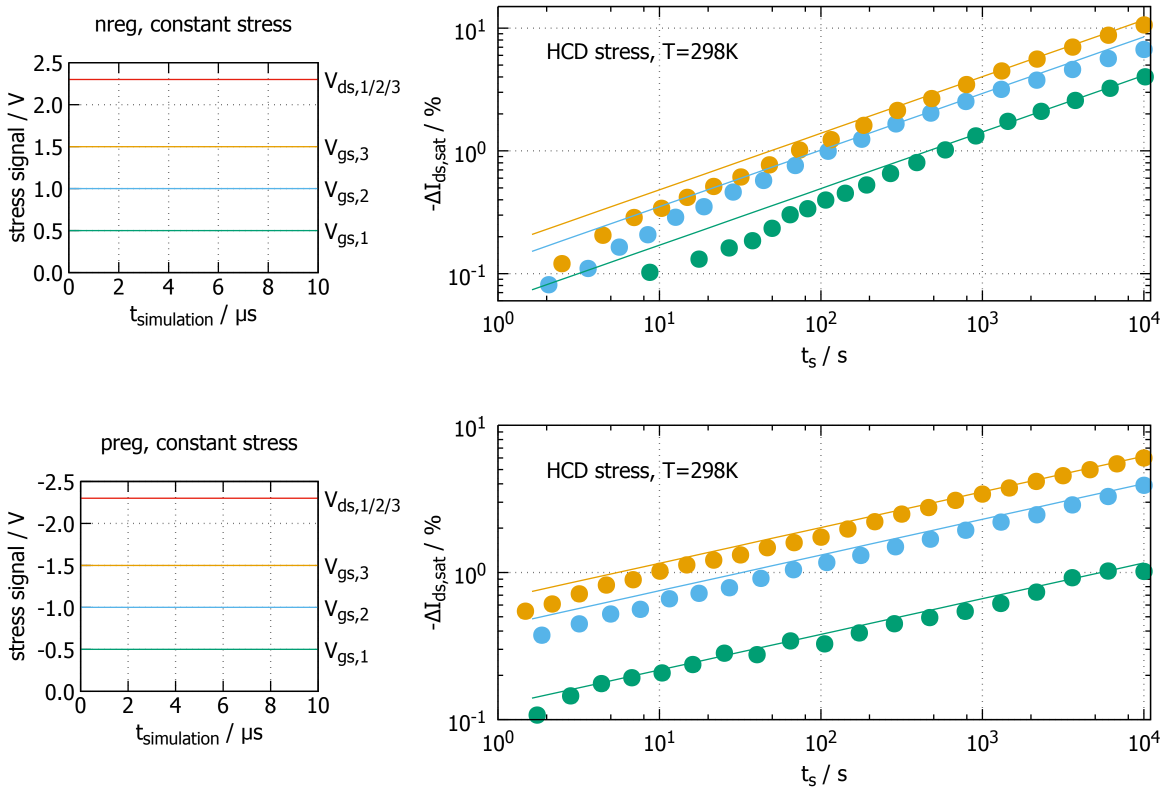


Figure 5.7.: Right: Comparison between measurement data (smbols) and aging simulations (lines) of the drain-source current drift at several gate-source voltages and fixed drain-source voltage of the core n-MOSFET (top) and p-MOSFET (bottom). The simulated stress period is shown on the left side and used to extrapolate the drift $\Delta P \propto t_s^n$ to the stress time t_s depicted in the right plot. To emphasize HCD and suppress BTI a stress temperature of 298 K was chosen.

Influence of the duty cycle

If the constant stress bias is applied only for a certain amount of time to the transistor, the stress duty factor $D = \frac{t_{\text{stress}}}{t_{\text{operation}}}$ can be utilized following eq. 5.5. Simulation results of the core n-MOSFET for different values of D are shown together with the according stress signals and measurement data in Fig. 5.8. Like for the case of constant stress (see Fig. 5.7) also the duty cycle dependence indicates that the simulated drift overestimates the early and wear out phase of the device under test which here again is caused by the power law extrapolation of the drift via the total stress time.

Arbitrary stress signals

In an analog application the importance of the effective duty factor D_{deg} (eq. 5.13) becomes obvious because the stress duty factor cannot be estimated from the stress pattern directly. Now the acceleration of the terminal voltage also needs to be taken into

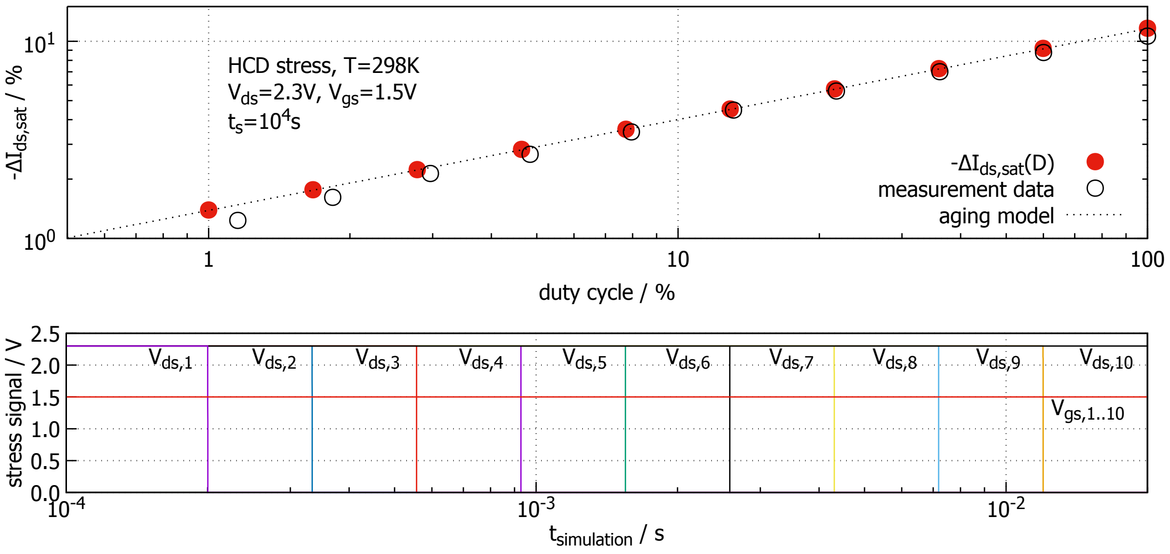


Figure 5.8.: Results of 10 aging simulations applying different duty cycles (signals are shown in the bottom graph) to the core n-MOSFET stressed for 10^4 s at HCD conditions, $T = 298\,\mathrm{K}, V_{ds} = 2.3\,\mathrm{V}, V_{gs} = 1.5\,\mathrm{V}, D = (1.0, 1.7, 2.8, 4.6, 7.7, 12.9, 21.5, 35.9, 60, 100)\,\%$, compared with measurement data. As guide to the eye the aging model is included in the top graph.

account to determine the parameter drift. Results for sine, square wave and sawtooth stress at the drain-terminal at constant gate bias are presented in Fig. 5.9.

NBTI aging simulation

To show the importance of a compact model of NBTI including recovery, the simulation of the experiment discussed in Section 4.3.3 has been repeated neglecting the recovery of the threshold voltage and is compared to those results in Fig. 5.10. Here, the drift over stress time is overestimated for each signal shape by at least a factor of 2 for long stress times, leading to a significant underestimation of the lifetime. A neglect of recovery effects might lead to a re-design of the whole circuit. Of course, we take the recovery effects into account for our Design for Reliability (DfR) approach in order to reduce the need of re-designs and simulate realistic drift information over the lifetime of each single transistor.

5.3.2. Circuit degradation

Because the simulation results of the single transistor reflect the measured drift data very well, also the degradation of circuits can be considered. For static aging mechanisms like HCD the fit model (eq. 5.1) yields results comparable to the experimental data and should be used due to its simplicity. If the stress applied to a single transistor is causing

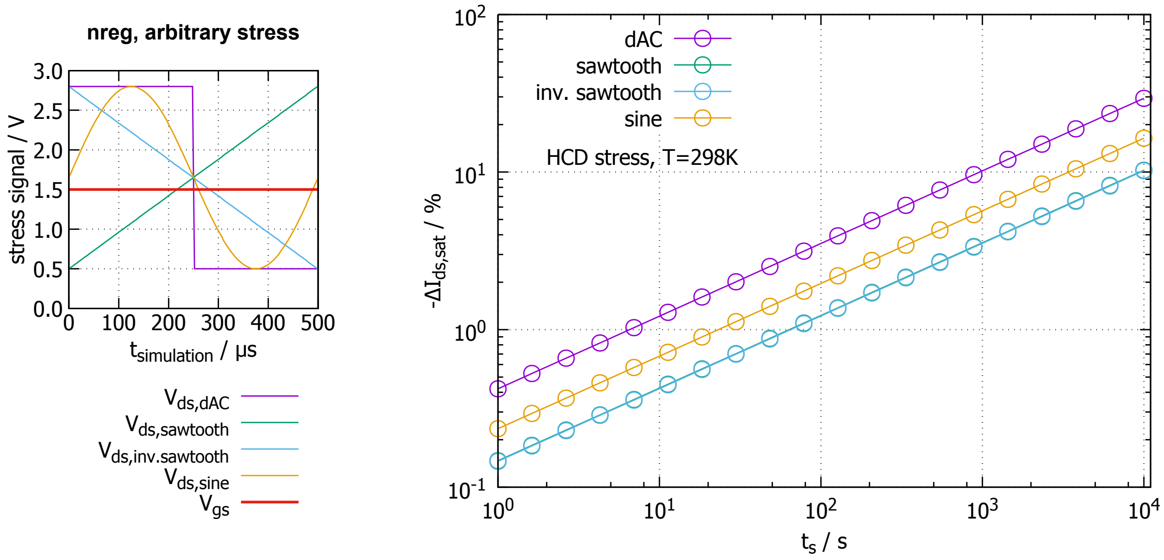


Figure 5.9.: Left: Stress signals applied during the simulation to the core n-MOSFET. **Right:** The drift caused by the four different stress signals shows a strong dependence on each stress signal. The largest drift is obtained with the dAC (square wave with offset) signal, followed by the sine signal and lowest in the case of the two overlaying sawtooth signals. As expected, the effective duty factor D_{deg} (eq. 5.13) leads to a parallel shift of the drift curves in their log-log representation.

BTI one has to use a more sophisticated aging model which also considers recovery effects (e.g. the model presented in [168]).

To show the validity of this hypothesis for HCD, extracted voltage signals of an inverter circuit have been applied to the core n-MOSFET at room temperature with an arbitrary waveform generator (experiment performed by Fabian Proebster). The comparison between measurement data and simulation is shown in Fig. 5.11 together with the applied stress signals and demonstrates the good agreement of both. Once the gate-source voltage dependence (eq. 5.6) is neglected ($C = 1, \mathrm{const.}$), the experimentally determined drift is overestimated about four times by the simulation. This example clearly highlights the importance of an aging model which takes all terminal voltages into account to calculate more realistic drift results compared to simple worst-case aging models.

During the early degradation phase the simulation generally overestimates the drift of the transistor. This can also be observed in the comparison graph for constant stress biases (Fig. 5.7) and is a typical result if the aging model is proportional to a power law of the stress time. On the other hand, for very long aging times when the degradation starts to saturate this power law dependence will also lead to an overestimation of the drift. Anyhow, this model type never underestimates the drift over the circuit lifetime which is a crucial key factor in terms of a risk assessment and guarantees a fail-safe

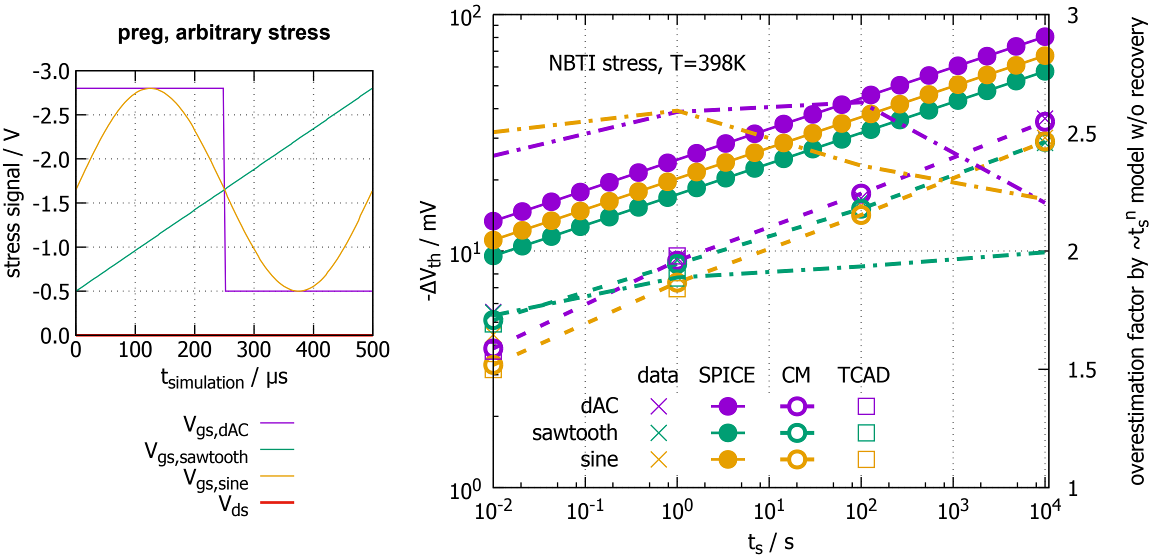


Figure 5.10.: Left: Stress signals applied to the core p-MOSFET during the simulation. **Right:** Comparison of the simulated drift (SPICE model ($\propto t_\mathrm{s}^n$), neglecting recovery) with the measurement data demonstrates the necessity of an aging model (e.g. compact model (CM) [168] or TCAD model [196]) which also takes recovery of the transistor into account, otherwise the actual threshold voltage drift will by overestimated by at least a factor of 2 (dash-dotted lines serve as guides to the eye) for long stress times.

operation.

