CARL VON OSSIETZKY universität OLDENBURG

Fakultät II
Department für Informatik
Abteilung Eingebettete Hardware-/Software-Systeme

Abstraction of Aging Models for High Level Degradation Prediction

Dissertation
zur Erlangung des
Doktorgrades der Naturwissenschaften
(Doctor rerum naturalium)

vorgelegt von
Reef Janes Eilers
geboren am 22.07.1980 in Oldenburg

Oldenburg 2017

To my wife and our daughters,
for your support and love

Submitted: February 28, 2017

Advisor: Prof. Dr.-Ing. Wolfgang Nebel
Second Reader: Prof. Dr.-Ing. Steffen Paul

Defended: May 31, 2017

Abstract

The aging effect “Negative Bias Temperature Instability”, which is highly dependent on device history, has a direct impact on the design of integrated circuits. In order to make realistic predictions available in the design process, an analysis procedure in conjunction with major industrial tools is required and simulation durations of existing history aware models must be significantly reduced. Therefore, a gate level methodology relying on a performance-oriented, yet accurate abstraction of the switching trap NBTI model is presented within this thesis. Evaluation results for various stress scenarios demonstrate very precise simulations and a major improvement to another performance-oriented model abstraction. In this way, simulation durations facilitate realistic aging predictions of larger components in a reasonable period of time. Various aging assessments finally demonstrate that a circuit's realistic aging is strongly overestimated by static worst case assumptions.

Keywords: Negative Bias Temperature Instability, Reliability, Aging, Mission Scenario, Gate Level

Kurzzusammenfassung

Der Alterungseffekt “Negative Bias Temperature Instability”, der stark von der Verwendungsweise eines Geräts abhängt, muss beim Entwurf von integrierten Schaltungen berücksichtigt werden. Hierzu wird ein Verfahren für die Alterungsvorhersage benötigt, das geringe Simulationsdauern aufweist und auf industriellen Standardwerkzeugen beruht. Diese Arbeit stellt daher ein solches Verfahren auf Gatterebene vor, das auf einer leistungsorientierten und dennoch genauen Abstraktion des physikalischen Alterungsmodells beruht. Die Simulationsergebnisse verschiedener Einsatzszenarien zeigen sehr hohe Genauigkeiten und eine wesentliche Verbesserung gegenüber einer alternativen Abstraktionsmethode, so dass realistische Alterungsvorhersagen größerer Komponenten innerhalb sinnvoller Zeiträume ermöglicht werden. Verschiedene Alterungsbewertungen zeigen schließlich, dass realistische Alterungen eines Schaltkreises durch statische Worst-Case-Annahmen stark überschätzt werden.

Schlagwörter: Negative Bias Temperature Instability, Zuverlässigkeit, Alterung, Einsatzszenario, Gatterebene

Contents

CHAPTER 1

Introduction

1.1 Motivation

With the pace of Moore's law, industry is driving technology dimensions further towards the atomic regime. With this scaling, the technology picked up more and more physical artifacts, influencing the usage of such devices. The advent of various flavors of static currents such as sub-threshold leakage in 90nm, gate tunneling in 65nm and gate induced drain leakage in 45nm [1], also introduced an increasing susceptibility to process variations as well as an electro-thermal coupling. This resulted in quite a number of challenging problems for the design of integrated circuits in the last decade. Currently, it seems, as if degradation effects, which cause transformations of an integrated circuit's characteristics after delivery, could become one of the main challenges for this decade. Similar to the static currents, there is not just a new physical phenomenon - there is rather a vast selection of degradation mechanisms.

These degradation mechanisms can be separated into two different classes. At first, there are degradation effects that directly cause permanent local failures within digital circuits or even induce a complete destruction of a device. The microcracking effect, which characterizes the loss of electrical connections due to diverging thermal expansions of different materials, is thus a fundamental example for this class of degradation effects. In addition, electro-migration (EM) specifies a force, introduced by high current densities, that can dislocate interconnect material at elevated temperatures [2]. After a phase of mild degradation of the interconnect's resistance, the electro-migration effect will enter a run-away situation, resulting in a sudden total connection-loss. Time dependent dielectricity breakdown (TDDB) may occur in oxides, having collected a vast number of trapped oxide charges, forming a conductive path through the oxide and thus to a permanent device failure [3]. Finally, there are radiation induced permanent failures like single event latchup and others, each of which finally leads to a thermal destruction of a device as a result of ionizing radiation.

Mechanisms slowly varying relevant process parameters over time form the second class of degradation effects. At high temperatures, the most relevant degradation effect of this class is negative bias temperature instability (NBTI), where chemical traps in the gate oxide can capture and emit charges, thus causing an increase of the transistor's threshold voltage [4]. At lower temperatures hot carrier degradation (HCD) dominates, where fast (hot) carriers can get trapped in the oxide thus again influencing the threshold voltage of the transistor [5]. From an abstract view, all degradation effects of this class result in a change of power demand and path timing. As soon as the available slack within one path is exceeded, degradation will also lead to a timing failure. A distinct classification of all degradation effects within this schema may sometimes not be possible, as TDDB causes a slight timing degradation prior to the creation of a conductive path for instance [6].

Research and industry have a long tradition of the handling of failures due to process variations in the way that redundancies may be employed [7] and testing procedures are realized prior to delivery. Hence, a particular error margin of the produced devices' failure rate is economically still acceptable. However, this failure rate increases for recent technologies after delivery due to degradation effects as shown in Figure 1.1. With technology scaling, the failure rates even tend to surpass the error margin after increasingly shorter lifetimes of the devices, causing violations of the economically acceptable error margin within operation time of avionics, automotive or medical applications. Possible degradation induced failures of numerous devices within guaranteed operation times may generate tremendous cost. For instance, Intel Corporation admits a $700 million hit for Sandy Bridge chipset problems, which are probably caused by time dependent dielectricity breakdown of a single transistor and only affect approximately 5% of the systems over three years [8]. Hence, research and industry are currently trying to develop tools and methodologies, helping to cope with degradation at all design levels from system design, where redundancy may be employed and parameter adaptions of a performance-reliability trade-off can be realized, down to devices, where direct reduction of the effects are the main focus.

In order to assess these tools and methodologies as well as a device's overall susceptibility to degradation, an estimation of the expected degradation behavior is needed prior to delivery of a device. Degradation measurements based on prototypes are thus a common approach, which is though associated with considerable cost and may cause delays within product development cycles. Furthermore, degradations due to realistic stress scenarios can surely not be measured

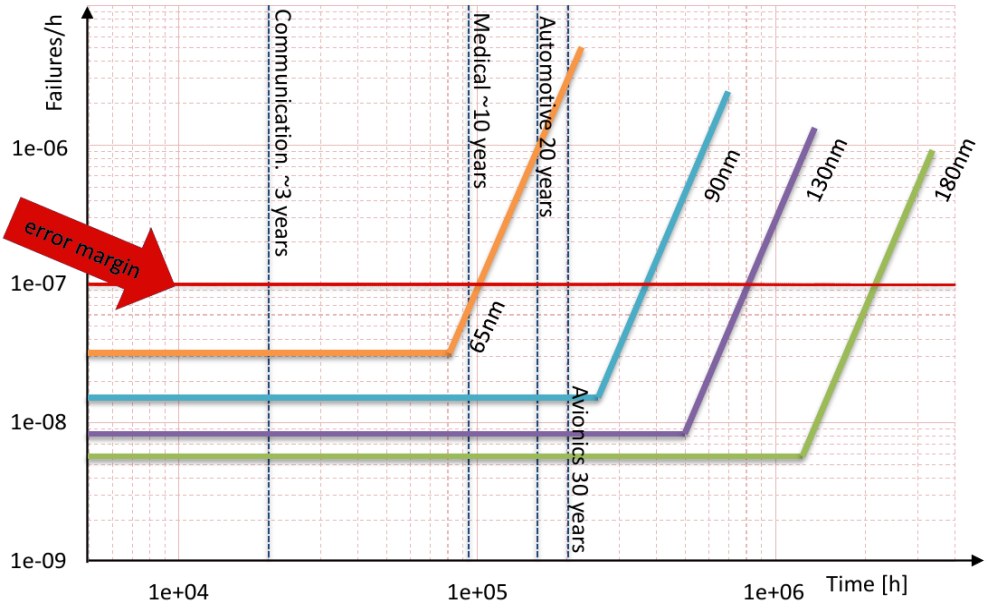


Figure 1.1: Failure rates of different manufacturing technologies as function of produced devices' lifetime (Source: [9])

due to slow acting transformations in the course of several years, leaving measurements of short or mid term scenarios in combination with accelerated degradation conditions on account of very high temperatures or supply voltages and extrapolations for long term stress scenarios. Thus, considerable inaccuracies may be introduced by these conversions from mid term accelerated degradation conditions to realistic long term scenarios.

Instead of costly degradation measurements, assessment of the expected degradation behavior may be analyzed based on simulations. This approach is already successfully applied to estimate the failure rates of integrated circuits due to electro-migration and time dependent dielectricity breakdown. In detail, mean time to failure of the complete integrated circuit is modeled on the basis of lognormal or Weibull device failure distributions for these degradation effects and a grid based approach is used to consider variations of influencing variables (e.g. temperature) across the circuit [10, 11]. A similar grid based approach is also proposed to estimate the elementary impact of on-chip temperature variations on the timing degradation effect NBTI [12]. However, occurrence of NBTI induced failures can only be specified within this approach by exceedance of a pre-defined threshold voltage boundary. As a major threshold voltage shift of a single transistor does not necessarily cause a timing failure, the approach's simulated NBTI induced failure rates may thus be very imprecise.