As stated before, the drift of all transistors, which face BTI stress during operation, should be simulated employing an aging model, which takes recovery effects into account. Because the results of the single device data were reproduced with high consistency by the NBTI recovery model from [168], this compact model is a very promising candidate for circuit aging simulations. Especially for the investigated 130 nm technology a realistic drift information is obtainable because the compact model is based on both microscopic single trap data (TDDS) and macroscopic device data (CET maps).

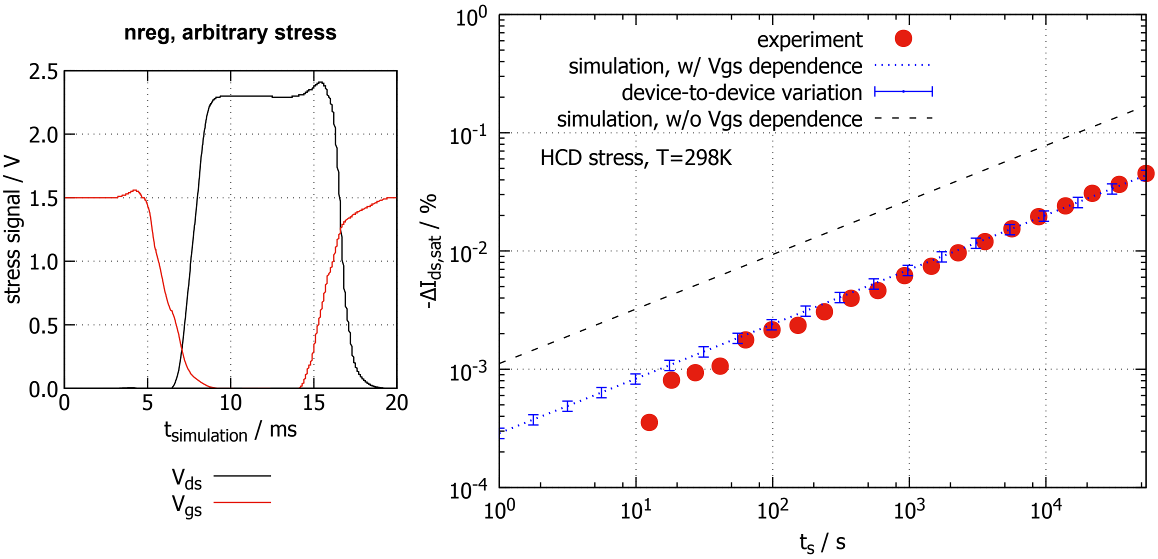


Figure 5.11.: Left: Stress signals applied by the arbitrary waveform generator to the core n-MOSFET and used for the aging simulation as stimuli. To emphasize HCD the drain-source voltage is increased to 2.3 V compared to the gate-source voltage which is limited to $\approx 1.5\,\mathrm{V}$. **Right:** The measured (symbols) and simulated drift (short dashes) of the device under test show consistent results for stress times longer than 50 s. If the gate-source voltage dependence is neglected ($A\,(V_{\mathrm{gs}} = 1.5\,\mathrm{V}) = 1$, long dashes) the experimentally determined drift is overestimated about four times. The deviations for shorter stress times can be related to the early aging phase of the constant stress bias data presented in Fig. 5.7 which also lead to an overestimation of the drain current drift by the simulation of the arbitrary signal stress.

6. Summary and conclusions

In the scope of this work transistors of a 130 nm technology (see Section 2.1) were subjected to in-depth reliability experiments. For that purpose a database of a dense 2-dimensional stress voltage matrix including different stress temperatures (see Chapter 3) has been built up and analyzed. Besides the measurements with the standard reliability lab equipment (see Section 2.3) an ultra-fast measurement equipment was assembled (after Reisinger) and extended to enable homogenous as well as inhomogeneous NBTI stress experiments with an ultra-short delay after stress of about 1 µs (see Section 2.5). In order to perform low noise experiments the internal power supplies of the probe station had to be replaced by external modules, which reduced the noise below the detection limit of a Tetronix DPO (see Section 2.2). For the statistical data acquisition a prober control module was implemented into the control software of the ultra-fast measurement equipment which allows a fully automatic measurement of a whole wafer. To handle all the recorded data a software environment (based on open source components) was programmed and adapted to all situations of this research (see Section 2.6). In this context a lot of effort was spent for the post-processing of CET map data including different types of data fitting (see sections 2.6.1 and 2.6.2) and the evaluation of the used parameters (see Section 4.8).

On the one hand the findings about the analyzed n-MOSFETs confirm the expected degradation behavior reported in literature. For the long channel transistors even the rather old "Hu model" from 1985 can be used to explain the aging caused by single carriers which overcome the critical energy [15] and dissociate Si-H bonds. Of course, for shorter channels (e.g. the analyzed core n-MOSFET) this model fails but the degradation can adequately be explained by successor models from Hess [11] or Bravaix et al. [20, 56].

On the other hand many interesting and new aspects about the degradation and recovery of the examined p-MOSFETs were found. The analog p-MOSFET with a channel length of 370 nm shows two hot-carrier degradation mechanisms. The first one is considered to be caused by single carrier effects which lead to a maximum of the gate current (therefore termed $D_{\mathrm{Igate,max}}$) at $V_{\mathrm{gs,s}} = -0.75\,\mathrm{V}$ and causes a reverse drift, the decrease of the absolute threshold voltage. The second degradation mechanism has its maximum effect at $V_{\mathrm{ds,s}} = V_{\mathrm{gs,s}}$ and is related to the interplay of multiple carriers which is emphasized at the maximum carrier flux and therefore called "$D_{\Psi,\mathrm{max}}$" in this work. Both degradation mechanisms can be explained by the models of Bravaix et al. [20, 56] and Tyaginov et al. [69–72]. Interestingly, $D_{\mathrm{Igate,max}}$ can be compensated by $D_{\Psi,\mathrm{max}}$ (see Fig. 4.21) if the gate-source voltage is chosen between the conditions of both degradation mechanisms, $V_{\mathrm{gs,s}} = -1.75\,\mathrm{V}$, independently of the stress temperature or the specific device geometry. An entirely physics based model published by Tyaginov et al. [69–72] (see Section 1.1.6) which takes all known hot-carrier degradation effects into account and explains them via the carrier distribution function is able to explain the HCD related results presented in this work.

In addition to the stress time dependence, the threshold voltage recovery after HCD stress of the core p-MOSFET has been analyzed and discussed in Section 4.5. As depicted in Fig. 4.52 the long term recovery signal shows no change in the threshold voltage drift. Therefore, no HCD related recovery is expected. If the gate-source voltage is increased to about $-1.5\,\mathrm{V}$ a small recovery about $3\,\mathrm{mV}$ and 10^4 second after stress can be observed (see Fig. 4.53). Interestingly, post-stress aging can be triggered by a strongly accelerated drain-source voltage and a gate-source voltage chosen slightly above the threshold voltage. The CET maps presented in Fig. 4.54 reveal the details of these findings.

The core p-MOSFET was in the focus of all research related to recovery effects in this work. Stress and recovery data for homogenous NBTI (see e.g. Fig. 4.34) have been recorded for various gate-source stress voltages (see Fig. 4.55), drain-source currents (see Fig. 4.35), drain-source read-out voltages (see Fig. 4.36), gate-source recovery voltages (see Fig. 4.38), accumulative pulses (see Fig. 4.40) and temperatures. To show the influence which each parameter has on the drift and recovery behavior, many data sets have been transformed to CET maps. In addition to that, aging and recovery data have been demonstrated not only for the threshold voltage but also for the drain-source current (see Fig. 4.41, Fig. 4.42). Furthermore, it has been revealed in Section 4.3.2 that the temperature time described in [181] cannot accurately describe the transformation of the data to a different temperature by a constant τ_0 because the differential CET map picture shows distinct differences (see Fig. 4.44), i.e. the generation and vanishing of traps of a certain $\tau_{\mathrm{s}}, \tau_{\mathrm{r}}$ class. One explanation could be that the barrier energy of the ensemble [181] changes with the temperature leading to a change of the probability of the recovery of defect classes for a given capture and emission time constant.

The statistical data of homogeneous NBTI experiments presented in Fig. 4.46 demonstrate that the variability of the threshold voltage drift decreases with the recovery time. On the other hand, this effect will be compensated for long stress times ($t_{\mathrm{s}} > 10^3\,\mathrm{s}$). For such long stress times the variability of the permanent component is larger compared to the variability of the recoverable component. This observation is also consistent with the results presented in Section 4.7. Here, for stress times in the range of $10^3\,\mathrm{s}$, the variability of the recoverable part of the threshold voltage superimposes onto the variability of the permanent component as shown in Fig. 4.70. In addition to that, the relative variability can be approximated with high accuracy by a power law of the stress time, including an offset of about 3%, which is depicted in Fig. 4.47. This offset is caused by the variability of the permanent traps which endure over a very long recovery time.

Especially, the interaction of HCD and NBTI was of high interest in this research work. That is why the 2-dimensional stress voltage matrix also contains many conditions to cover this degradation area which are discussed in Section 4.6. A drift minimum has been identified at $V_{\mathrm{ds,s}} = -0.5\,\mathrm{V}, V_{\mathrm{gs,s}} = -2.8\,\mathrm{V}$ and it was demonstrated that only the representation of the drift and recovery data via CET maps can show the shift of the NBTI related recoverable component to shorter stress and longer recovery times via an increase of the HCD contribution.

To assess the interplay of HCD and NBTI in more detail, statistical information has been recorded about the preexisting damage, degradation and recovery phase of both aging types (see Section 4.6.2). It was revealed that devices which were pre-stressed with HCD conditions show a smaller NBTI drift and also that the recovery of NBTI is impeded by HCD for a certain amount of time. Only a bake at 573 K (with the device not being connected to any voltage source) can remove this HCD related damage thus leading to the same threshold voltage drift of both experimental groups in the end ($t_{\mathrm{r,eff}} \approx 2 \times 10^{12}\,\mathrm{s}$, see Fig. 4.66).

The final analytical CET map model presented in Section 4.8 consists of parameters which depend on the technology and stress conditions. These stress conditions include the drain- and gate-source stress voltages, the stress and recovery time as well as the stress temperature (see Eq. 4.2 on page 100) and were also parametrized. Because the stress time was used to parametrize the four basic parameters of the S-shape curve, the stress time dependent fitting method of the CET map data (see Section 2.6.2) yields continuous CET maps which can be extrapolated to any reasonable stress and recovery time. The drift and recovery data of all characterized stress conditions including homogeneous and inhomogeneous NBTI as well as the mixture of HCD and NBTI can be described consistently by that CET map model.

The SPICE based aging simulations presented in Section 5.3 are based on worst-case aging models for HCD and NBTI. In particular, they facilitate a reuse of the aging models generated during a qualification of a transistor. In order to obtain more realistic drift results the worst-case HCD model was extended by a gate-source voltage and channel length dependence. As presented in Fig. 5.6 to Fig. 5.11 the aging model with a power law dependence of the stress time yields consistent results to the measurement data if no recovery is expected during the aging phase. Once recovery effects have to be taken into account the compact model from [196] as well as the microscopic TCAD model from the

TU Wien (see Section 1.2.3) lead to realistic estimations of the drift and recovery data of the core p-MOSFET. With respect to the computational effort the compact model is a highly promising candidate especially for SPICE based aging simulations.

7. Outlook

Some of the interesting aspects of the measured data, which were found through the last years, are not understood completely. In addition, ideas for future research to further challenge the current state of scientific knowledge include an extension of the CET map model presented in Section 4.8 with a broad range of stress and recovery temperatures. To allow a more effective recording of CET map data, the presented model should also be used to analyze datasets with fewer recovery traces, thus reducing the total measurement effort several times.