In contrast to this inexact NBTI approach, precise simulations of integrated circuit's failure rates due to the timing degradation effects NBTI and HCD have to rely on the timing degradations of various signal paths between registers as timing failures are provoked by the belated access of a register. Hence, a two stage modeling approach, consisting of threshold voltage estimations and respective implications on the timing of register transfer (RT) level components, has to be applied to assess the noncompliance of each RT component's timing with predetermined specifications and to thus simulate a circuit's overall failure rate. In particular, first stage of this modeling approach has to also respect individual stress scenarios of each transistor due to differing signal probabilities that lead to considerably diverse threshold voltage shifts [13]. Thus, NBTI and HCD simulations have to be performed for each transistor within the approach's first stage and these individual threshold voltage shifts have to be incorporated within the timing estimations of the second stage, which can be performed using the common analog electronic circuit simulation program SPICE [14].

Main additional requirement of the described NBTI failure rate analysis is the approach's overall performance, as the analysis should finally be integrated in the product development cycle and assessments of tools and methodologies to cope with degradation should not introduce

considerable development delays. In this way, simulation durations in the order of several weeks or month are not acceptable for a single degradation assessment. Already proposed methodologies that implement the described NBTI failure rate analysis [15, 16, 17] thus utilize various simplifications to improve the method's performance. At first, direct analog electronic circuit simulations are avoided within the approaches' second stage by using common gate level timing estimation procedures that rely on pre-characterized lookup tables. Additional simplifications that greatly reduce the number of signal path that have to be estimated [15] unfortunately receive little industrial support as extensively verified industrial timing estimation procedures would have to be replaced completely. Since threshold voltage degradations have to be estimated individually for every transistor, major performance improvements arise due to the replacement of accurate physical degradation models by simple analytical models. However, these analytical models rely on constant stress scenarios and thus cause a severe restriction of possible stress scenarios in the way that varying scenarios as induced by system idle phases or techniques as power gating cannot be supported and the analysis is effectively limited to worst case assumptions. Since NBTI degradation is highly dependent on temperature and regenerates during idle phases, these worst case assumptions may cause a severe overestimation of the circuit's failure rate. In consequence, production cost may be increased due to employed redundancies associated with a larger chip size or the integrated circuit's performance is reduced due to additional slack.

Best to my knowledge, all existing NBTI models are either too slow for a full circuit and/or full lifetime simulation, cannot handle varying stress conditions as typically occurring in most systems, or they are not accurately following silicon measurements. Hence, a new technique to simulate degradation induced threshold voltage shifts fulfilling the requirements of sufficient performance for full circuit analysis and support of varying stress scenarios with idle phases and changing influencing variables (e.g. temperature) is needed. Besides, calculation of the degraded timing of signal paths between registers should be based on already proposed efficient methodologies by using verified industrial timing estimation procedures whenever possible in order to increase the industrial relevance of the degradation assessment approach as degraded and non degraded timings would be calculated similarly. In this way, some of the worst case assumptions in integrated circuit design may be disregarded thus causing slightly reduced production cost or facilitating improvements of a device's performance.

Which methodologies may be applied to facilitate a performance oriented yet varying mission scenario aware abstraction of physical timing degradation models is thus the main research question of this thesis. Furthermore, the question in which way industrial timing estimation procedures can be utilized within the estimation of degraded timings of signal paths between registers is investigated. In this way, the issue of what is the impact of static worst case assumptions concerning idle phases, temperatures and supply voltages on simulated timing degradations of RT level components and thus failure rates of integrated circuits, is examined.

1.2 Preliminary Inspection

In order to preliminary assess the implication of the usage of varying stress scenarios during timing degradation analysis and thus to tentatively evaluate the importance of the proposed methodologies of this thesis, a small-sized initial inspection is performed. Therefore, NBTI threshold voltage degradations are simulated for different long term static stress conditions based on an implementation of a published NBTI model (see Section 5.1.2) and characterization data of a 22 nm FinFET technology (see Section 5.1.1). In particular, static temperatures of 300, 350 and 400 K, gate voltages of -0.8 and -1 V and a maximal simulation time of 10 years are utilized. As presented in Figure 1.2, shift in threshold voltage is up to two times larger due to the increase in gate voltage by 0.2 V and up to 8 times larger due to the increase in temperature by 100 K. Besides, enormous threshold voltage shifts larger than 400 mV occur for the worst case condition of 400 K and -1 V.

These results confirm that a degradation simulation based on worst case assumptions will greatly overestimate the degradation effect when a system is subjected to varying stress with changing temperatures and gate voltages. It is thus necessary to simulate the degradation based on a method that supports varying temperatures and gate voltages. In addition, the

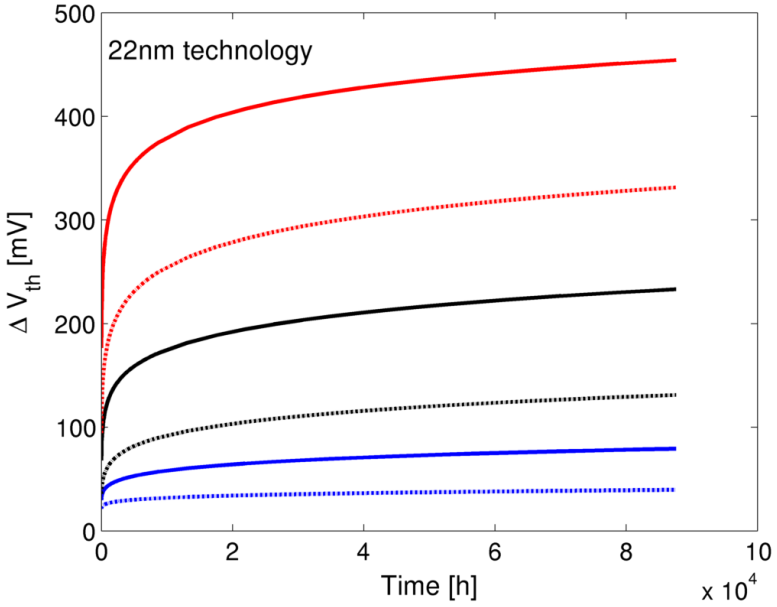


Figure 1.2: NBTI threshold voltage degradation for a 22 nm technology in the course of 10 years caused by permanent stress with gate voltage -0.8 V (dashed) and -1 V (solid) at 300 (blue), 350 (black) and 400 K (red)

enormous maximal threshold voltage degradation generally emphasizes the significance of the NBTI degradation effect and highlights the need of a degradation analysis within the design of integrated circuits.

1.3 Objective and Application of the Thesis

Overall objective of this thesis is thus the developed of a methodology that facilitates an accurate, efficient and reliable computation of the degraded timings of RT level components subjected to varying stress. A separation of this main objective into sub-goals for three different parts of the methodology further particularizes the demands.

At first, the methodology shall be integrated as an additional component within an industrial tool flow. In this way, timings of complete signal path shall be calculated using industrial timing estimation procedures, since these procedures are efficient and generate reliable results due to extensive testing. As degraded and non degraded timings are thus calculated similarly, the industrial relevance of the methodology may be increased. Furthermore, determining factors of the degradation shall also be estimated by industrial tools whenever possible. Secondly, main requirements of the methodology's technique to estimate degraded threshold voltages are performance demands to facilitate numerous computations within reasonable time and support of realistic stress scenarios as induced by varying temperatures, changing supply voltages due to dynamic voltage and frequency scaling (DVFS) and idle phases caused by overall off states or power gating. As the technique's accuracy is a key objective and a systematical underestimation of the threshold voltage degradation must not occur, the technique's evaluation with precise reference models is an essential part of this thesis. Finally, overall performance requirements call for efficient conversions between threshold voltage shifts and timing degradations of single devices. Likewise, accuracy of these conversions has to be evaluated. Hence, efficient techniques are used for each part of the methodology and the overall accuracy is guaranteed by the parts' evaluations.

In this way, the main application of this thesis is the computation of degraded timings of RT level components subjected to varying stress. Based on these results, the overall failure rate of an integrated circuit may afterwards be easily estimated by the noncompliance of each RT component's degraded timing with predetermined specifications.

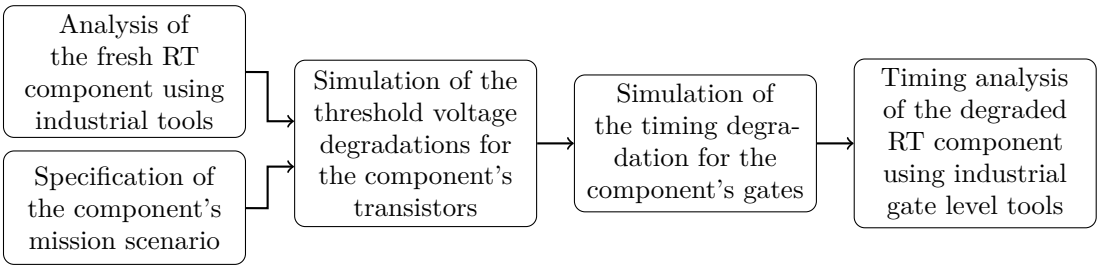


Figure 1.3: Basic concept of the proposed degradation analysis flow

1.4 Basic Concept of the Proposed Timing Degradation Analysis

According to the main objective of the development of a methodology to simulate degraded timings of RT level components, a basic concept of an appropriate tool flow is presented in Figure 1.3. As starting point, the basic concept uses industrial tools for an initial analysis of the unaged "fresh" RT component to estimate relevant characteristics for the simulation of the timing degradation. In addition, the RT component's mission scenario is precisely specified as the component's degradation clearly depends on the utilized scenario (see Section 1.2). Afterwards, degraded threshold voltages are estimated for the relevant transistors within the RT component and a subsequent simulation of the resulting timing degradation is carried out for each device. In particular, timing degradations are estimated for the RT component's gates in order to finally utilize industrial gate level tools for the timing analysis of the complete RT component. In this way, the proposed methodology is integrated as an additional component within an industrial tool flow.