Also the physical modeling of the interplay of HCD and NBTI is still at an early stage, and, therefore, a unified model of both degradation mechanisms is of high interest to the whole reliability community. Interesting ideas are discussed, which e.g. take the change of the activation energy over stress time into account, but still many questions remain open.

The current physical NBTI models are based on a measurement delay after stress about 1 µs. To challenge these models measurements of stress and recovery data should be performed which employ a poly-heater setup to rapidly cool down the device under test at the end of the stress phase. This would yield a much shorter effective delay following the formulation of the temperature time discussed in [181] and Section 4.3.2.

8. Acknowledgments

This thesis presents the complex and challenging research in the field of single transistor reliability including recovery effects and its application to SPICE simulations. It has been conducted for about five years between 2012 and 2017 at Infineon Technologies AG (Neubiberg, Germany) and was supervised by Prof. Dr. Tibor Grasser (TU Wien, Austria). All investigations were in the context of three European funding projects (RELY, MoRV and RESIST) and have been represented to as well as discussed with the scientific community at various conferences (IEEE International Integrated Reliability Workshop (IIRW, 2013), 25th European Symposium on Reliability of Electron Devices, Failure Physics and Analysis (ESREF, 2014), User meeting of the ITG technical group 8.5.6 (2015, 2017), IEEE International Reliability Physics Symposium (IRPS, 2016)). The closest cooperations were in terms of the physical effects of hot-carrier degradation (HCD) and negative bias temperature instability (NBTI) with the Institute of Microelectronics (TU Wien) and in terms of a SPICE compatible compact model of NBTI considering recovery effects with the Fraunhofer Gesellschaft IIS/EAS (Dresden, Germany).

At this point I would like to thank Prof. Dr. Tibor Grasser (Head of Institute for Microelectronics, TU Wien) very much for his supervision of my research within the scope of this PhD thesis, all the interesting and pioneering technical discussions, his everlasting support and the always motivating work atmosphere. Especially his availability via eMail and telephone - which is very important for an external student - was of unprecedented scale. With pleasure I remember the personal meetings with him and also the PhD exchange meetings with the members of his working group.

In addition, I would like to thank my former line manager at Infineon Technologies AG, Dr. Wolfgang Gustin, who gave me the opportunity to work as a PhD student in his department between 2012 and 2014 and to finalize my research during the last years in combination with very interesting design for reliability related European funding projects (RELY, MoRV and RESIST). Without the great support of my two supervisors at Infineon, Dr. Hans Reisinger and Dr. Heiko Nielen, I think this research would have never come to the present state. At this, Hans was the expert in all discussions regarding device physics and measurement setups and Heiko the expert in aging simulations and higher abstraction levels.

Especially I would like to thank my wife, Karina, who was also working as a PhD student at the same department. Without her support and understanding this PhD thesis would not have been possible. Besides the personal time which I needed during the weekends and holidays for the finalization of this thesis she always was a brilliant discussion partner for me due to her knowledge and technical experience in the same field of research.

I also would like to thank the members of my former working group (IFAG BE QM REL DSD) and all other colleagues at Infineon who supported me during my PhD work and the related funding projects. Without the support of Dr. Günther Schindler the measurements with the standard lab equipment would not have been possible. His software suite (DART) enabled the semi-automatic measurements for all the entries of the 2-dimensional stress matrix at HCD conditions with the standard equipment. In addition to that, I would like to thank my former colleague Fabian Proebster who was a Bachelor student in our working group and measured the data of the core n-MOSFET presented in Fig. 5.11. Whenever a problem occurred with the probe station or other lab instruments it was Ulrich Brunner who always had time and a helping hand to solve it. Also, I would like to thank the team assistant of IFAG BE QM and REL, Sabine Bentenrieder, cordially for her support regarding the daily business (e.g. business trips, order process, etc.) and her constant effort for a good team atmosphere. Dr. Peter Rotter was the crucial interface from the reliability department to Infineon's design groups. He enabled the implementation of the worst-case aging models into SPICE assertions as described in Section 5.1 and always supported me when I had questions about the simulation flow. In this context, he was an excellent counterpart in the European funding projects RELY, MoRV and RESIST. In addition to that I would like to thank Dr. Karl Hofmann who supported the realization of the circuit presented in Section 2.9 and Georg Georgakos who was a valuable discussion partner in the whole DfR context and during the European funding projects RELY and RESIST. Furthermore, I would like to thank my colleague Dr. Stefano Aresu, a very experienced discussion partner in terms of technology qualifications and a great moral support through the last years.

Many members of the working group of Prof. Dr. Tibor Grasser were important discussion partners for the whole time of this PhD research. Therefore, I would like to thank (in place of the whole working group) Gerhard Rzepa for the interesting exchange and discussions about the modeling of CET maps, the variability of NBTI and the outstanding collaboration during the European project MoRV, Michael Waltl for the detailed discussions about measurement techniques of NBTI, Stanislav Tyaginov and Markus Jech for the discussions regarding all aspects of HCD and Bianka Ullmann for the exchange of ideas about inhomogeneous NBTI stress. I also would like to thank the

former group members Franz Schanovsky, Wolfgang Goes and Paul-Jürgen Wagner for the discussions at the initial phase of my research.

Within the scope of the European funding projects I had many contacts to researchers of other companies whom I also would like to thank (as representatives of their companies) for all the valuable discussions during and after the project phases, namely Roland Jancke and Kay-Uwe Giering (Fraunhofer Gesellschaft), Ben Kaczer (imec), Domenik Helms and Reef Eilers (OFFIS) as well as Christian Kernstock and Oskar Baumgartner (Global TCAD Solutions).

This work was supported by the European funding projects RELY (Design for RELIABILITY of SoCs for Applications like Transportation, Medical, and Industrial Automation, GA:01M3091), MoRV (Modelling reliability under variability, GA:619234) and RESIST (RESilient Integrated SysTems, BMBF F&E 16ES0301-16ES03010, CATRENE).

A. Additional figures

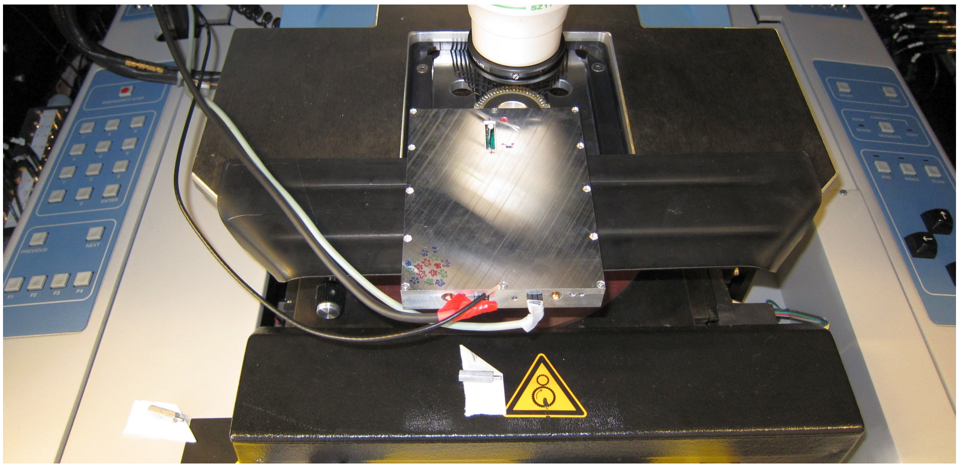


Figure A.1.: The ultra-fast measurement equipment is mounted on top of the probe card stage. To reduce the signal background noise all ground signals are connected to the hot-chuck wafer substrate terminal.

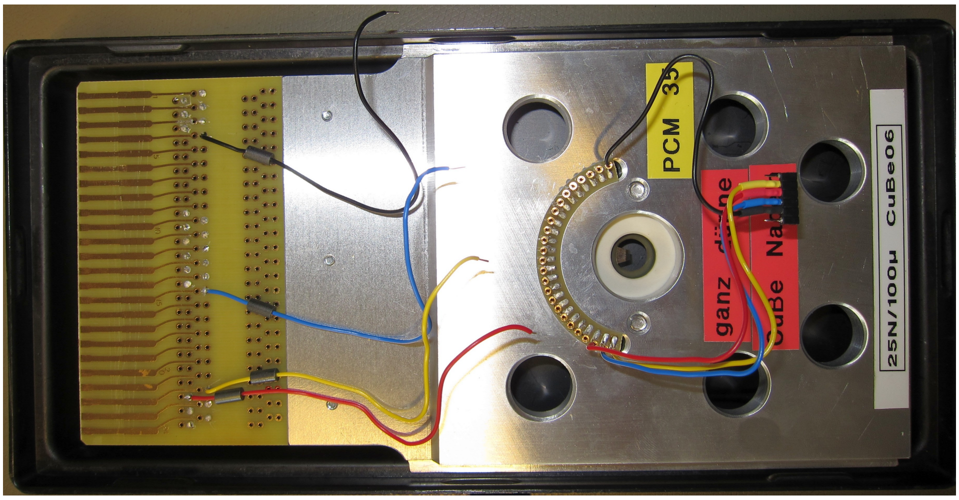
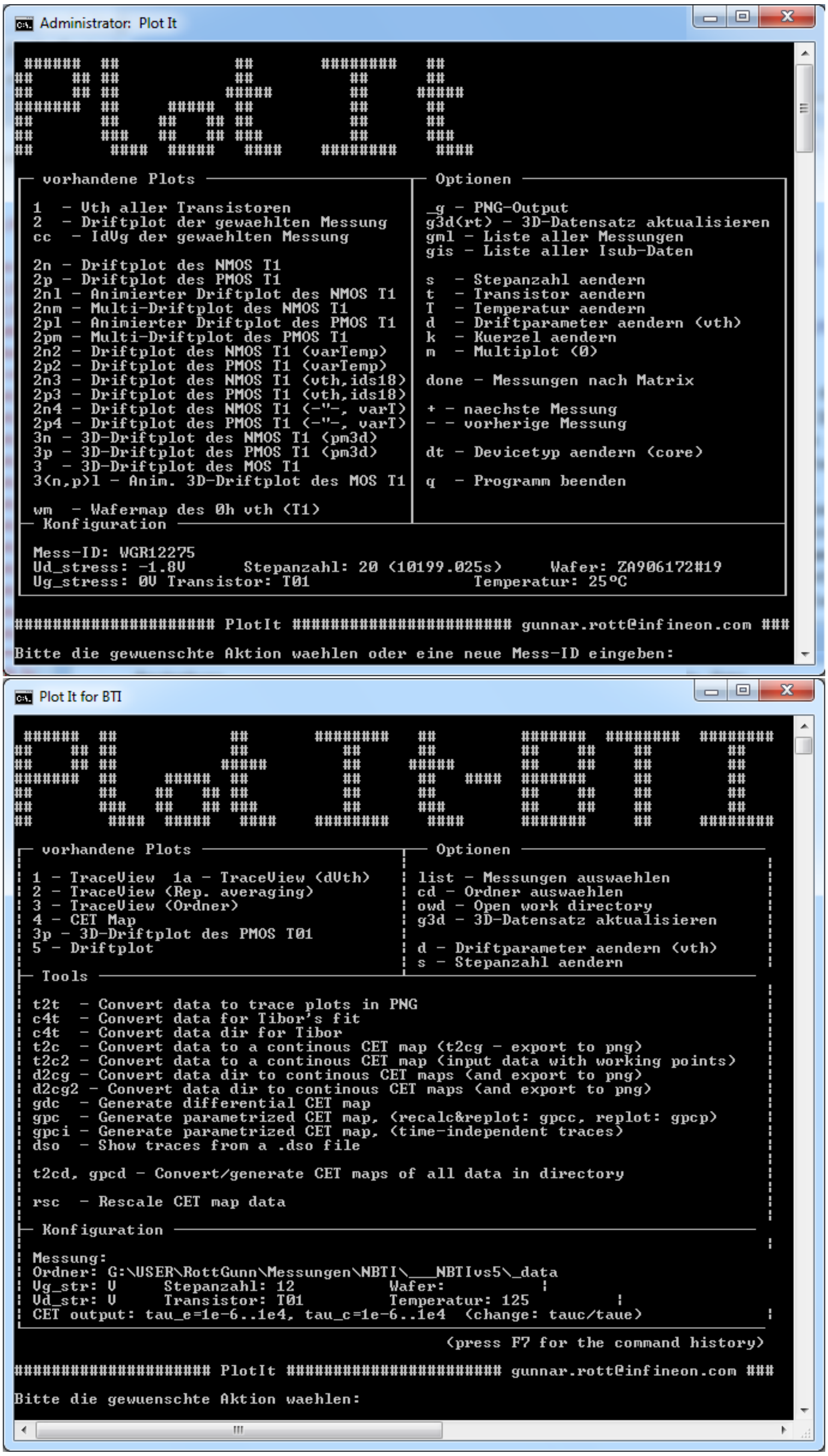