In order to prepare the implementation of the basic tool flow concept, relevant fundamentals concerning design and analysis of integrated circuits are stated within the following chapters to identify possible interfaces for the integration of the simulation of degraded timings as an additional component within an industrial tool flow. Furthermore, the different published modeling techniques of the timing degradation effects must be known as a performance oriented yet varying stress scenario aware simulation of a transistor's threshold voltage degradation is required within the proposed modeling technique.

1.5 Structure of the Thesis

Before the proposed timing degradation analysis procedure, which is integrated within an industrial tool flow, is presented in detail within Chapter 4, relevant fundamentals are described in Chapter 2 and Chapter 3 summarizes the recent state-of-the-art in timing degradation modeling. Afterwards, Chapter 5 proposes a novel, efficient and varying stress scenario aware method to calculate a transistor's degraded threshold voltage and presents an extensive evaluation with state-of-the-art reference models. Timing degradation of single gates is covered in Chapter 6 by generation and evaluation of abstract timing specifications based on SPICE measurements. In addition, summaries within Chapter 5 and 6 highlight the integration of the developed methodologies within the proposed timing degradation analysis procedure. As an example of use and to verify the applicability of the proposed procedure, the timing degradation analysis is executed in Chapter 7. Finally, Chapter 8 concludes the thesis and presents an outlook on future work.

CHAPTER 2

Fundamentals

In order to develop a simulation methodology for the timing degradation of register transfer (RT) level components that relies on the proposed basic concept of Section 1.4, the relevant fundamental preconditions have to be regarded. At first, the physical origins of major degradation effects that cause timing degradation are presented and the respective implications on timing characteristics are stated. Besides, the typical process of the design of integrated circuits is examined with a focus on simulation methodologies to estimate the circuit's timing. Furthermore, typical calculation methods of parameters that are relevant for the simulation of degraded timings are regarded.

2.1 Transistor Degradation Effects

As stated in Chapter 1, there are several completely different transistor degradation effects. However, main degradation effects that have to be considered during a gate level analysis of timing degradation are negative bias temperature instability and hot carrier degradation, which are described in Sections 2.1.1 and 2.1.2.

2.1.1 Negative Bias Temperature Instability (NBTI)

The term negative bias temperature instability is a phenomenological description of a degradation effect. If a PMOS transistor is subjected to a negative gate-source voltage (negative bias) at an elevated temperature, the transistor's properties may change (instability). NBTI is already characterized several years ago [18], but the actual physical origin of the effect is still controversial. This topic is further discussed Section 3.1.

In detail, stress phases with negative gate-source voltage, hereafter referred to as gate voltage, cause an increase of the transistor's threshold voltage V_{th}. As shown in the left panel of Figure 2.1, the degradation effect increases with increasing temperature. Furthermore, relaxation phases without negative gate voltage cause a partial regeneration of the threshold voltage degradation. Again, the actual relaxation curve depends on temperature (see right panel of Figure 2.1). While conventional NBTI studies have analyzed the shape of the V_{th} envelope induced by different long term stress scenarios, more recent studies examine long term relaxation scenarios [19] or short time effects [20]. Here, it is proven that a significant share of degradation and relaxation process occurs on short time scales.

When using very small devices (see Figure 2.2), NBTI degradation and relaxation curves consist of several individual ΔV_{th} steps [21]. Hence, NBTI must be caused by several individual events and is determined by statistical processes. It is also interesting to note that NBTI was thought to be independent of a stress signal's frequency [20], which could be rebutted by [21].

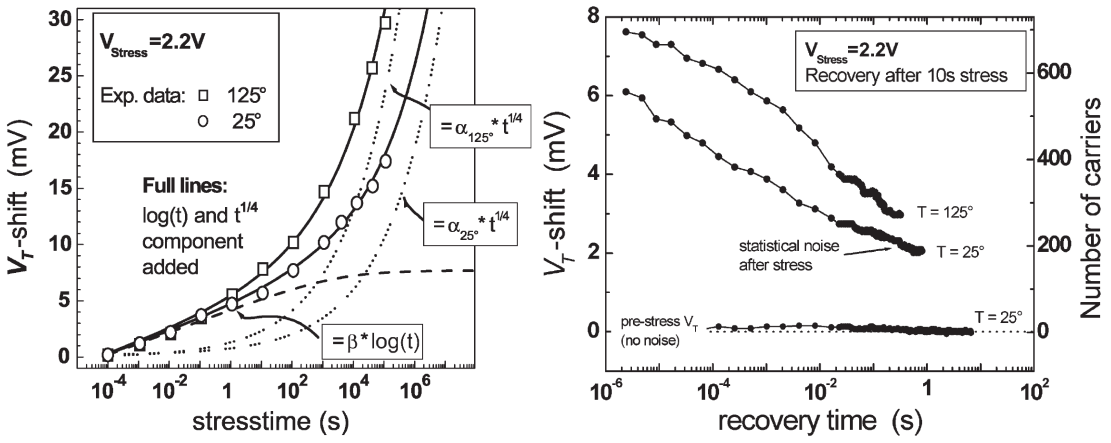


Figure 2.1: Measurement data of NBTI degradation (left panel) and relaxation (right panel) curves (Source: [19])

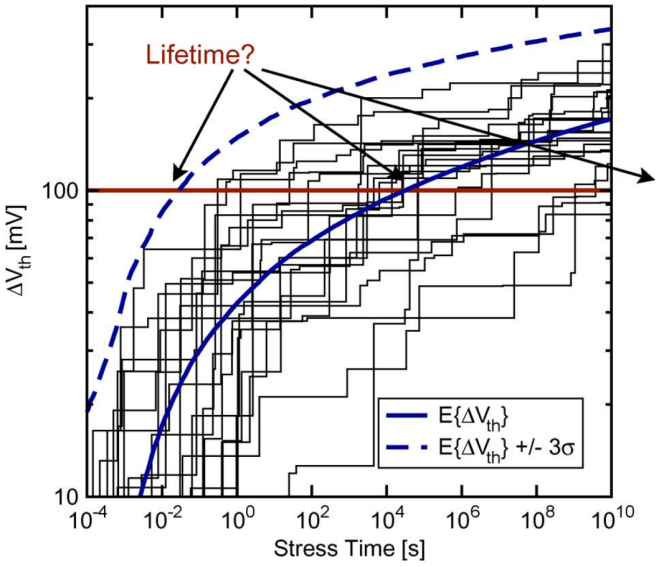


Figure 2.2: Variations of NBTI degradation curves due to individual events and statistical processes (Source: [22])

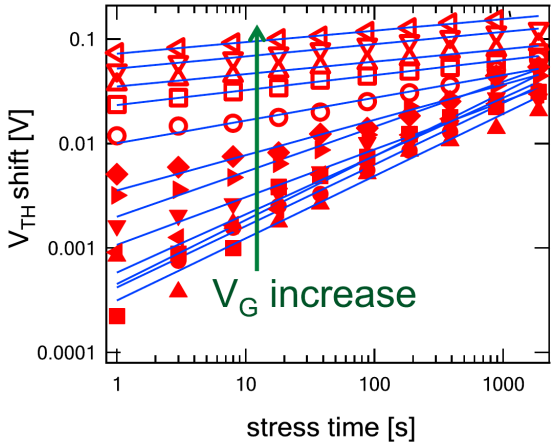


Figure 2.3: Threshold voltage degradation caused by source drain currents (Source: Adapted from [23])

2.1.2 Hot-Carrier Degradation (HCD)

Hot-carrier degradation describes the effect that the source drain current causes a significant increase in gate leakage [1] and leads to ΔV_{th} degradation of both PMOS and NMOS transistors [23]. This characteristic is presented in Figure 2.3. Recovery of HCD degradation is in general rather weak, if there is any recovery at all [5]. In contrast to NBTI, hot-carrier degradation also hardly depends on temperature.

2.1.3 Impact on Combinational Gates

The increase of a transistor's threshold voltage, which is caused by negative bias temperature instability or hot-carrier degradation, clearly affects the transistor's timing characteristics. Consequently, timing characteristics of combinational gates may also be altered by degradation effects. First of all, the gate's parameters output delay and output slew rate, which are defined in Figure 6.1, are directly increased due to the threshold voltage degradation [24]. Furthermore, a second order timing degradation occurs. The increased output slew rate causes an additional growth of a subsequent gate's delay, as the input slew rate of this gate is reduced. Since the

NBTI effect only increases the threshold voltage of PMOS transistors, the high to low switching delay of an inverter is though not affected by NBTI [24].

2.2 Design of Integrated Circuits

Main objective of the design of integrated circuits is to generate a circuitry of transistors that directly performs particular computations. Since a large number of transistors and a complex circuitry is needed in general, positioning and connection of every transistor is error-prone and must thus be executed automatically based on an abstract specification of the functional behavior. Additional degrees of freedom exist for the total number as well as positioning and connection of transistors, as different integrated circuits may feature the same functional behavior. These degrees of freedom can be used to optimize non functional properties like performance characteristics, power consumption, degradation behavior or physical size of the integrated circuit, which roughly corresponds to manufacturing cost. In order to modify the integrated circuit's non functional properties and to systematically generate the circuitry of transistors based on a functional specification, abstract descriptions of the circuitry are used.

Main abstractions that are relevant within this thesis are gate level and register transfer (RT) level descriptions. At gate level the circuitry of transistors is grouped by logic gates, which are defined as small circuits that perform logical operations on one or more inputs and produce a single logical output (e.g. NOT or NAND). Due to the basic CMOS technology each gate consists of complementary and symmetrical pairs of PMOS and NMOS transistors. In contrast, blocks of combinational logic that perform operations between registers are characterized at RT level (e.g. adder). Intermediate results of the design process that are relevant for this thesis are thus register transfer level (RTL) design, which is characterized by a hardware description language (e.g. Verilog) and gate level netlist. As shown in Figure 2.4, the gate level netlist can directly be generated based on the RTL design by a logic synthesis. Within the synthesis process, additional information about available gates and the gate's characteristics are needed, which is provided by the liberty library (.lib) for a specific manufacturing technology [25].