Figure A.2.: A multi-purpose probe card which can either connect the device under test to the standard lab equipment (left four wires soldered to the PC board of the probe card) or to the ultra-fast measurement equipment (right four wires with 6-pin interface block).



```
 ######  ##            ##       ########    ##
##    ## ##            ##           ##      ##
##    ## ##           #####         ##     #####
#######  ##     #####  ##           ##      ##
##       ##    ##   ## ##           ##      ##
##       ###   ##   ## ###          ##      ###
##        ####  #####   ####    ########     ####

 ┌ vorhandene Plots ─────────────────────────┬ Optionen ───────────────────────────┐
 │                                           │                                     │
 │ 1  - Uth aller Transistoren               │ _g - PNG-Output                     │
 │ 2  - Driftplot der gewaehlten Messung     │ g3d(rt) - 3D-Datensatz aktualisieren│
 │ cc  - IdUg der gewaehlten Messung         │ gml - Liste aller Messungen         │
 │                                           │ gis - Liste aller Isub-Daten        │
 │ 2n - Driftplot des NMOS T1                │                                     │
 │ 2p - Driftplot des PMOS T1                │ s  - Stepanzahl aendern             │
 │ 2nl - Animierter Driftplot des NMOS T1    │ t  - Transistor aendern             │
 │ 2nm - Multi-Driftplot des NMOS T1         │ T  - Temperatur aendern             │
 │ 2pl - Animierter Driftplot des PMOS T1    │ d  - Driftparameter aendern (vth)   │
 │ 2pm - Multi-Driftplot des PMOS T1         │ k  - Kuerzel aendern                │
 │ 2n2 - Driftplot des NMOS T1 (varTemp)     │ m  - Multiplot (0)                  │
 │ 2p2 - Driftplot des PMOS T1 (varTemp)     │                                     │
 │ 2n3 - Driftplot des NMOS T1 (vth,ids18)   │ done - Messungen nach Matrix        │
 │ 2p3 - Driftplot des PMOS T1 (vth,ids18)   │                                     │
 │ 2n4 - Driftplot des NMOS T1 (-"-, varT)   │ + - naechste Messung                │
 │ 2p4 - Driftplot des PMOS T1 (-"-, varT)   │ - - vorherige Messung               │
 │ 3n - 3D-Driftplot des NMOS T1 (pm3d)      │                                     │
 │ 3p - 3D-Driftplot des PMOS T1 (pm3d)      │ dt - Devicetyp aendern (core)       │
 │ 3  - 3D-Driftplot des MOS T1              │                                     │
 │ 3(n,p)l - Anim. 3D-Driftplot des MOS T1   │ q  - Programm beenden               │
 │                                           │                                     │
 │ wm  - Wafermap des 0h vth (T1)            │                                     │
 ├ Konfiguration ────────────────────────────┴─────────────────────────────────────┤
 │                                                                                 │
 │ Mess-ID: WGR12275                                                               │
 │ Ud_stress: -1.8V      Stepanzahl: 20 (10199.025s)      Wafer: ZA906172#19       │
 │ Ug_stress: 0V Transistor: T01                          Temperatur: 25°C         │
 └─────────────────────────────────────────────────────────────────────────────────┘

##################### PlotIt ###################### gunnar.rott@infineon.com ###

Bitte die gewuenschte Aktion waehlen oder eine neue Mess-ID eingeben:
```

```
 ######  ##            ##       ########  ##         #######  ########  ########
##    ## ##            ##           ##    ##         ##    ##    ##        ##
##    ## ##           #####         ##   #####       ##    ##    ##        ##
#######  ##     #####  ##           ##    ##   ####  #######     ##        ##
##       ##    ##   ## ##           ##    ##         ##    ##    ##        ##
##       ###   ##   ## ###          ##    ###        ##    ##    ##        ##
##        ####  #####   ####    ########   ####      #######     ##     ########

┌ vorhandene Plots ───────────────────────┬─ Optionen ─────────────────────────
¦                                         ¦                                     ¦
¦ 1 - TraceView  1a - TraceView (dVth)    ¦ list - Messungen auswaehlen         ¦
¦ 2 - TraceView (Rep. averaging)          ¦ cd - Ordner auswaehlen              ¦
¦ 3 - TraceView (Ordner)                  ¦ owd - Open work directory           ¦
¦ 4 - CET Map                             ¦ g3d - 3D-Datensatz aktualisieren    ¦
¦ 3p - 3D-Driftplot des PMOS T01          ¦                                     ¦
¦ 5 - Driftplot                           ¦ d - Driftparameter aendern (vth)    ¦
¦                                         ¦ s - Stepanzahl aendern              ¦
├ Tools ──────────────────────────────────┴────────────────────────────────────
¦                                                                               ¦
¦ t2t  - Convert data to trace plots in PNG                                     ¦
¦ c4t  - Convert data for Tibor's fit                                           ¦
¦ c4t  - Convert data dir for Tibor                                             ¦
¦ t2c  - Convert data to a continous CET map (t2cg - export to png)             ¦
¦ t2c2 - Convert data to a continous CET map (input data with working points)   ¦
¦ d2cg - Convert data dir to continous CET maps (and export to png)             ¦
¦ d2cg2 - Convert data dir to continous CET maps (and export to png)            ¦
¦ gdc  - Generate differential CET map                                          ¦
¦ gpc  - Generate parametrized CET map, (recalc&replot: gpcc, replot: gpcp)     ¦
¦ gpci - Generate parametrized CET map, (time-independent traces)               ¦
¦ dso  - Show traces from a .dso file                                           ¦
¦                                                                               ¦
¦ t2cd, gpcd - Convert/generate CET maps of all data in directory               ¦
¦                                                                               ¦
¦ rsc  - Rescale CET map data                                                   ¦
¦                                                                               ¦
├ Konfiguration ───────────────────────────────────────────────────────────────
¦                                                                               ¦
¦ Messung:
¦ Ordner: G:\USER\RottGunn\Messungen\NBTI\___NBTIvs5\_data
¦ Ug_str: V     Stepanzahl: 12          Wafer:          ¦
¦ Ud_str: V     Transistor: T01         Temperatur: 125        ¦
¦ CET output: tau_e=1e-6..1e4, tau_c=1e-6..1e4  (change: tauc/taue)             ¦
└──────────────────────────────────────────────────────────────────────────────
                                            (press F7 for the command history)

##################### PlotIt ###################### gunnar.rott@infineon.com ###

Bitte die gewuenschte Aktion waehlen:
```

Figure A.3.: Main window of PlotIt and PlotIt-BTI, two software environments developed during the course of this thesis. They are based on open-source tools and post-process the recorded measurement data as well as visualize the complex multi-dimensional space of device aging. Due to their flexible program structure both environments can easily be adapted to each experimental setup in this work.

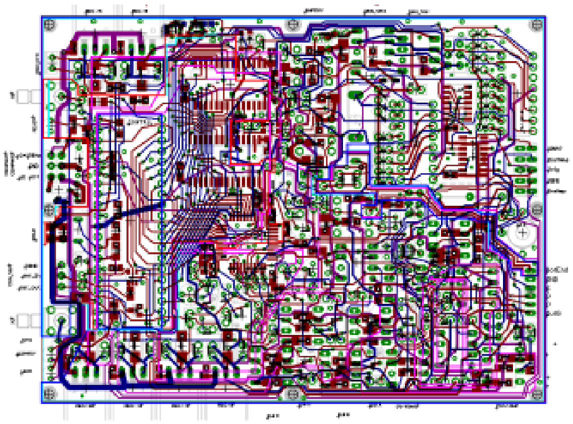


Figure A.4.: The layout of the PCB of the ultra-fast measurement equipment shows four different metal layers and includes more than 100 single components.

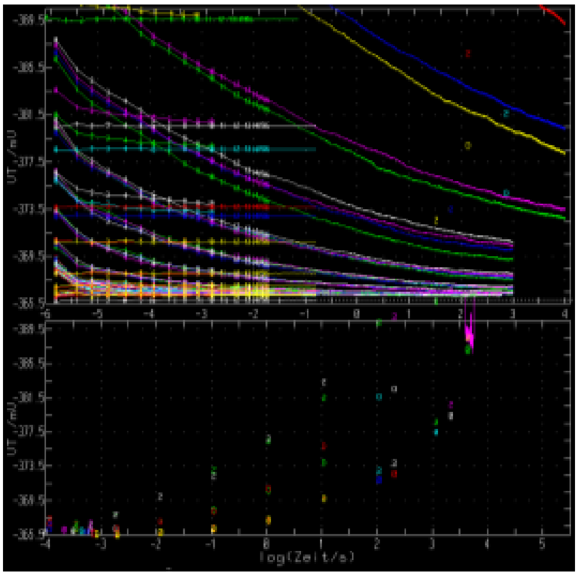


Figure A.5.: A test measurement performed on the dual gate oxide p-MOSFET of the examined smart power technology test module. The top of the graph shows the recorded recovery traces and the bottom the threshold voltage drift over stress time.

B. Additional tables

t_s / s	t_r / s	**recorded data points**
0	10^{0}	612
10^{-6}	10^{1}	1710
10^{-5}	10^{1}	1710
10^{-4}	10^{1}	1710
10^{-3}	10^{1}	1710
10^{-2}	10^{2}	3804
10^{-1}	10^{3}	6666
10^{0}	10^{4}	9726
10^{1}	10^{4}	9726
10^{2}	10^{4}	9726
10^{3}	10^{4}	9726
10^{4}	10^{4}	9726

Table B.1.: Timing and data density of a full CET map dataset of the core p-MOSFET stressed at 398 K. The recovery times were chosen to allow a reasonable recovery of the collected stress. Starting from $t_\mathrm{s} = 1\,\mathrm{s}$ a full recovery of the threshold voltage was not possible within an acceptable time. Consequently, the limitation of $t_\mathrm{r} < 1 \times 10^{4}\,\mathrm{s}$ for longer stress times causes a build-up of interface states with large recovery time constants, which leads to an offset of the recorded recovery traces.

step	mode	duration / s	meas. par.
0	measure	10^0	V_{th}
1	measure	1.8×10^{-5}	$I_{\mathrm{ds,lin}}$
2	measure	1.3×10^{-5}	$I_{\mathrm{ds,ana}}$
3	measure	1.1×10^{-5}	$I_{\mathrm{ds,sat}}$
4	stress	1.0×10^{-6}	-
5	measure	1.8×10^{-6}	V_{th}
6	measure	1.3×10^{-5}	$I_{\mathrm{ds,ana}}$
7	measure	1.1×10^{-5}	$I_{\mathrm{ds,sat}}$
8	measure	1.8×10^{-5}	$I_{\mathrm{ds,lin}}$
9	measure	1×10^{1}	V_{th}
10	measure	3.3×10^{-3}	$I_{\mathrm{ds,ana}}$
11	measure	3.3×10^{-3}	$I_{\mathrm{ds,sat}}$
12	measure	3.3×10^{-3}	$I_{\mathrm{ds,lin}}$
13	stress	1.0×10^{-5}	-
14	measure	1.8×10^{-6}	V_{th}
15	measure	1.1×10^{-5}	$I_{\mathrm{ds,sat}}$
16	measure	1.8×10^{-5}	$I_{\mathrm{ds,lin}}$
17	measure	1.3×10^{-5}	$I_{\mathrm{ds,ana}}$
18	measure	1.0×10^{1}	V_{th}
19	measure	3.3×10^{-3}	$I_{\mathrm{ds,sat}}$
20	measure	3.3×10^{-3}	$I_{\mathrm{ds,lin}}$
21	measure	3.3×10^{-3}	$I_{\mathrm{ds,ana}}$
22	stress	1.0×10^{-4}	-
23	measure	1.8×10^{-6}	V_{th}
24	measure	1.8×10^{-5}	$I_{\mathrm{ds,lin}}$
25	measure	1.3×10^{-5}	$I_{\mathrm{ds,ana}}$
26	measure	1.1×10^{-5}	$I_{\mathrm{ds,sat}}$
27	measure	1.0×10^{1}	V_{th}
28	measure	3.3×10^{-3}	$I_{\mathrm{ds,lin}}$
29	measure	3.3×10^{-3}	$I_{\mathrm{ds,ana}}$
30	measure	3.3×10^{-3}	$I_{\mathrm{ds,sat}}$
...			

Table B.2.: First lines of a recipe of a CET map measurement to record threshold voltage and operating point recovery information of the same device under test. Due to the alternation of the operating point measurements the recovery caused by the settling time of the SMU can be subtracted out for the second and third operating point.