Design variations on RT and gate level abstraction levels, like usage of additional RT components and alteration of utilized gates, can be realized by modifications of the Verilog design and gate level netlist, respectively. These variations may induce considerable transformations of the integrated circuit's non functional properties due to an indirect modification of the circuitry of transistors. Hence, pre-production analysis of the impact of design variations on the circuit's non functional properties is a matter of great interest. For this purpose, analog electronic circuit simulations with the program SPICE [14] can be performed in order to predict the entire behavior of the circuitry of transistors.

2.2.1 Gate Level Analysis

Instead of an analysis that relies on the basic circuitry of transistors, the analysis may also be directly performed based on the gate level netlist. Main advantages of the gate level analysis are faster computation and possible omission of the generation of the basic circuitry of transistors. Thus, non functional properties may be assessed in an earlier stage of development of an integrated circuit.

A main non functional property is an integrated circuit's timing performance. As this property cannot be calculated solely based on the gate level netlist, additional characterization data that is provided by the liberty library is also needed for various gates. By using this characterization data, the delay of each gate within the netlist is estimated and made available for subsequent analysis tools within a standard delay format (SDF) file [26] as presented in Figure 2.4. Since a gate's delay depends on whether a rising or a falling output transition is triggered as well as which input pin has provoked the alteration, not a single delay value but at least two times number of input pins different delay values are needed for each gate. In the process of generating the SDF file, the capacitances that are connected to the gates' output pins, referred to as load capacitances, have to be estimated based on the liberty library data. As shown in Figure 2.4, processed capacitances can certainly also be made available for other analysis tools. Besides load capacitances, slew rates of the input signals, which are precisely

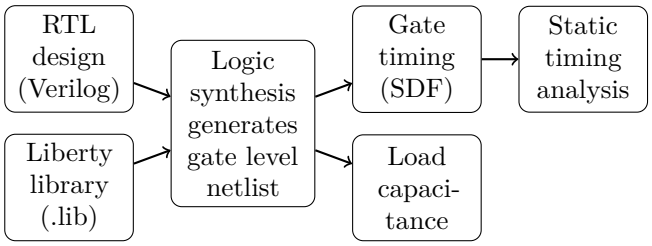


Figure 2.4: Flowchart of gate level logic synthesis with static timing analysis with major output arguments for this thesis

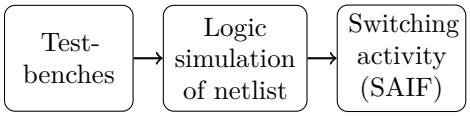


Figure 2.5: Flowchart of gate level logic simulation with major output arguments for this thesis

defined in Figure 6.1, greatly affect the gates' delays and must thus also be estimated for each gate based on liberty library data. Thereby, propagation of slew rate values has to be considered as the estimated output slew rate of a preceding gate must be used as input slew rate of the subsequent gate. Although a gate's delay also depends on temperature, supply voltage and process parameters like transistor sizings and threshold voltage, these determining factors are not incorporate within the liberty library. Thus, variations of these factors can only be respected for the whole integrated circuit by using a different liberty library. In order to calculate the timing of the complete integrated circuit, a static timing analysis (STA) [27] is finally performed based on the generated SDF file as presented in Figure 2.4.

Another fundamental gate level analysis is the simulation of each gate's pin activity, which is a major determining factor for degradation effects as stated in Section 2.1. For this purpose, a logic simulation computes the propagation of input register test data within the gate level netlist of a RT component. Thus, signal probability and transition probability of each gate's input pin are estimated and made available using the switching activity interchange format (SAIF) [28] as presented in Figure 2.5.

CHAPTER 3

Related Work

Current state-of-the-art in modeling of an integrated circuit's delay degradation is described within this chapter. This includes different modeling techniques of the physical origins of the major degradation effects negative bias temperature instability (NBTI) and hot-carrier degradation (HCD) as well as differing abstraction techniques of the fundamental models. Furthermore, a classification of published procedures to estimate the delay degradation of complete integrated circuits is carried out. Modeling techniques of the degradation effects and procedures for integrated circuits are used as starting point for the development of the proposed gate level degradation analysis as well as reference methodologies to evaluate the accuracy of the proposed analysis within the following chapters.

3.1 Modeling of Negative Bias Temperature Instability

3.1.1 Reaction-Diffusion Model

A first NBTI model, called reaction-diffusion model was proposed by [29]. It explained NBTI with a good agreement to the measurement data available at that time. The model relies on the Si–H bonds at the interface between monocrystalline silicon and silicon dioxide that is used as dielectric material. As lattice constants of Si and Si_2 do not match perfectly, Si^+ traps occur at the interface. Since these interface traps negatively affect the transistor's characteristics by increasing the threshold voltage, hydrogen is used within the production to neutralize the interface traps. However, the hydrogen may be dissolved in the course of time as illustrated in Figure 3.1 and may thus cause threshold voltage degradation. This hydrogen dissociation process at the Si-SiO_2 interface is thus used as fundamental cause of NBTI within the reaction-diffusion model.

Four interdependent processes, which are characterized by a system of differential equations, are used within the reaction-diffusion model. The top panel of Figure 3.2 introduces a schematic description of these processes. Negative gate bias causes the breakage of Si–H bonds at the Si–SiO_2 interface and interface traps as well as freed hydrogen atoms are created (reaction). The freed hydrogen drifts through the oxide (diffusion) and some H may convert to H_2, which causes higher diffusion rates. The actual diffusion of hydrogen atoms or molecules is further examined within successive publications [30]. The third process characterizes the interface between silicon dioxide and polycrystalline silicon that is used as gate material. Here, the different diffusion rates within the materials have a distinct impact on the overall drift of hydrogen molecules. Finally, the last process specifies the recovery of Si-H bonds for time intervals without negative gate bias.

Possible resulting hydrogen profiles in the oxide during NBTI stress are shown in the bottom panel of Figure 3.2. The initial interface trap generation rate depends on Si–H bond dissociation (profile 1 and 2), while the later rate depends on hydrogen diffusion (3,4 and 5). Once the diffusion front reaches the SiO_2-polysilicon interface (profile 5), the diffusivity of H_2 within the polycrystalline silicon has to also be regarded. The area under the hydrogen profile equals the generated interface traps and can thus directly be translated to threshold voltage degradation. As the reaction-diffusion model's performance is not sufficient for various degradation analysis approaches, performance oriented model simplifications do exist like the NBTI model of [32] that is adapted for analogue stress scenarios with constant temperature.

Different enhancements of the reaction-diffusion model are proposed as the basic model cannot simulate the fast degradation and relaxation processes that are apparent within new

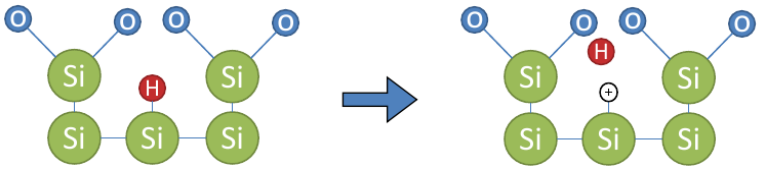


Figure 3.1: Dissociation of hydrogen at the Si-SiO_2 interface is used as fundamental cause of NBTI within the reaction-diffusion model

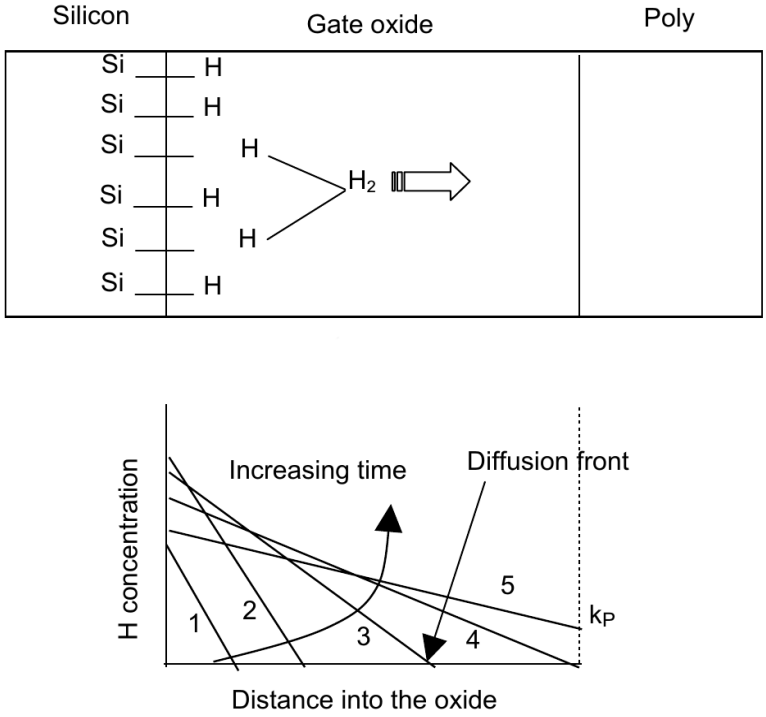


Figure 3.2: Schematic description of the reaction–diffusion model (top panel) and related hydrogen profiles in the oxide during NBTI stress (bottom panel) (Source: [31])

NBTI measurements [20]. While [33] proposes a hole assisted reduction of the activation energy, an additional hole trapping process is introduced by [34]. This hole trapping process relies on a different type of pre-existing defects and is only determined by tunneling probabilities. As shown by [19], there is also a significant difference between measurement data and the corresponding reaction-diffusion simulations for long time relaxation scenarios, which is a major characteristic of the model and cannot be modified with a model enhancement. Furthermore, simulation results of the reaction-diffusion model are independent of the stress signal's frequency [20], which does not correspond to measurement data [21].

3.1.2 Switching Trap Model

Due to the discrepancies between new measurement data and simulation results of the reaction-diffusion model, a fundamentally different NBTI model called switching trap model is proposed by [4]. This model is able to simulate correct long time relaxation curves [19] and has caused a complete paradigm shift in NBTI modeling as the reaction-diffusion model is only rarely used nowadays [22].