C. List of own publications

- B. Ullmann, M. Jech, K. Puschkarsky, G. Rott, M. Waltl, Y. Illarionov, H. Reisinger, T. Grasser "Impact of Mixed Negative Bias Temperature Instability and Hot Carrier Stress on MOSFET Characteristics - Part I: Experimental". Submitted to IEEE Transactions on Electron Devices, 2018.
- K.-U. Giering, G. Rott, G. Rzepa, H. Reisinger, A.K. Puppala, T. Reich, W. Gustin, T. Grasser, R. Jancke "Analog-circuit NBTI degradation and time-dependent NBTI variability: An efficient physics-based compact model". In: IEEE International Reliability Physics Symposium (IRPS), 2016.
- M. Metzdorf, R. Eilers, D. Helms, K.-U. Giering, R. Jancke, G. Rzepa, T. Grasser, M. Karner, P. Raghavan, D. Alexandrescu, A. Evans, G. Rott, H. Reisinger, W. Gustin "MoRV: Modelling Reliability under Variability". Silicon Errors in Logic – System Effects (SELSE) 12, 2016.
- G. Rzepa, W. Goes, G. Rott, K. Rott, M. Karner, C. Kernstock, B. Kaczer, H. Reisinger, T. Grasser "Physical Modeling of NBTI: From Individual Defects to Devices". International Conference on Simulation of Semiconductor Processes and Devices (SISPAD), 2014.
- T. Grasser, G. Rzepa, M. Waltl, W. Goes, K. Rott, G. Rott, H. Reisinger, J. Franco, B. Kaczer "Characterization and Modeling of Charge Trapping: From Single Defects to Devices" International Conference on IC Design and Technology (ICICDT), 2014.
- G.A. Rott, K. Rott, H. Reisinger, W. Gustin, T. Grasser "Mixture of Negative Bias Temperature Instability and Hot-Carrier driven Threshold Voltage Degradation of 130nm Technology P-Channel Transistors". Microelectronics Reliability 54, pp. 2310–2314, 2014.
- G.A. Rott, H. Nielen, H. Reisinger, W. Gustin, S. Tyaginov, T. Grasser "Drift Compensating Effect during Hot-Carrier Degradation of 130nm Technology Dual Gate Oxide P-Channel Transistors". IEEE International Integrated Reliability Workshop Final Report (IIRW), pp. 73-77, 2013
- K. Rott, D. Schmitt-Landsiedel, H. Reisinger, G. Rott, G. Georgakos, C. Schlünder, S. Aresu, W. Gustin, T. Grasser "Impact and measurement of short term

threshold instabilities in MOSFETs of analog circuits". IEEE International Integrated Reliability Workshop Final Report, pp. 31-34., 2012

- G.A. Rott, F. Zhang, Y. Haba, W. Kröger, E. Burkel "Dielectric Properties of Porous Calcium Titanate ($CaTiO_3$)". Biomaterials Science - Processing, Properties and Applications: Ceramic Transactions, 228, 2011.
- F. Zhang, E. Otterstein, G.A. Rott, U. Beck, D.G. Weiss, E. Burkel. "Preparation and surface modification of TiMn foams for bone implants". Biomedical Engineering, 55 (1), 2010.
- G.A. Rott, F. Zhang, Y. Haba, E. Burkel "Dielektrische Eigenschaften von porösem Calciumtitanat ($CaTiO_3$)". Biomed Tech, 55 (Suppl. 1), 2010.
- G.A. Rott, F. Zhang, Y. Haba and E. Burkel "Dielectric properties of porous calcium titanate". Biomedical Engineering, 55 (1), 2010.
- G.A. Rott, F. Zhang, Y. Haba, W. Kröger and E. Burkel "Dielectric Spectroscopy and Microstructure of Sintered Calcium Titanate ($CaTiO_3$) Samples With Different Porosities". Material Science and Engineering (Darmstadt), 2010.
- U. van Rienen, R. Appali, R. Bader, W. Baumann, U. Beck, D. Behrend, R. Benecke, K. Biala, H. Birkholz, E. Burkel, G. Engel, J. Gimsa, U. Gimsa, E. Gongadze, A. Grünbaum, Y. Haba, F. Liese, B. Liu, M. Lüder, C. Matschegewski, W. Mittelmeier, E. Mix, J.B. Nebe, K.A. Nowak, J. Pahnke, H.W. Pau, S. Pauleweit, S. Petersen, S. Polnick, T. Reimer, G.A. Rott, R. Salomon, E. Vinter, T. Weihe and O. Wolkenhauer, "Analysis and Simulation of Electrical Interactions of Implants with Biosystems". Biomedical Engineering, 55 (1), pp. 167–170, 2010.

D. Talks at conferences

- VDE-ITG 8.5.6 2017: “Continuous^2 Capture-Emission-Time Maps for the interaction of HCD and NBTI including recovery effects”
- VDE-ITG 8.5.6 2015: “Mischung von Hot-Carrier Degradation und Negative Bias Temperature Instability in p-MOSFETs einer 130nm Technologie”
- European Symposium on Reliability of Electron Devices, Failure Physics and Analysis (ESREF) 2014: “Mixture of Negative Bias Temperature Instability and Hot-Carrier driven Threshold Voltage Degradation of 130nm Technology P-Channel Transistors”
- IEEE International Integrated Reliability Workshop (IIRW) 2013: “Drift Compensating Effect during Hot-Carrier Degradation of 130nm Technology Dual Gate Oxide P-Channel Transistors”
- VDE-BioMedTec (BMT) 2011: “Dielectric properties of porous calcium titanate”
- Materials Science and Engineering (MSE) 2010: “Dielectric Spectroscopy and Microstructure of Sintered Calcium Titanate ($CaTiO_3$) Samples With Different Porosities”
- Materials Science & Technology (MS&T) 2010: “Dielectric Properties of Porous Calcium Titanate ($CaTiO_3$)”

References

[1] Roger Newman. “Visible Light from a Silicon $p-n$ Junction”. In: *Phys. Rev.* 100 (2 Oct. 1955), pp. 700–703.

[2] M. Herzog and F. Koch. “Hot-carrier light emission from silicon metal-oxide-semiconductor devices”. In: *Applied Physics Letters* 53.26 (1988), pp. 2620–2622.

[3] L. Vivien and L. Pavesi. *Handbook of Silicon Photonics*. Series in Optics and Optoelectronics. CRC Press, 2016.

[4] E. Macii, V. Paliouras, and O. Koufopavlou. *Integrated Circuit and System Design: Power and Timing Modeling, Optimization and Simulation; 14th International Workshop, PATMOS 2004, Santorini, Greece, September 15-17, 2004, Proceedings*. Lecture Notes in Computer Science Bd. 14. Springer, 2004.

[5] Eiji Takeda, Norio Suzuki, and Takaaki Hagiwara. “Device performance degradation to hot-carrier injection at energies below the $Si - SiO_2$ energy barrier”. In: *Electron Devices Meeting, 1983 International*. Vol. 29. IEEE. 1983, pp. 396–399.

[6] P. E. Cotrell, R. R. Troutman, and T. H. Ning. “Hot-electron emission in n-channel IGFETs”. In: *IEEE Transactions On Electron Devices, Vol. ED-26* (1979), pp. 520–533.

[7] S. Tam, P. K. Ko, and C. Hu. “Lucky-electron model of channel hot electron injection in MOSFET’s”. In: *IEEE Transactions On Electron Devices* (1984), pp. 1116–1125.

[8] M. S. Liang et al. “Hot carriers induced degradation in thin gate oxide MOSFETs”. In: *Electron Devices Meeting, 1983 International*. Vol. 29. 1983, pp. 186–189.

[9] Susumu Kohyama et al. “Non-Thermal Carrier Generation in MOS Structures”. In: *Japanese Journal of Applied Physics* 19.S1 (1980), p. 85.

[10] S. Tyaginov et al. “Understanding and Modeling the Temperature Behavior of Hot-Carrier Degradation in SiON nMOSFETs”. In: *IEEE Electron Device Letters* 37.1 (Jan. 2016), pp. 84–87.

[11] K. Hess et al. “The physics of determining chip reliability”. In: *Circuits and Devices Magazine, IEEE* 17.3 (May 2001), pp. 33–38.

[12] William McMahon et al. “From Atoms to Circuits: Theoretical and Empirical Modeling of Hot Carrier Degradation”. In: *Hot Carrier Degradation in Semiconductor Devices*. Ed. by Tibor Grasser. New York: Springer International Publishing, 2014, pp. 3–27.

[13] Alain Bravaix et al. “Hot-Carrier Degradation in Decananometer CMOS Nodes: From an Energy-Driven to a Unified Current Degradation Modeling by a Multiple-Carrier Degradation Process”. In: *Hot Carrier Degradation in Semiconductor Devices*. Ed. by Tibor Grasser. New York: Springer International Publishing, 2014, pp. 57–103.

[14] Stanislav Tyaginov. “Physics-Based Modeling of Hot-Carrier Degradation”. In: *Hot Carrier Degradation in Semiconductor Devices*. Ed. by Tibor Grasser. New York: Springer International Publishing, 2014, pp. 105–150.

[15] C. Hu et al. “Hot-Electron-Induced MOSFET Degradation-Model, Monitor, and Improvement”. In: *IEEE Transactions On Electron Devices, Vol. ED-32, No. 2* (1985), pp. 375–385.

[16] H. Gesch, J. P. Leburton, and G. E. Dorda. “Generation of interface states by hot-hole injection in MOSFET’s”. In: *IEEE Transactions On Electron Devices, Vol. ED-29* (1982), p. 913.

[17] E. Takeda and N. Suzuki. “An empirical model for device degradation due to hot-carrier injection”. In: *IEEE Electron Device Letters, Vol. EDL-4* (1983), pp. 111–113.

[18] F. C. Hsu and S. Tam. “Relationship between MOSFET degradation and hot-electron induced interface-state generation”. In: *IEEE Electron Device Letters, Vol. EDL-5* (1984), p. 50.

[19] K. O. Jeppson and C. M. Svensson. “Negative bias stress of MOS devices at high electric fields and degradation of NMOS devices”. In: *Journal of Applied Physics, Vol. 48* (1977), pp. 2004–2014.

[20] C. Guerin, V. Huard, and A. Bravaix. “General framework about defect creation at the Si-SiO2 interface”. In: *Journal of Applied Physics* 105.11, 114513 (2009), pp. 1–12.

[21] B.N.J. Persson and Ph. Avouris. “Local bond breaking via STM-induced excitations: the role of temperature”. In: *Surface Science* 390.1–3 (1997), pp. 45–54.

[22] K. L. Brower. “Dissociation kinetics of hydrogen-passivated (111) Si-SiO$_2$ interface defects”. In: *Phys. Rev. B* 42 (6 Aug. 1990), pp. 3444–3453.

[23] A. Stesmans. “Passivation of Pb0 and Pb1 interface defects in thermal (100) Si/SiO2 with molecular hydrogen”. In: *Applied Physics Letters* 68.15 (1996), pp. 2076–2078.

[24] T. H. Ning et al. “1um MOSFET VLSI technology: Part IV - Hot-electron design constraints”. In: *IEEE Transactions On Electron Devices, Vol. ED-26* (1979), pp. 346–353.

[25] W. Shockley. “Problems related to p-n-junctions in silicon”. In: *Solid State Electronics, Vol. 2* (1961), pp. 35–67.

[26] C. Hu. “Hot-electron effects in MOSFETs”. In: *Electron Devices Meeting, 1983 International*. Vol. 29. 1983, pp. 176–181.

[27] Y. A. El Mansy and A. R. Boothroyd. “A simple two-dimensional model for IGFET operation in the saturation region”. In: *IEEE Transactions On Electron Devices, Vol. ED-24* (1977), p. 254.

[28] P.K. Ko, R.S. Muller, and C. Hu. “A unified model for hot-electron currents in MOSFET’s”. In: *Electron Devices Meeting, 1981 International*. Vol. 27. 1981, pp. 600–603.

[29] T. N. Nguyen and J. D. Plummer. “Physical mechanisms responsible for short channel effects in MOS devices”. In: *Electron Devices Meeting, 1981 International*. Vol. 27. 1981, pp. 596–599.

[30] K. Hess et al. “Theory of channel hot-carrier degradation in {MOSFETs}”. In: *Physica B: Condensed Matter* 272.1-4 (1999), pp. 527–531.

[31] T. Erdogan et al. “Decay of ultraviolet-induced fiber Bragg gratings”. In: *Journal of Applied Physics* 76.1 (1994), pp. 73–80.

[32] A. Stesmans. “Structural relaxation of P_b defects at the (111)Si/SiO$_2$ interface as a function of oxidation temperature: The P_b-generation - stress relationship”. In: *Phys. Rev. B* 48 (4 1993), pp. 2418–2435.

[33] B. Tuttle and C. G. Van de Walle. “Structure, energetics, and vibrational properties of Si-H bond dissociation in silicon”. In: *Phys. Rev. B* 59 (20 1999), pp. 12884–12889.