Two different physical causes of NBTI are used within the switching trap model. The first one is the breakage of Si-H bonds in the Si-SiO_2 interface as utilized within the reaction-diffusion model. However, these defects are only modeled with a two state approach that covers the bond breakage and annealing processes, but does not incorporate the hydrogen diffusion process. Si-Si bonds that always exist in the silicon dioxide are used as the second physical cause of NBTI. As shown in Figure 3.3, these Si-Si can be weakened and one of both silicon atoms may enter into a connection with a different atom leaving the other silicon atom within a Si^+ trap state (right panel of Figure 3.3). These Si-Si defects are modeled with a four state approach that is presented in Figure 3.4. Starting with the initial neutral stable state, a charge exchange due to tunneling weakens the Si-Si bond and causes a positive metastable state. A subsequent temperature activated structural relaxation causes the Si-Si bond breakage described above (positive stable state). Alternatively, the structural relaxation may occur before charge exchange via a neutral metastable state. Naturally, each state transformation may also happen the other way round as a recovery process.

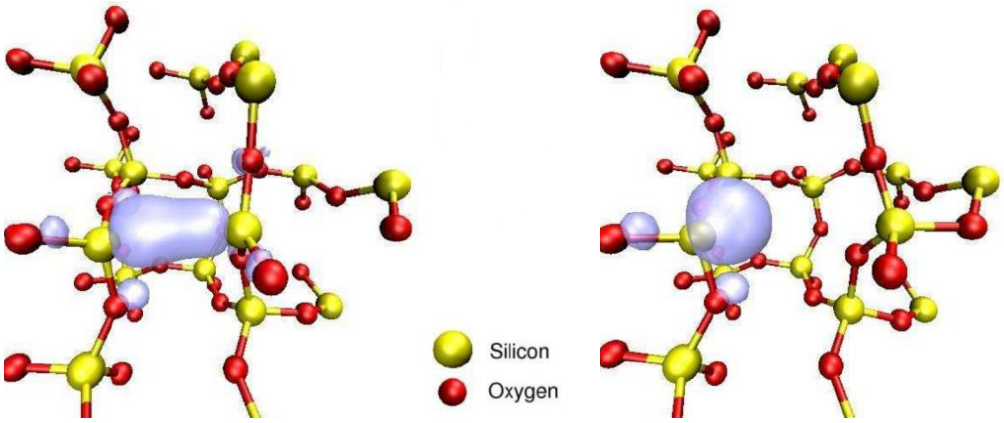


Figure 3.3: Breakage of Si-Si bonds within the silicon dioxide is used as additional cause of NBTI within the switching trap model (Source: Provided by the author of [22]))

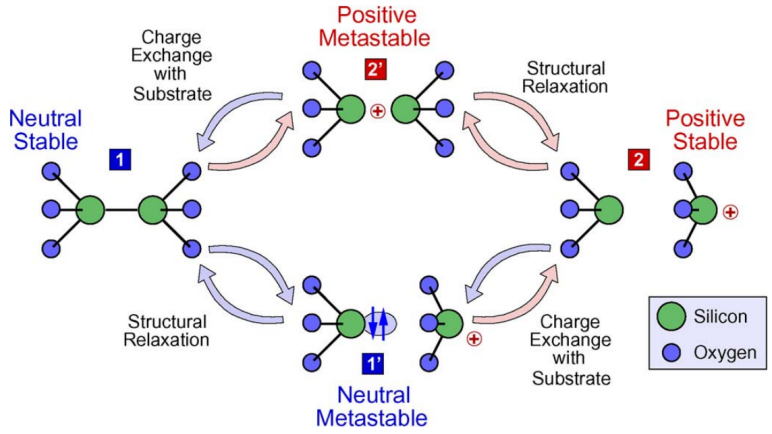


Figure 3.4: Four state modeling approach for Si-Si defects within the switching trap model (Source: [22])

A transistor's overall threshold voltage is theoretically affected in a unique way by each trap being in a positive or occupied state. If these traps are considered in the physical Poison equation of the transistor, the corresponding individual threshold voltage shifts can be obtained [35]. However, this computation is very time-consuming for devices with many traps and an approximation based on the assumption that a trap charge can be described as a homogeneous charge density is often used to assess each trap's impact on the threshold voltage [35].

NBTI simulations rely on the individual two and four state physical models and track each trap's probability of being in a positive or occupied state. For each constant stress scenario, there are always non zero probabilities for both occupation and recovery of each trap. Thus, each trap's probability of being in an occupied state has to be estimated based on very short time steps. Consequently, the computational effort of switching trap NBTI simulations is very high as a supposably large number of traps is processed separately with short time steps. Scenarios with varying temperatures or gate voltages are directly supported, since each trap's parameters are adjusted correspondingly. As a conventional occupation state is used for each trap, which can either be fully charged or fully uncharged, the statistical nature of the degradation process is simulated (see Figure 2.2). Therefore, several statistical simulations need to be averaged in order to estimate the expectation value of the threshold voltage degradation.

Each trap can also be described by the mean time to enter and to leave a positive or occupied state, called capture time τ_C and emission time τ_E, respectively. Capture and emission times heavily depend on gate voltage as well as temperature. For various combinations of gate voltage and temperature, the values of τ_C and τ_E are significantly different. However, very similar capture and emission times can occur for particular combinations of temperature and gate

voltage.

Based on measurement data of NBTI degradation and relaxation curves for various gate voltages and temperatures, the total number of defects as well as each defect's parameter values of two or four state approach are adjusted until measurement data and simulation results coincide. The resulting parameter values as well as capture and emission times for different gate voltages and temperatures are then specified within a trap list. Here, only two temperatures are characterized as capture and emission times of an arbitrary temperature can be calculated based on the characterized values due to the temperature activated nature of the structural relaxation.

Recent examinations show that the breakage of O-H bonds within the silicon dioxide may also be involved in NBTI degradation [36]. This additional type of defects can also be integrated within the trap list using capture and emission time values. In this way, abstractions or simplifications of the switching trap model that rely on capture and emission times of the trap list are not hindered by these recent examinations and directly benefit from an increased accuracy of the switching trap approach.

In conclusion, the switching trap model precisely simulates the NBTI degradation for arbritary stress scenarios, but the performance requirements stated in Chapter 1 cannot be fulfilled by the model. Even a performance oriented model simplification [37] cannot be used for the described operation conditions as a huge number of defects still has to be tracked separately.

3.1.3 Capture-Emission Time Model (CET)

The initial model assumption, as proposed by [38], is that the discrete switching traps, which can either be fully charged or fully uncharged, can be replaced by a continuous statistical process. Hence, occupation probabilities are used to directly calculate the expectation value of the threshold voltage degradation. This step is necessary when compacting the explicit traps into a two-dimensional trap distribution $\Delta V_{th}(\tau_C, \tau_E)$ as presented in Figure 3.5. This distribution is hereinafter referred to as capture-emission time (CET) map. A CET map is generated for a specific temperature, gate voltage in a stress condition and gate voltage in a relaxation condition by first estimating capture times of all traps for the stress condition and emission times of all traps for the relaxation condition. Each trap's expected impact on threshold voltage degradation, as stated in Section 3.1.2, is subsequently binned by estimated capture and emission time within a selected τ_C-τ_E grid.

Within Figure 3.5, this grid of different capture and emission time bins is provided by the resolution that is utilized for the capture and emission time axes. Besides, the impact of each τ_C-τ_E bin on the threshold voltage degradation, which is represented by the color coding in Figure 3.5, is determined by the summation of the threshold voltage shifts of all traps with corresponding capture and emission time values. The color code of each bin thus represents the threshold voltage shift that occurs if all traps with corresponding time constants are in a charged or occupied state.

The effect that capture and emission process may occur at the same gate voltage is accounted for by a multiplication of threshold voltage impact and equilibrium occupancy difference (EOD) for each trap [35]. EOD is defined as probability that a trap is charged due to a scenario with specific gate voltage and temperature as well as infinite degradation time. The impact of traps with equal capture and emission times is thus diminished by 50%. Finally, a Gaussian convolution is applied to the CET map to soften inaccuracies due to the binning.

As can be seen in Figure 3.5, capture and emission times are weakly correlated. While capture and emission times spread over several orders of magnitude, usually capture and emission time are in the same order of magnitude. Key benefits of the CET approach are thus the direct simulation of the threshold voltage expectation value and the possibility of improved performance due to total number of τ_C-τ_E bins of the CET map being significantly smaller than the number of traps.