[34] A. Haggag et al. “High-performance chip reliability from short-time-tests-statistical models for optical interconnect and HCI/TDDB/NBTI deep-submicron transistor failures”. In: *Reliability Physics Symposium, 2001. Proceedings. 39th Annual. 2001 IEEE International*. 2001, pp. 271–279.

[35] W. McMahon, A. Haggag, and K. Hess. “Reliability scaling issues for nanoscale devices”. In: *Nanotechnology, IEEE Transactions on* 2.1 (Mar. 2003), pp. 33–38.

[36] T. -C. Shen et al. “Atomic-Scale Desorption Through Electronic and Vibrational Excitation Mechanisms”. In: *Science* 268.5217 (1995), pp. 1590–1592.

[37] Yuan Taur and Tak H Ning. *Fundamentals of modern VLSI devices.* Cambridge university press, 2009.

[38] W. McMahon et al. “The effects of a multiple carrier model of interface trap generation on lifetime extraction for MOSFETs”. In: *MSM Conf., Puerto Rico.* 2002, pp. 576–579.

[39] K. Hess et al. “Impact of nanostructure research on conventional solid-state electronics: The giant isotope effect in hydrogen desorption and CMOS lifetime”. In: *Physica E: Low-dimensional Systems and Nanostructures* 3.1–3 (1998), pp. 1–7.

[40] Stanislav E Tyaginov et al. “Physics-Based Hot-Carrier Degradation Models”. In: *ECS Transactions* 35.4 (2011), pp. 321–352.

[41] Oleg Penzin et al. “MOSFET degradation kinetics and its simulation”. In: *Electron Devices, IEEE Transactions on* 50.6 (2003), pp. 1445–1450.

[42] N. Watanabe et al. “Structural origin of carrier trap levels at SiO_2 Si interface”. In: *Proc. Extended Abstracts 1998 Int. Conf. Solid State Device Mater.* 1998, pp. 130–131.

[43] A. Plonka. *Time-dependent reactivity of species in condensed media.* Springer-Verlag New York Inc.,New York, NY, Jan. 1986.

[44] F. Ootsuka. “The evaluation of the activation energy of interface state generation by hot-electron injection”. In: *Electron Devices, IEEE Transactions on* 38.6 (June 1991), pp. 1477–1483.

[45] J.E. Chung, P.-K. Ko, and Chenming Hu. “A model for hot-electron-induced MOSFET linear-current degradation based on mobility reduction due to interface-state generation”. In: *Electron Devices, IEEE Transactions on* 38.6 (June 1991), pp. 1362–1370.

[46] Paul Heremans et al. “Consistent model for the hot-carrier degradation in n-channel and p-channel MOSFETs”. In: *Electron Devices, IEEE Transactions on* 35.12 (Dec. 1988), pp. 2194–2209.

[47] J. W. McPherson, R. B. Khamankar, and A. Shanware. “Complementary model for intrinsic time-dependent dielectric breakdown in SiO2 dielectrics”. In: *Journal of Applied Physics* 88.9 (2000), pp. 5351–5359.

[48] J. Albohn et al. “Capture cross sections of defect states at the Si/SiO2 interface”. In: *Journal of Applied Physics* 88.2 (2000), pp. 842–849.

[49] R. A. B. Devine et al. “Interfacial hardness enhancement in deuterium annealed 0.25 μm channel metal oxide semiconductor transistors”. In: *Applied Physics Letters* 70.22 (1997), pp. 2999–3001.

[50] Zhi Chen et al. “On the mechanism for interface trap generation in MOS transistors due to channel hot carrier stressing”. In: *Electron Device Letters, IEEE* 21.1 (Jan. 2000), pp. 24–26.

[51] Haldun Kufluoglu and Muhammad Ashraful Alam. “A Computational Model of NBTI and Hot Carrier Injection Time-Exponents for MOSFET Reliability”. In: *Journal of Computational Electronics* 3.3-4 (2004), pp. 165–169.

[52] H. Kufluoglu and M.A. Alam. “A geometrical unification of the theories of NBTI and HCI time-exponents and its implications for ultra-scaled planar and surround-gate MOSFETs”. In: *Electron Devices Meeting, 2004. IEDM Technical Digest. IEEE International.* Dec. 2004, pp. 113–116.

[53] Kjell O. Jeppson and Christer M. Svensson. “Negative bias stress of MOS devices at high electric fields and degradation of MNOS devices”. In: *Journal of Applied Physics* 48.5 (1977), pp. 2004–2014.

[54] S. Mahapatra, P.B. Kumar, and M.A. Alam. “A new observation of enhanced bias temperature instability in thin gate oxide p-MOSFETs”. In: *Electron Devices Meeting, 2003. IEDM ’03 Technical Digest. IEEE International.* Dec. 2003, pp. 14.2.1–14.2.4.

[55] S.E. Rauch and G. La Rosa. “The energy-driven paradigm of NMOSFET hot-carrier effects”. In: *Device and Materials Reliability, IEEE Transactions on* 5.4 (Dec. 2005), pp. 701–705.

[56] A. Bravaix et al. “Hot-Carrier acceleration factors for low power management in DC-AC stressed 40nm NMOS node at high temperature”. In: *Reliability Physics Symposium, 2009 IEEE International.* Apr. 2009, pp. 531–548.

[57] K. Stokbro et al. “STM-Induced Hydrogen Desorption via a Hole Resonance”. In: *Phys. Rev. Lett.* 80 (12 Mar. 1998), pp. 2618–2621.

[58] K. Stokbro et al. “First-principles theory of inelastic currents in a scanning tunneling microscope”. In: *Phys. Rev. B* 58 (12 Sept. 1998), pp. 8038–8041.

[59] T. -C. Shen et al. “Atomic-Scale Desorption Through Electronic and Vibrational Excitation Mechanisms”. In: 268.5217 (1995), pp. 1590–1592.

[60] In: *Fujitsu Sci. Tech. J.* 39.1 (2003), pp. 106–118.

[61] Chris G. Van de Walle and R. A. Street. “Structure, energetics, and dissociation of Si-H bonds at dangling bonds in silicon”. In: *Phys. Rev. B* 49 (20 May 1994), pp. 14766–14769.

[62] Blair Tuttle and Chris G. Van de Walle. “Structure, energetics, and vibrational properties of Si-H bond dissociation in silicon”. In: *Phys. Rev. B* 59 (20 May 1999), pp. 12884–12889.

[63] Blair Tuttle. “Hydrogen and P_b defects at the (111)$\mathrm{Si} - \mathrm{SiO_2}$ interface: An *ab initio* cluster study”. In: *Phys. Rev. B* 60 (4 July 1999), pp. 2631–2637.

[64] Ivan Andrianov and Peter Saalfrank. “Theoretical study of vibration-phonon coupling of H adsorbed on a Si(100) surface”. In: *The Journal of Chemical Physics* 124.3 (2006), p. 034710.

[65] B. Sun et al. “Vibrational Lifetimes and Frequency-Gap Law of Hydrogen Bending Modes in Semiconductors”. In: *Phys. Rev. Lett.* 96 (3 Jan. 2006), p. 035501.

[66] B.S. Doyle et al. “The generation and characterization of electron and hole traps created by hole injection during low gate voltage hot-carrier stressing of n-MOS transistors”. In: *Electron Devices, IEEE Transactions on* 37.8 (Aug. 1990), pp. 1869–1876.

[67] Dominique Vuillaume and Alain Bravaix. “Charging and discharging properties of electron traps created by hot-carrier injections in gate oxide of n-channel metal oxide semiconductor field effect transistor”. In: *Journal of Applied Physics* 73.5 (1993), pp. 2559–2563.

[68] Yoann Mamy Randriamihaja et al. “Microscopic scale characterization and modeling of transistor degradation under {HC} stress”. In: *Microelectronics Reliability* 52.11 (2012), pp. 2513–2520.

[69] SE Tyaginov et al. “Hot-carrier degradation modeling using full-band Monte-Carlo simulations”. In: *Physical and Failure Analysis of Integrated Circuits (IPFA), 2010 17th IEEE International Symposium on the*. IEEE. 2010, pp. 1–5.

[70] S.E. Tyaginov et al. “Interface traps density-of-states as a vital component for hot-carrier degradation modeling”. In: *Microelectronics Reliability* 50.9–11 (2010). 21st European Symposium on the Reliability of Electron Devices, Failure Physics and Analysis, pp. 1267–1272. ISSN: 0026-2714.

[71] IA Starkov et al. “Analysis of worst-case hot-carrier conditions for high voltage transistors based on full-band monte-carlo simulations”. In: *Physical and Failure Analysis of Integrated Circuits (IPFA), 2010 17th IEEE International Symposium on the*. IEEE. 2010, pp. 1–6.

[72] I. Starkov et al. “Hot-carrier degradation caused interface state profile - Simulation versus experiment”. In: *Journal of Vacuum Science & Technology B* 29.1, 01AB09 (2011), pp. 1–8.

[73] Christoph Jungemann and Bernd Meinerzhagen. *Hierarchical Device Simulation.* Springer-Verlag Wien, 2003.

[74] K. Rupp et al. “Inclusion of carrier-carrier-scattering into arbitrary-order spherical harmonics expansions of the Boltzmann transport equation”. In: *Computational Electronics (IWCE), 2012 15th International Workshop on.* May 2012, pp. 1–4.

[75] S. Tyaginov et al. “Physical modeling of hot-carrier degradation for short- and long-channel MOSFETs”. In: *Reliability Physics Symposium, 2014 IEEE International.* June 2014, XT.16.1–XT.16.8.

[76] Karl Rupp, Tibor Grasser, and Ansgar Jüngel. “On the feasibility of spherical harmonics expansions of the Boltzmann transport equation for three-dimensional device geometries”. In: *IEDM Techn. Digest* (2011).

[77] M. Bina et al. “Modeling of hot carrier degradation using a spherical harmonics expansion of the bipolar Boltzmann transport equation”. In: *Electron Devices Meeting (IEDM), 2012 IEEE International.* Dec. 2012, pp. 30.5.1–30.5.4.

[78] S. Tyaginov et al. “(Late) Essential ingredients for modeling of hot-carrier degradation in ultra-scaled MOSFETs”. In: *Integrated Reliability Workshop Final Report (IRW), 2013 IEEE International.* Oct. 2013, pp. 98–101.

[79] Sung-Min Hong, Anh-Tuan Pham, and Christoph Jungemann. *Deterministic solvers for the Boltzmann transport equation.* Springer Science & Business Media, 2011.

[80] Antonio Gnudi, Davide Ventura, and Giorgio Baccarani. “One-dimensional simulation of a bipolar transistor by means of spherical harmonics expansion of the Boltzmann transport equation”. In: *Proc. SISDEP.* Vol. 4. 1991, pp. 205–213.

[81] A. Gnudi et al. “Two-dimensional MOSFET simulation by means of a multidimensional spherical harmonics expansion of the Boltzmann transport equation”. In: *Solid-State Electronics* 36.4 (1993), pp. 575–581.

[82] Andreas Kerber and Eduard Cartier. “Bias Temperature Instability Characterization Methods”. In: *Bias Temperature Instability for Devices and Circuits.* Ed. by Tibor Grasser. New York: Springer Science & Business Media, 2013, pp. 3–31.

[83] Thomas Aichinger, Gregor Pobegen, and Michael Nelhiebel. “Application of On-Chip Device Heating for BTI Investigations”. In: *Bias Temperature Instability for Devices and Circuits*. Ed. by Tibor Grasser. New York: Springer Science & Business Media, 2013, pp. 33–51.

[84] Tahui Wang, Jung-Piao Chiu, and Yu-Heng Liu. “Statistical Characterization of BTI-Induced High-k Dielectric Traps in Nanoscale Transistors”. In: *Bias Temperature Instability for Devices and Circuits*. Ed. by Tibor Grasser. New York: Springer Science & Business Media, 2013, pp. 53–74.

[85] T. Grasser et al. “A Unified Perspective of RTN and BTI”. In: *Reliability Physics Symposium (IRPS), 2014 IEEE International*. 2014.

[86] Stewart E. Rauch III. “BTI-Induced Statistical Variations”. In: *Bias Temperature Instability for Devices and Circuits*. Ed. by Tibor Grasser. New York: Springer Science & Business Media, 2013, pp. 135–160.

[87] B. Kaczer et al. “Statistical Distribution of Defect Parameters”. In: *Bias Temperature Instability for Devices and Circuits*. Ed. by Tibor Grasser. New York: Springer Science & Business Media, 2013, pp. 161–176.

[88] Jason P. Campbell and Patrick M. Lenahan. “Atomic-Scale Defects Associated with the Negative Bias Temperature Instability”. In: *Bias Temperature Instability for Devices and Circuits*. Ed. by Tibor Grasser. New York: Springer Science & Business Media, 2013, pp. 177–228.

[89] V.V. Afanas’ev, M. Houssa, and A. Stesmans. “Charge Properties of Paramagnetic Defects in Semiconductor/Oxide Structures”. In: *Bias Temperature Instability for Devices and Circuits*. Ed. by Tibor Grasser. New York: Springer Science & Business Media, 2013, pp. 229–252.