During CET simulations, occupation probabilities $P(\tau_C, \tau_E, t)$ of every τ_C-τ_E bin are calculated based on given stress scenarios. The occupation probability over time is described by

$$P(\tau_C, \tau_E, t + \Delta t) = 1 - (1 - P(\tau_C, \tau_E, t)) \cdot e^{-\frac{\Delta t}{\tau_C}} \tag{3.1}$$

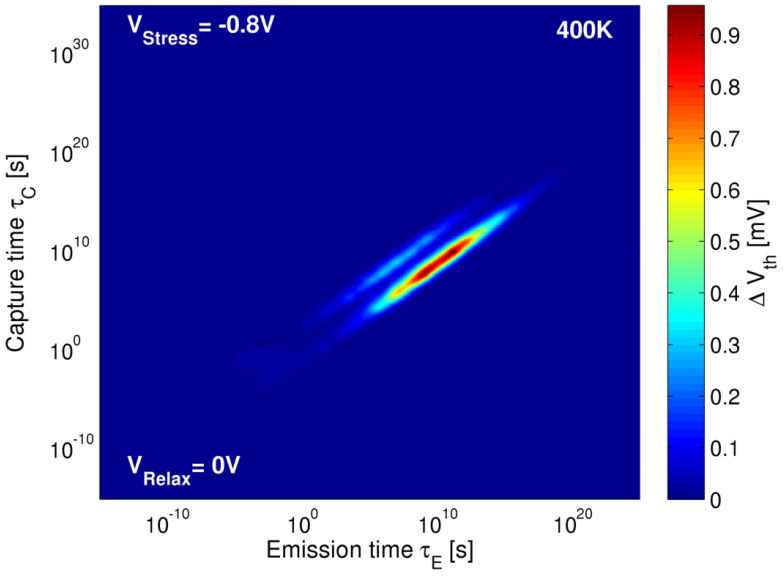


Figure 3.5: CET map of the 22 nm technology for 400 K and stress voltage -0.8 V with a resolution of 320 capture and 256 emission time bins. The color coding of each bin represents the combined threshold voltage shift that can be triggered by all traps with corresponding time constants. Provided trap list of the switching trap model (see Section 5.1.1) are used to derive the CET data.

if the system is stressed for a time Δt and

$$P\left(\tau_C, \tau_E, t + \Delta t\right) = P\left(\tau_C, \tau_E, t\right) \cdot e^{-\frac{\Delta t}{\tau_E}} \tag{3.2}$$

if the stress is removed for a time Δt. The threshold voltage degradation due to NBTI over time is computed as

$$\Delta V_{th\ tran}(t + \Delta t) = \int d\tau_C \int d\tau_E P\left(\tau_C, \tau_E, t + \Delta t\right) \cdot \Delta V_{th}\left(\tau_C, \tau_E\right) \tag{3.3}$$

Resulting occupation probabilities $P\left(\tau_C, \tau_E\right)$ for elementary degradation scenarios are demonstrated within Figures 3.6 and 3.7. During a stress interval, the occupation probability builds up in the capture time dimension and is thus uniformly distributed in the τ_E dimension as shown in Figure 3.6. A subsequent relaxation interval causes a decline of the occupation probability that starts at very low emission times and is uniformly distributed in the τ_C dimension. In this way, rectangularly shaped occupation probabilities $P\left(\tau_C, \tau_E\right)$ are produced for specific degradation scenarios with a single stress interval that is followed by a single relaxation interval as shown in Figure 3.7.

Calculation of threshold voltage degradation can also be implemented using RC-circuits [38]. For each τ_C-τ_E bin a RC-circuit as presented in Figure 3.8 is used. While the capacitance C_{ref} characterizes the bin's impact on threshold voltage degradation, mean capture and emission times are determined by two resistances connected in series with diodes. Thus, the overall threshold voltage degradation can be simulated using a parallel connection of several RC-circuits.

3.1.4 Analytical Approach

Precise threshold voltage degradation traces that are simulated by the NBTI models described above, contain fine structures of short time degradation and relaxation processes for alternating stress signals (see Figure 5.22). These fine structures are disregarded by analytical NBTI models and thus only the high side envelope of the actual degradation trace shall be considered. In addition, the time dependence of the high side envelope is characterized with a simple analytical function. In this way, the complexity of the NBTI models is greatly reduced and threshold voltage shifts due to degradation times of several years can be simulated almost instantly.

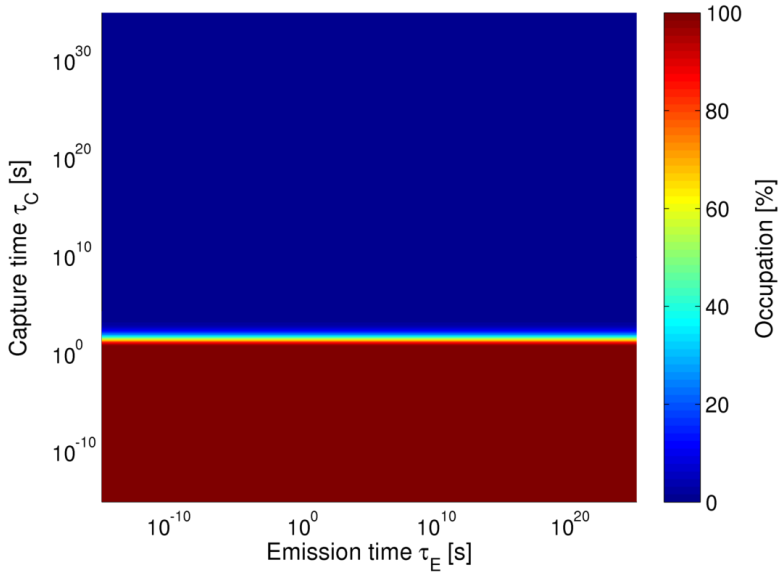


Figure 3.6: Occupation probability $P(\tau_C, \tau_E)$ for a specific degradation scenario with a single stress interval

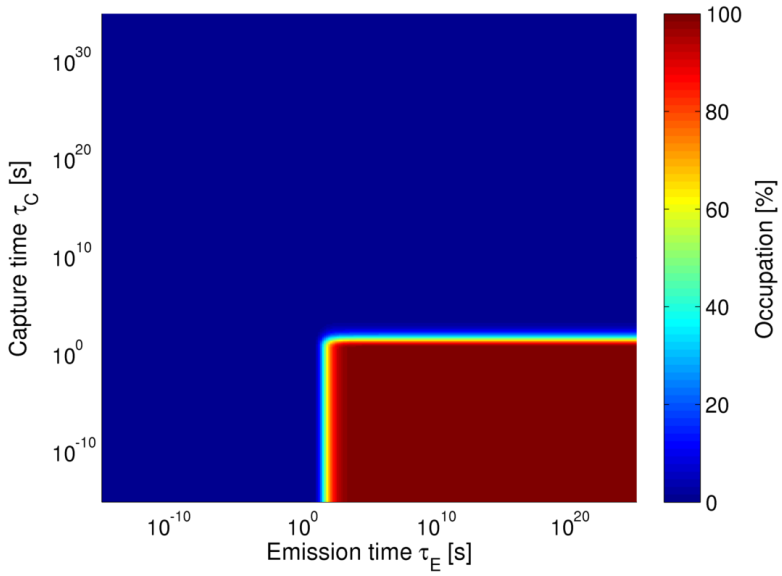


Figure 3.7: Occupation probability $P(\tau_C, \tau_E)$ for a specific degradation scenario with a single stress interval that is followed by a single relaxation interval

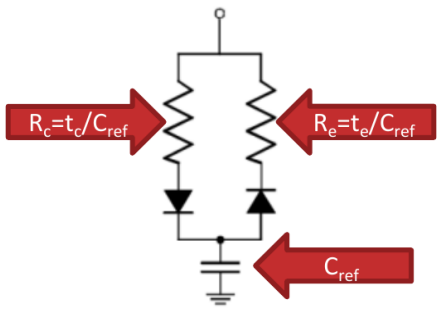


Figure 3.8: RC-circuit based modeling of the threshold voltage shift caused by a single bin of the CET map

However, varying stress scenarios, which may include longer relaxation phases at arbitrary points in time, cannot be supported by these analytical models, as the assumption of explicit time dependence would be violated.

Several analytical NBTI models, which can be divided into two groups, are proposed in the literature. The first group is originally derived from the reaction-diffusion model and uses a power law $\Delta V_{th} \approx a \cdot t^b$ as analytical function [39]. In contrast, a logarithmic function is used for the time dependence of the threshold voltage degradation within the second group of models [17]. This group of models relies on the long time characteristics of the switching trap model. Parameter values of the analytical functions are usually estimated by the use of NBTI simulations for long time scenarios. However, parameter values of the logarithmic function can even be directly determined based on trap lists of the switching trap model [40].

3.1.5 Mission Scenario Aware Analytical Approach

This section refers to the analytical model of [41] that is called "mission scenario aware analytical approach" (MSA) throughout this thesis. The model relies on the switching trap model and logarithmic fits of the high side envelopes of the ΔV_{th} curves as [40]. However, the regular analytical approach is enhanced with a modeling technique to support varying stress scenarios.

At first, capture time τ_C, emission time τ_E and activation energy E_0 are averaged for numerous defects and the model parameters K, B and β are specified for the respective technology. Afterwards, the logarithmic degradation curves are estimated for constant scenarios using Equation 3.4.

$$\Delta V_{th} = A(V,T) \cdot \log\left(1 + \frac{t_s}{\tau_C} \cdot \left(\frac{1}{1 + \left(\frac{t_r}{\tau_E}\right)^{\beta}}\right)\right) \tag{3.4}$$

t_s and t_r refer to the scenario's overall stress and relaxation time, respectively. Impact of temperature T and gate voltage V is incorporated using the multiplication factor $A(V,T)$, which also requires the Boltzmann constant k_B and the technology's oxide thickness t_{ox}.

$$A(V,T) = K \cdot \exp\left(\frac{-E_0}{k_B \cdot T}\right) \cdot \exp\left(\frac{B \cdot V}{t_{ox} \cdot k_B \cdot T}\right) \tag{3.5}$$

Additionally, a basic logarithmic relaxation curve is specified for each constant scenario that can be used for various initial threshold voltages at the beginning of a relaxation interval by using a simple scaling approach.

The modeling technique to support varying stress scenarios is presented in Figure 3.9. Within this example, a constant high degradation scenario persists during the intervals prior to t_1 and after t_2 (labeled with up_0 and up_2). Between t_1 and t_2 the device is subjected to a constant low degradation scenario (interval up_1). The threshold voltage degradation of the up_0 interval is directly specified by the high degradation scenario's logarithmic curve. As the estimated value at t_1 is larger than the t_1 value of the low degradation scenario's logarithmic curve (marked by dashed line and $\Delta V^*_{th,1}$), there should first be relaxation in the interval up_1. Therefore, the respective pre-characterized relaxation curve is scaled to the estimated threshold voltage at t_1 (line 6). However, relaxation can only occur within the MSA model as long as the threshold voltage is larger than the value of the low degradation scenario's logarithmic curve (dashed line) that corresponds to the same degradation time. Thus, the relaxation curve (line 6) is passed over to the respective degradation curve (line 7). As the estimated value at t_2 is smaller than the t_2 value of the high degradation scenario's logarithmic curve, there should only be degradation in the interval up_2. This characteristic is implemented within the MSA model by using the high degradation scenario's logarithmic curve starting with the threshold voltage at t_2 for the interval up_2 (line 11).