[90] Jian F. Zhang. “Oxide Defects”. In: *Bias Temperature Instability for Devices and Circuits*. Ed. by Tibor Grasser. New York: Springer Science & Business Media, 2013, pp. 253–285.

[91] Diing Shenp Ang. “Understanding Negative-Bias Temperature Instability from Dynamic Stress Experiments”. In: *Bias Temperature Instability for Devices and Circuits*. Ed. by Tibor Grasser. New York: Springer Science & Business Media, 2013, pp. 287–302.

[92] Gregor Pobegen, Thomas Aichinger, and Michael Nelhiebel. “Impact of Hydrogen on the Bias Temperature Instability”. In: *Bias Temperature Instability for Devices and Circuits*. Ed. by Tibor Grasser. New York: Springer Science & Business Media, 2013, pp. 485–505.

[93] Souvik Mahapatra. “A Comprehensive Modeling Framework for DC and AC NBTI”. In: *Bias Temperature Instability for Devices and Circuits*. Ed. by Tibor Grasser. New York: Springer Science & Business Media, 2013, pp. 349–378.

[94] M. Toledano-Luque and B. Kaczer. “Characterization of Individual Traps in High-k Oxides”. In: *Bias Temperature Instability for Devices and Circuits*. Ed. by Tibor Grasser. New York: Springer Science & Business Media, 2013, pp. 597–614.

[95] Jacopo Franco and Ben Kaczer. “NBTI in (Si)Ge Channel Devices”. In: *Bias Temperature Instability for Devices and Circuits*. Ed. by Tibor Grasser. New York: Springer Science & Business Media, 2013, pp. 615–641.

[96] John Keane et al. “On-Chip Silicon Odometers for Circuit Aging Characterization”. In: *Bias Temperature Instability for Devices and Circuits*. Ed. by Tibor Grasser. New York: Springer Science & Business Media, 2013, pp. 679–717.

[97] Javier Martin-Martinez, Rosana Rodriguez, and Montse Nafria. “Simulation of BTI-Related Time-Dependent Variability in CMOS Circuits”. In: *Bias Temperature Instability for Devices and Circuits*. Ed. by Tibor Grasser. New York: Springer Science & Business Media, 2013, pp. 783–810.

[98] S. Mahapatra et al. “A Comparative Study of Different Physics-Based NBTI Models”. In: *Electron Devices, IEEE Transactions on* 60.3 (Mar. 2013), pp. 901–916.

[99] S. Chakravarthi et al. “A comprehensive framework for predictive modeling of negative bias temperature instability”. In: *Reliability Physics Symposium Proceedings, 2004. 42nd Annual. 2004 IEEE International*. Apr. 2004, pp. 273–282.

[100] A.T. Krishnan et al. “Material dependence of hydrogen diffusion: implications for NBTI degradation”. In: *Electron Devices Meeting, 2005. IEDM Technical Digest. IEEE International*. Dec. 2005, pages.

[101] John R Jameson et al. “A semiclassical model of dielectric relaxation in glasses”. In: *Journal of applied physics* 100.12 (2006), p. 124104.

[102] S. Desai et al. “A comprehensive AC / DC NBTI model: Stress, recovery, frequency, duty cycle and process dependence”. In: *Reliability Physics Symposium (IRPS), 2013 IEEE International*. Apr. 2013, XT.2.1–XT.2.11.

[103] K. Joshi et al. “A consistent physical framework for N and P BTI in HKMG MOSFETs”. In: *Reliability Physics Symposium (IRPS), 2012 IEEE International*. Apr. 2012, 5A.3.1–5A.3.10.

[104] A.E. Islam et al. “Gate Leakage vs. NBTI in Plasma Nitrided Oxides: Characterization, Physical Principles, and Optimization”. In: *Electron Devices Meeting, 2006. IEDM ’06. International.* Dec. 2006, pp. 1–4.

[105] A.E. Islam et al. “Recent Issues in Negative-Bias Temperature Instability: Initial Degradation, Field Dependence of Interface Trap Generation, Hole Trapping Effects, and Relaxation”. In: *Electron Devices, IEEE Transactions on* 54.9 (Sept. 2007), pp. 2143–2154.

[106] Franz Schanovsky and Tibor Grasser. “On the Microscopic Limit of the RD Model”. In: *Bias Temperature Instability for Devices and Circuits.* Ed. by Tibor Grasser. New York: Springer Science & Business Media, 2013, pp. 379–408.

[107] Dieter K. Schroder. “Negative bias temperature instability: What do we understand?” In: *Microelectronics Reliability* 47.6 (2007). Modelling the Negative Bias Temperature Instability, pp. 841–852.

[108] H. Kufluoglu and M.A. Alam. “Theory of interface-trap-induced NBTI degradation for reduced cross section MOSFETs”. In: *Electron Devices, IEEE Transactions on* 53.5 (May 2006), pp. 1120–1130.

[109] T. Grasser, W. Gos, and B. Kaczer. “Dispersive Transport and Negative Bias Temperature Instability: Boundary Conditions, Initial Conditions, and Transport Models”. In: *Device and Materials Reliability, IEEE Transactions on* 8.1 (Mar. 2008), pp. 79–97.

[110] Shigeo Ogawa and Noboru Shiono. “Generalized diffusion-reaction model for the low-field charge-buildup instability at the Si-SiO_2 interface”. In: *Phys. Rev. B* 51 (7 Feb. 1995), pp. 4218–4230.

[111] B. Kaczer et al. “Disorder-controlled-kinetics model for negative bias temperature instability and its experimental verification”. In: *Reliability Physics Symposium, 2005. Proceedings. 43rd Annual. 2005 IEEE International.* Apr. 2005, pp. 381–387.

[112] H. Reisinger et al. “Analysis of NBTI Degradation- and Recovery-Behavior Based on Ultra Fast VT-Measurements”. In: *Reliability Physics Symposium Proceedings, 2006. 44th Annual., IEEE International.* Mar. 2006, pp. 448–453.

[113] S. Mahapatra et al. “A critical re-evaluation of the usefulness of R-D framework in predicting NBTI stress and recovery”. In: *Reliability Physics Symposium (IRPS), 2011 IEEE International.* Apr. 2011, 6A.3.1–6A.3.10.

[114] F. Schanovsky and T. Grasser. “On the microscopic limit of the reaction-diffusion model for the negative bias temperature instability”. In: *Integrated Reliability Workshop Final Report (IRW), 2011 IEEE International.* Oct. 2011, pp. 17–21.

[115] F. Schanovsky and T. Grasser. “On the microscopic limit of the modified reaction-diffusion model for the negative bias temperature instability”. In: *Reliability Physics Symposium (IRPS), 2012 IEEE International.* Apr. 2012, XT.10.1–XT.10.6.

[116] Donald A. McQuarrie. “Stochastic Approach to Chemical Kinetics”. In: *Journal of Applied Probability* 4.3 (1967), pp. 413–478.

[117] Peter Hänggi, Peter Talkner, and Michal Borkovec. “Reaction-rate theory: fifty years after Kramers”. In: *Rev. Mod. Phys.* 62 (2 Apr. 1990), pp. 251–341.

[118] Ahmad Ehteshamul Islam et al. “Critical analysis of short-term negative bias temperature instability measurements: Explaining the effect of time-zero delay for on-the-fly measurements”. In: *Applied Physics Letters* 90.8, 083505 (2007), pp. 1–3.

[119] Ahmad Ehteshamul Islam and Muhammad Ashraful Alam. “Analyzing the distribution of threshold voltage degradation in nanoscale transistors by using reaction-diffusion and percolation theory”. In: *Journal of Computational Electronics* 10.4 (2011), pp. 341–351.

[120] Steven S Andrews and Dennis Bray. “Stochastic simulation of chemical reactions with spatial resolution and single molecule detail”. In: *Physical Biology* 1.3 (2004), p. 137.

[121] Radek Erban and S Jonathan Chapman. “Stochastic modelling of reaction-diffusion processes: algorithms for bimolecular reactions”. In: *Physical Biology* 6.4 (2009), p. 046001.

[122] Gianluca Malavasi et al. “Void size distribution in MD-modelled silica glass structures”. In: *Journal of Non-Crystalline Solids* 352.3 (2006), pp. 285–296.

[123] Wolfgang Goes. “Advanced Modeling of Oxide Defects”. In: *Bias Temperature Instability for Devices and Circuits.* Ed. by Tibor Grasser. New York: Springer Science & Business Media, 2013, pp. 409–446.

[124] Hans Reisinger. “The time-dependent defect spectroscopy”. In: *Bias Temperature Instability for Devices and Circuits.* Ed. by Tibor Grasser. New York: Springer Science & Business Media, 2013, pp. 75–109.

[125] V. Huard et al. “New characterization and modeling approach for NBTI degradation from transistor to product level”. In: *Electron Devices Meeting, 2007. IEDM 2007. IEEE International.* Dec. 2007, pp. 797–800.

[126] Ingemar Lundström and Christer Svensson. “Tunneling to traps in insulators”. In: *Journal of Applied Physics* 43.12 (1972), pp. 5045–5047.

[127] F.P. Heiman and G. Warfield. “The effects of oxide traps on the MOS capacitance”. In: *Electron Devices, IEEE Transactions on* 12.4 (Apr. 1965), pp. 167–178.

[128] S. Christensson, I. Lundström, and C. Svensson. “Low frequency noise in MOS transistors - I Theory”. In: *Solid-State Electronics* 11.9 (1968), pp. 797–812.

[129] W. Shockley and W. T. Read. “Statistics of the Recombinations of Holes and Electrons”. In: *Phys. Rev.* 87 (5 1952), pp. 835–842.

[130] Alan Louis McWhorter et al. “1/f noise and related surface effects in germanium.” In: (1955).

[131] 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.

[132] Muhammad Masuduzzaman, A.E. Islam, and M.A. Alam. “Exploring the Capability of Multifrequency Charge Pumping in Resolving Location and Energy Levels of Traps Within Dielectric”. In: *Electron Devices, IEEE Transactions on* 55.12 (Dec. 2008), pp. 3421–3431.

[133] A. Avellán, D. Schroeder, and W. Krautschneider. “Modeling random telegraph signals in the gate current of metal–oxide–semiconductor field effect transistors after oxide breakdown”. In: *Journal of Applied Physics* 94.1 (2003), pp. 703–708.

[134] Mark Isler and Dennis Liebig. “Enhanced multiphonon capture of hot electrons by deep centers with strong lattice coupling: A Monte Carlo study of InP:Fe”. In: *Phys. Rev. B* 61 (11 2000), pp. 7483–7488.

[135] N. Zanolla et al. “Measurement and simulation of gate voltage dependence of RTS emission and capture time constants in MOSFETs”. In: *Ultimate Integration of Silicon, 2008. ULIS 2008. 9th International Conference on.* Mar. 2008, pp. 137–140.

[136] Kun Huang and Avril Rhys. “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* 204.1078 (1950), pp. 406–423.

[137] Kurt V. Mikkelsen and Mark A. Ratner. “Electron tunneling in solid-state electron-transfer reactions”. In: *Chemical Reviews* 87.1 (1987), pp. 113–153.

[138] Melvin Lax. “The Franck-Condon Principle and Its Application to Crystals”. In: *The Journal of Chemical Physics* 20.11 (1952), pp. 1752–1760.

[139] Thomas H. Keil. “Shapes of Impurity Absorption Bands in Solids”. In: *Phys. Rev.* 140 (2A Oct. 1965), A601–A617.

[140] F. Schanovsky et al. “A multi scale modeling approach to non-radiative multi phonon transitions at oxide defects in MOS structures”. In: *Journal of Computational Electronics* 11.3 (2012), pp. 218–224.

[141] Supriyo Datta. *Quantum transport: atom to transistor*. Cambridge University Press, 2005.

[142] T. Grasser et al. “On the volatility of oxide defects: Activation, deactivation, and transformation”. In: *2015 IEEE International Reliability Physics Symposium*. Apr. 2015, 5A.3.1–5A.3.8.

[143] Markus Wilde and Katsuyuki Fukutani. “Hydrogen detection near surfaces and shallow interfaces with resonant nuclear reaction analysis”. In: *Surface Science Reports* 69.4 (2014), pp. 196–295.

[144] Ziyuan Liu et al. “Current Understanding of the Transport Behavior of Hydrogen Species in MOS Stacks and Their Relation to Reliability Degradation”. In: *ECS Transactions* 35.4 (2011), pp. 55–72.

[145] J. H. Stathis and D. J. DiMaria. “Identification of an interface defect generated by hot electrons in SiO2”. In: *Applied Physics Letters* 61.24 (1992), pp. 2887–2889.