In this way, threshold voltage degradation due to varying stress scenarios that may consist of more than two constant stress conditions, can efficiently be simulated with an analytical approach.

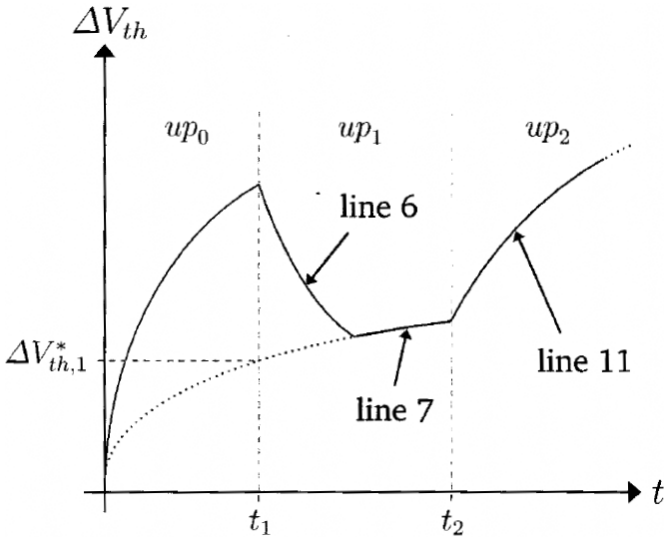


Figure 3.9: Modeling technique to support varying stress scenarios based on transitions between pre-characterized analytical degradation and relaxation curves (Source: [42])

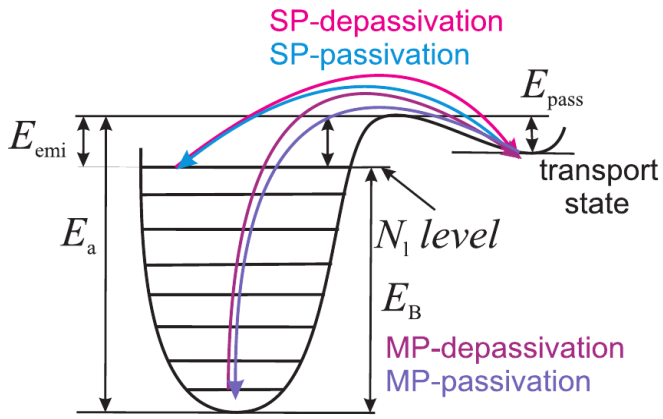


Figure 3.10: Multi-stage process of recent hot-carrier degradation models (Source: [5])

3.2 Modeling of Hot-Carrier Degradation (HCD)

Hot-carrier degradation presumably also relies on Si^+ traps at the Si-SiO_2 interface (see Figure 3.1). However, a different process causes the activation of the traps. While the NBTI models are based on a temperature activated process, the hot-carrier effect is thought to be caused by fast "hot" charge carriers within the channel of a transistor. The first proposed model for hot-carrier degradation (HCD), referred to as lucky electron model [43], characterizes the possibility that the fast charge carriers are redirected, traverse the silicon dioxide and directly cause the activation of traps. As recent technologies use very low gate voltages, Si-H bonds cannot be broken by a single fast charge carrier. Therefore, the initial hot-carrier model is enhanced by a multi-stage activation process of Si-H bonds like a quantum harmonic oscillator [5] and several fast charge carriers are thus needed to cause a depassivation of a Si-H bond (see Figure 3.10).

As NBTI and HCD are each based on the same type of defects, recent studies analyse whether NBTI and HCD are dependent on one another and must thus be simulated in a combined way [44]. Furthermore, it is also possible to characterize the hot-carrier degradation or the combined NBTI and HCD degradation using the CET map approach [45].

3.3 Degradation Analysis of Digital Circuits

3.3.1 Stress Conditions in CMOS Logic Gates

[46] already accurately specifies stress conditions in CMOS logic gates that induce NBTI or HCD. A PMOS transistor degrades due to NBTI, when the transistor is in inversion and the gate terminal is negatively biased with respect to source and drain. The probability that the PMOS transistor is in inversion is of course directly determined by the signal probability SP of the gate terminal in the way that inversion probability $P_{Inversion}$ equals $1 - SP$. If the transistor is in inversion, the transistor's source has also to be at supply voltage V_{DD} in order to fulfill the stated NBTI stress condition. Thus, all transistors that are located in a transistor stack between analyzed transistor and a supply line, have to be in inversion. The overall probability of a NBTI stress condition P_{NBTI} of an analyzed transistor is hence dependent on inversion probabilities of the analyzed transistor and all transistors located above in the stack.

$$P_{NBTI}(\text{Analyzed transistor}) = \prod_i P_{Inversion}(\text{Transistor}_i) \tag{3.6}$$

In detail, Equation 3.6 is only applicable for independent gate input signals for each transistor. If this requirement cannot be fulfilled, a worst case assumption is that all transistors in the stack tend to conduct at the same time. Hence, the probability of a NBTI stress condition P_{NBTI} is determined by the minimal inversion probability $P_{Inversion}$ of all involved transistors.

For CMOS logic gates, the probability of NBTI stress conditions for each PMOS transistor within the gate may thus be dependent solely on the signal probability of single input pins for particular gates (e.g. NAND). However, some gates require the consideration of signal probabilities of all input pins in order to estimate the probability of NBTI stress conditions for each PMOS transistor. As can be seen in Figure 6.7b, NOR gates belong to this kind of gates for instance.

For hot carrier degradation, a considerable current must flow in the channel of a transistor as described in Section 2.1.2. Thus, a transition from "off"- to "on"-state must occur at a transistor in the transistor stack and all other transistors in the stack must be "on" [46]. In order to estimate the overall HCD probability of a transistor, probability of a state transition and probability of a conducting path from a supply line to the gate's output have to be multiplied. In detail, the probability of a state transition is determined by the sum of the transition probabilities of all transistors in the stack and the probability of a conducting path is characterized by the product of the inversion probabilities of the complete transistor stack similar to Equation 3.6.

As stated for NBTI, the probability of HCD stress conditions for each transistor within a CMOS logic gate may thus be dependent on the transition probability of only a single gate input pin or several pins. Due to this dependence of NBTI and HCD on signal and transition probabilities of several input pins for some gates, [13] proposes pin reordering for this kind of gates as a method to reduce the degradation that occurs due to disadvantageous probabilities at the gate's primary input pins.

3.3.2 Transistor Level Analysis

Timing degradation of a complete integrated circuit can be analyzed at transistor level and industrial analysis tools such as RelExpert [47] are already available. In detail, the transistor level degradation analysis is build up of three independent parts. At first, the fresh transistor netlist has to be simulated in order to compute per transistor stress conditions. Afterwards, threshold voltage shifts are estimated for each transistor and the simulation of the degraded transistor netlist is finally performed to obtain the circuit's degraded timing. Industrial tools such as RelExpert mainly offer an interface to facilitate this transistor level degradation analysis without relying on a particular method to estimate per transistor threshold voltage shifts.

Individual analogue stress conditions of each transistor have a huge impact on the degraded timing of a complete path [16] and have to thus be respected for a precise timing analysis. As SPICE circuit simulations at transistor level are very resource-intensive, only short time stress is normally simulated to compute these per transistor stress conditions. However, these short time stress profiles contrast the long term character of the degradation effects and a method

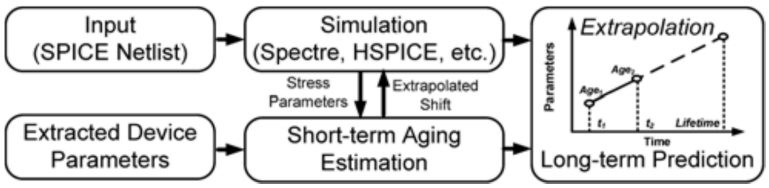


Figure 3.11: Transistor level degradation analysis based on extrapolations of transistor parameter drifts caused by short term stress (Source: [17])

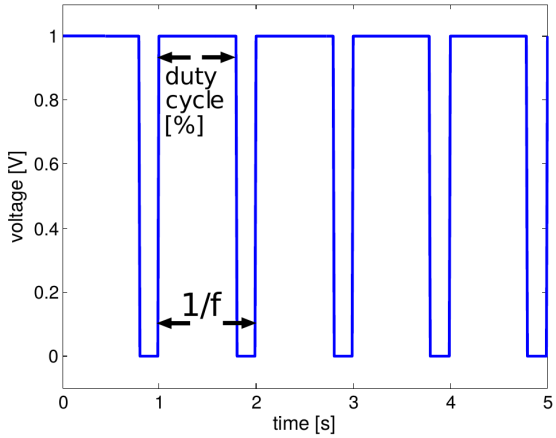


Figure 3.12: Definition of frequency f and duty cycle dc of a periodic stress signal

to extrapolate the degradation behavior is needed. In this regard, individual parameter drifts (threshold voltage) of each transistor can be computed for the precise analogue short time stress condition and the long term prediction is based on an extrapolation of these parameter drifts as shown in Figure 3.11. Although this method relies on very precise individual stress conditions, large inaccuracies are introduced by the extrapolation [17]. Another extrapolation method focuses on the abstraction of the individual stress conditions by using individual signal and transition probabilities for each transistor. The stress conditions can thus be characterized based on periodic signals specified by frequencies and duty cycles [13] as presented in Figure 3.12. Since dependence of the NBTI degradation on the stress signal's frequency is only a secondary effect, a fixed frequency and an abstraction of the stress condition that solely relies on duty cycle may even be used [39]. The subsequent long term extrapolation is realized as part of the computation of the threshold voltage shift for each transistor. In particular, analytical approaches that incorporate the impact of the stress signal's frequency and duty cycle for static stress conditions are utilized (see Section 3.1.4). In this way, the individual stress conditions are less precise but the extrapolation accuracy is improved considerably [17]. As a consequence, the analysis approach is bound to static overall stress conditions and variations of the usage pattern or temperature changes cannot be supported. Furthermore, the IR drop effect, which characterizes voltage variations for individual transistors due to resistances within the supply network, can also be incorporated in the transistor level degradation analysis as these individual stress conditions can also be estimated based on SPICE simulations [48].