[146] E. Cartier, J. H. Stathis, and D. A. Buchanan. “Passivation and depassivation of silicon dangling bonds at the Si/SiO2 interface by atomic hydrogen”. In: *Applied Physics Letters* 63.11 (1993), pp. 1510–1512.

[147] E. Cartier, D. A. Buchanan, and G. J. Dunn. “Atomic hydrogen-induced interface degradation of reoxidized-nitrided silicon dioxide on silicon”. In: *Applied Physics Letters* 64.7 (1994), pp. 901–903.

[148] E. Cartier and J.H. Stathis. “Insulating Films on Semiconductors Atomic hydrogen-induced degradation of the Si/SiO2 structure”. In: *Microelectronic Engineering* 28.1 (1995), pp. 3–10.

[149] J. M. M. de Nijs et al. “Hydrogen induced donor-type Si/SiO2 interface states”. In: *Applied Physics Letters* 65.19 (1994), pp. 2428–2430.

[150] M. Houssa et al. “Insights on the physical mechanism behind negative bias temperature instabilities”. In: *Applied Physics Letters* 90.4 (Jan. 2007), p. 043505.

[151] T. Grasser et al. “On the microscopic structure of hole traps in pMOSFETs”. In: *2014 IEEE International Electron Devices Meeting*. Dec. 2014, pp. 21.1.1–21.1.4.

[152] Al-Moatasem El-Sayed et al. “Hydrogen-Induced Rupture of Strained Si-O Bonds in Amorphous Silicon Dioxide”. In: *Phys. Rev. Lett.* 114 (11 Mar. 2015), p. 115503.

[153] T. Grasser et al. “The time dependent defect spectroscopy (TDDS) for the characterization of the bias temperature instability”. In: *Reliability Physics Symposium (IRPS), 2010 IEEE International*. May 2010, pp. 16–25.

[154] J. W. McPherson. “Quantum Mechanical Treatment of Si-O Bond Breakage in Silica Under Time Dependent Dielectric Breakdown Testing”. In: *2007 IEEE International Reliability Physics Symposium Proceedings. 45th Annual*. Apr. 2007, pp. 209–216.

[155] T. Grasser 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)*. Dec. 2015, pp. 20.1.1–20.1.4.

[156] T. Grasser et al. “NBTI in Nanoscale MOSFETs - The Ultimate Modeling Benchmark”. In: *IEEE Transactions on Electron Devices* 61.11 (Nov. 2014), pp. 3586–3593.

[157] Hans Reisinger et al. “The statistical analysis of individual defects constituting NBTI and its implications for modeling DC- and AC-stress”. In: *Reliability Physics Symposium (IRPS), 2010 IEEE International*. July 2010, 2A.1.1–2A.1.9.

[158] M. Cho et al. “Off-state stress degradation mechanism on advanced p-MOSFETs”. In: *2015 International Conference on IC Design Technology (ICICDT)*. June 2015, pp. 1–4.

[159] Ralf Rudolf et al. “Automotive 130 nm Smart-Power-Technology including embedded Flash Functionality”. In: *Proceedings of the 23rd International Symposium on Power Semiconductor Devices & IC's*. May 2011, pp. 20–23.

[160] T. Schiml et al. “A 0.13 /spl mu/m CMOS platform with Cu/low-k interconnects for system on chip applications”. In: *VLSI Technology, 2001. Digest of Technical Papers. 2001 Symposium on*. June 2001, pp. 101–102.

[161] Terence B. Hook, David Harmon, and Chuan Lin. “Plasma process-induced damage on thick (6.8 nm) and thin (3.5 nm) gate oxide: parametric shifts, hot-carrier response, and dielectric integrity degradation”. In: *Microelectronics Reliability* 41.5 (2001), pp. 751–765.

[162] A. Martin et al. “Reliability degradation of MOS transistors originated from plasma process-induced charging of circuit blocks and detected with fWLR methods”. In: *Integrated Reliability Workshop Final Report (IRW), 2011 IEEE International*. Oct. 2011, pp. 82–86.

[163] Hans Reisinger et al. “The Impact of Recovery on BTI Reliability Assessments”. In: *Integrated Reliability Workshop Final Report (IRW), 2010 IEEE International.* 2010, pp. 12–16.

[164] T. Grasser et al. “Recent advances in understanding the bias temperature instability”. In: *Electron Devices Meeting (IEDM), 2010 IEEE International.* Dec. 2010, pp. 4.4.1–4.4.4.

[165] M. Simicic et al. “Advanced MOSFET variability and reliability characterization array”. In: *2015 IEEE International Integrated Reliability Workshop (IIRW).* Oct. 2015, pp. 73–76.

[166] T. Grasser et al. “Analytic modeling of the bias temperature instability using capture/emission time maps”. In: *Electron Devices Meeting (IEDM), 2011 IEEE International.* Dec. 2011, pp. 27.4.1–27.4.4.

[167] Reinhold Paul. *MOS-Feldeffekttransistoren.* Springer-Verlag Berlin Heidelberg, 1994, p. 155.

[168] K. U. Giering et al. “NBTI modeling in analog circuits and its application to long-term aging simulations”. In: *Integrated Reliability Workshop Final Report (IIRW), 2014 IEEE International.* Oct. 2014, pp. 29–34.

[169] Alexander Acovic, Giuseppe La Rosa, and Yuan-Chen Sun. “A review of hot-carrier degradation mechanisms in MOSFETs”. In: *Microelectronics Reliability* 36.7-8 (1996), pp. 845–869.

[170] E Takeda. “Hot-carrier effects in submicrometre MOS VLSIs”. In: *IEE Proceedings I (Solid-State and Electron Devices)* 131.5 (1984), pp. 153–162.

[171] Jeff D Bude. “Gate current by impact ionization feedback in sub-micron MOSFET technologies”. In: *VLSI Technology, 1995. Digest of Technical Papers. 1995 Symposium on.* IEEE. 1995, pp. 101–102.

[172] Franco Venturi, Enrico Sangiorgi, and Bruno Riccò. “The impact of voltage scaling on electron heating and device performance of submicrometer MOSFETs”. In: *IEEE Transactions on Electron Devices* 38.8 (1991), pp. 1895–1904.

[173] William McMahon and Karl Hess. “A multi-carrier model for interface trap generation”. In: *Journal of Computational Electronics* 1.3 (2002), pp. 395–398.

[174] James E Chung et al. “Low-voltage hot-electron currents and degradation in deep-submicrometer MOSFETs”. In: *IEEE Transactions on electron devices* 37.7 (1990), pp. 1651–1657.

[175] SE Rauch, G La Rosa, and FJ Guarin. “Role of EE scattering in the enhancement of channel hot carrier degradation of deep-submicron NMOSFETs at high V GS conditions”. In: *IEEE Transactions on Device and Materials Reliability* 1.2 (2001), pp. 113–119.

[176] Chloé Guérin, Vincent Huard, and Alain Bravaix. “The energy-driven hot-carrier degradation modes of nMOSFETs”. In: *IEEE Transactions on Device and Materials Reliability* 7.2 (2007), pp. 225–235.

[177] Stewart E Rauch, Fernando J Guarin, and Giuseppe LaRosa. “Impact of EE scattering to the hot carrier degradation of deep submicron NMOSFETs”. In: *IEEE Electron Device Letters* 19.12 (1998), pp. 463–465.

[178] T. Grasser et al. “On the microscopic origin of the frequency dependence of hole trapping in pMOSFETs”. In: *2012 International Electron Devices Meeting*. Dec. 2012, pp. 19.6.1–19.6.4.

[179] T. Aichinger, M. Nelhiebel, and T. Grasser. “Unambiguous identification of the NBTI recovery mechanism using ultra-fast temperature changes”. In: *2009 IEEE International Reliability Physics Symposium*. Apr. 2009, pp. 2–7.

[180] G. Pobegen et al. “Understanding temperature acceleration for NBTI”. In: *2011 International Electron Devices Meeting*. Dec. 2011, pp. 27.3.1–27.3.4.

[181] G. Pobegen and T. Grasser. “On the Distribution of NBTI Time Constants on a Long, Temperature-Accelerated Time Scale”. In: *IEEE Transactions on Electron Devices* 60.7 (July 2013), pp. 2148–2155.

[182] H. Reisinger et al. “Understanding and modeling AC BTI”. In: *2011 International Reliability Physics Symposium*. Apr. 2011, 6A.1.1–6A.1.8.

[183] R. Bellens et al. “On the hot-carrier-induced post-stress interface trap generation in n-channel MOS transistors”. In: *IEEE Transactions on Electron Devices* 41.3 (Mar. 1994), pp. 413–419.

[184] Gunnar Andreas Rott et al. “Mixture of negative bias temperature instability and hot-carrier driven threshold voltage degradation of 130nm technology p-channel transistors”. In: *Microelectronics Reliability* 54.9 (2014), pp. 2310–2314.

[185] B. Ullmann et al. “The impact of mixed negative bias temperature instability and hot carrier stress on single oxide defects”. In: *2017 IEEE International Reliability Physics Symposium (IRPS)*. Apr. 2017, pages.

[186] G. A. Rott et al. “Drift compensating effect during hot-carrier degradation of 130nm technology dual gate oxide P-channel transistors”. In: *Integrated Reliability Workshop Final Report (IRW), 2013 IEEE International*. 2013, pp. 73–77.

[187] K. Rott et al. “Impact and measurement of short term threshold instabilities in MOSFETs of analog circuits”. In: *2012 IEEE International Integrated Reliability Workshop Final Report*. Oct. 2012, pp. 31–34.

[188] Tibor Grasser et al. “A two-stage model for negative bias temperature instability”. In: *2009 IEEE International Reliability Physics Symposium*. 2009, pp. 33–44.

[189] Tibor Grasser et al. “The ‘permanent’component of NBTI: composition and annealing”. In: *Reliability Physics Symposium (IRPS), 2011 IEEE International*. 2011, 6A–2.

[190] Prasad Chaparala and Douglas Brisbin. “Impact of NBTI and HCI on PMOSFET threshold voltage drift”. In: *Microelectronics Reliability* 45.1 (2005), pp. 13–18.

[191] Yandong He and Ganggang Zhang. “Experimental insights on the degradation and recovery of pMOSFET under non-uniform NBTI stresses”. In: *Physical and Failure Analysis of Integrated Circuits (IPFA), 2011 18th IEEE International Symposium on the*. IEEE. 2011, pp. 1–6.

[192] G. Pobegen et al. “Observation of Normally Distributed Energies for Interface Trap Recovery After Hot-Carrier Degradation.” In: *IEEE Electron Device Letters 2013* 34 (2013), pp. 939–941.

[193] F Cacho et al. “HCI/BTI coupled model: The path for accurate and predictive reliability simulations”. In: *Reliability Physics Symposium, 2014 IEEE International*. IEEE. 2014, pp. 5D–4.

[194] David L. Griscom. “Diffusion of radiolytic molecular hydrogen as a mechanism for the post-irradiation buildup of interface states in SiO2-on-Si structures”. In: *Journal of Applied Physics* 58.7 (1985), pp. 2524–2533.

[195] *Virtuoso(R) Unified Reliability Interface Reference*. English. Version 12.1. Cadence Design Systems. Oct. 2012. 90 pp.

[196] K.-U. Giering et al. “Analog-circuit NBTI degradation and time-dependent NBTI variability: An efficient physics-based compact model”. In: *Reliability Physics Symposium (IRPS), 2016 IEEE International*. 2016.

List of Figures

List of Tables

Curriculum Vitae

Personal information

- 13.11.1985: Born in Flensburg, Germany

Education

- 1998-2004: High School: Mariengymnasium, Jever, Germany
- 2004: Received “Abitur / Allgemeine Hochschulreife” from Mariengymnasium, Jever, Germany
- 2004-2006: Studies of Physics: Universität Oldenburg, Germany (Grundstudium)
- 2004-2009: Received scholarship from the “Studienstiftung des Deutschen Volkes”
- 2006-2009: Studies of Physics: Universität Rostock, Germany (Hauptstudium)
- 2009: Received degree of “Diplom-Physiker” from the Universität Rostock, Germany
- 2009-2011: PhD student (received scholarship from the “Deutsche Forschungsgemeinschaft”, research training group “welisa”), Universität Rostock
- since 2012: PhD student, Institute for Microelectronics, TU Wien

Work Experience

- 2001: Formation of the web design company “seiten-bauer.com”
- 2006: Internship at Deutsche Lufthansa AG, Frankfurt, Germany
- 2012-2014: PhD student at Infineon Technologies AG, Neubiberg, Germany
- since 2014: Reliability engineer at Infineon Technologies AG, Neubiberg, Germany
- since 2017: Project manager at Infineon Technologies AG, Neubiberg, Germany

Declaration

Hereby I declare that I wrote this dissertation myself only with the help of the cited literature and auxiliary means. Up to now, this thesis was not published or presented to another examinations office in the same or similar shape.

Gunnar Rott
München, April 10$^{\text{th}}$, 2018