Different technical approaches can be used for the simulation of the circuit's degraded timing. An obvious approach is the creation of individual specification files for each transistor and the transformation of the respective threshold voltage parameters within these files. Main drawback of this approach is though a clearly inefficient utilization of a huge number of individual files. Alternatively, the threshold voltage of each transistor can also be adjusted using voltage controlled voltage sources within SPICE [17]. The implementation of a NBTI degradation model as a SPICE circuit that can be integrated within each transistor as proposed by [38] is a direct way to simulate the degradation effect on transistor level. However, impractical long term SPICE simulations are thus needed to model timing degradations caused by standard usage scenarios.

	NBTI model	Scenario support	Delay estimation	Timing analysis of a circuit
Wang et al. [39]	Analytical power law	Worst case	Polynomial fit of SPICE characterization	Distinct analysis
DeBole et al. [49]	Analytical power law	Worst case	Linear fit of SPICE characterization	Distinct analysis
Lorenz [50]	Analytical power law	Worst case	Linear fit of SPICE characterization	Distinct analysis with path reduction
Huard et al. [16]	Analytical power law	Worst case	Linear fit of SPICE characterization	Distinct analysis
Mintarno et al. [51]	Analytical power law	Worst case	SPICE recharacterization	Standard analysis with uniquified and modified .lib
Kükner et al. [52]	CET approach	Changing stress	SPICE recharacterization	Standard analysis with uniquified and modified .lib
Cao et al. [17]	Analytical logarithmic law	Worst case	Extraction based on delay for various V_{DD} values	Distinct analysis
Barke [42]	Piecewise analytical log. law	Varying scenarios	Linear fit of SPICE characterization	Distinct analysis with path reduction

Table 3.1: Different approaches for a gate level degradation analysis that are proposed within the literature. Different columns comprise summaries of the approaches' basic methods to estimate threshold voltage shifts, timings of single gates and delays of complete circuits as well as the possible support of varying mission scenarios.

3.3.3 Gate Level Analysis

As stated in Section 2.2.1, the timing analysis of a circuit may also be directly performed based on the gate level netlist. Various approaches have been proposed for a gate level degradation analysis, which are summarized in Table 3.1. Main commonality of all these approaches is that threshold voltage shifts of a single PMOS NMOS pair are used to estimate the degraded delay of a single path through a gate. In this way, degraded delays of an inverter and other single-stage gates (e.g. NAND) can be directly simulated [17] and multiple-stage gates have to be subdivided into several single-stage gates [46] for the timing analysis. Main differences of the proposed approaches are the utilized methods to estimate threshold voltage shifts, which directly affect the approaches' possible support of varying mission scenarios. Furthermore, different methods to estimate degraded timings of single gates as well as degraded delays of complete circuits are applied within the proposed gate level approaches. Utilized methods of each approach and the possible support of varying mission scenarios are also summarized in Table 3.1.

Many proposed gate level approaches use analytical functions to calculate NBTI induced threshold voltage shifts as described in Section 3.1.4. In particular, both power law and logarithmic functions are utilized for different gate level approaches, which thus rest upon long term characteristics of reaction-diffusion and switching trap model, respectively. Consequently, these gate level approaches are bond to static worst case conditions and do not support varying mission scenarios. Only marginal improvements can be achieved in this regard by respecting long term off states (mission profiles) within the computation of each gate's NBTI stress condition as implemented within the gate level approach of [16] that is presented in Figure 3.13. In detail, the already applied abstraction of a gate's stress condition that is based on periodic signals (see Section 3.3.2) is further simplified by changing the signal's duty cycle as function of long term off states [53].

By utilizing the CET model (see Section 3.1.3), the gate level approach of [52] fully supports varying NBTI stress conditions at the cost of highly reduced performance. Furthermore, varying temperatures and supply voltages can still not be respected within this approach. Finally, due

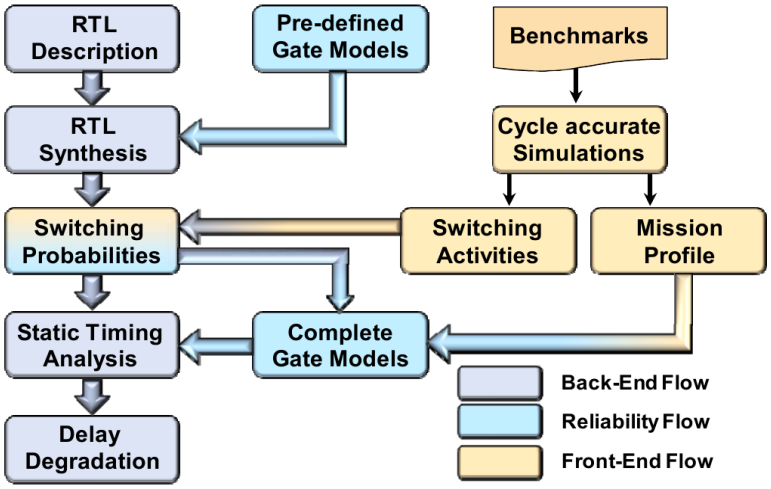


Figure 3.13: Flow chart of the gate level degradation analysis of [16]

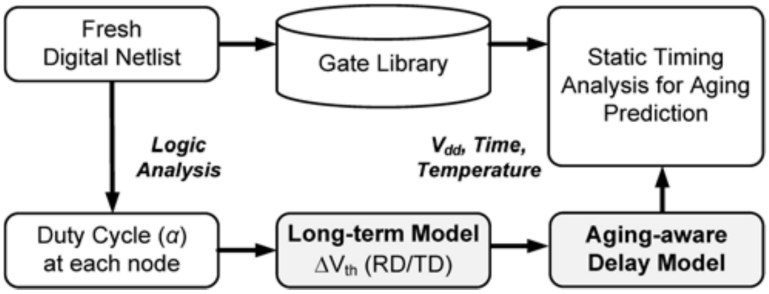


Figure 3.14: Flow chart of the gate level degradation analysis of [17]

to piecewise analytical functions within the mission scenario aware analytical NBTI model (see Section 3.1.5) varying mission scenarios are theoretically fully supported by the gate level degradation analysis of [42]. As the accuracy of this approach is though still debated, an extensive evaluation within Chapter 5 further examines the mission scenario aware analytical NBTI model.

SPICE measurements of single gates are performed by many proposed gate level approaches to estimate each gate's timing degradation as function of the respective threshold voltage degradation. Instead of tabularized values for various threshold voltage shifts, linear or quadratic regression techniques are used to minimize the size of the generated lookup tables. These generated lookup tables can then be used to efficiently estimate each gate's degraded timing during the gate level analysis. Unfortunately, no regression parameters of a gate's timing degradation as function of threshold voltage shift are included in any of the stated publications. An exception of this common procedure is the timing estimation method of [17], which uses already available timing data of a gate's delay for various supply voltages to identify the gate's delay as function of ΔV_{th}. Thus, the respective gate level degradation analysis approach, that is presented in Figure 3.14, does not incorporate SPICE measurements for the "Aging-aware Delay Model". The third proposed method to compute each gate's timing degradation simply relies on a SPICE based liberty library generation (see Section 2.2) for the degraded circuit. In general, two different gates of the same type (e.g. NAND2) degrade differently and can thus not be specified by the same entry of the liberty library. Hence, uniquified gates have to used within the gate level netlist and an appropriate liberty library entry has to be generated for every single gate. Main drawback is thus the potentially huge size of the liberty library and the approach's poor performance as the SPICE measurements are performed as part of every gate level degradation analysis.

By using the liberty library representation of each gate's degraded timing, final delay estimations of a complete circuit can be performed by standard static timing analysis tools (see

Section 2.2.1). Due to individual representations of the gates' timing degradations, all other proposed gate level analysis approaches use distinct timing analysis techniques for the complete circuit. In this regard, improved timing analysis methods can also be utilized to reduce the number of signal path that have to be estimated [54].

3.4 Summary

Different NBTI models are presented that will be used as starting point for the development of the NBTI simulation method within the proposed gate level degradation analysis. Besides, these models are utilized as reference models to evaluate the accuracy of the proposed methodologies within the following chapters. At first, two diverging physical explanations of the NBTI effect are introduced. The first approach, which is called reaction-diffusion model, uses a hydrogen dissociation process at the Si-SiO_2 interface and a subsequent recovery of Si-H bonds based on the hydrogen diffusion characteristics. On the other hand, the novel switching trap model asserts that the NBTI effect is mainly determined by the temperature activated bond breakage of Si-Si bonds within the gate oxide. This model is able to correctly simulate long time relaxation curves and has thus caused a complete paradigm shift in NBTI modeling as the reaction-diffusion model is only rarely used nowadays.

Varying methods for a performance oriented simplification of the fundamental switching trap model are also introduced. Initially, the capture-emission time (CET) model is presented that replaces the discrete switching traps by a continuous statistical process. Besides, the explicit traps are compacted into a two-dimensional trap distribution, which is called CET map. NBTI simulations thus rely on the calculation of occupation probabilities for each bin of the CET map. However, main drawback of this approach is that the fundamental CET map is bound to preassigned values of supply voltage and temperature. Furthermore, different analytical approaches are introduced that efficiently model the long time characteristics of the NBTI degradation but are bound to static stress conditions. However, the approach of [41] uses logarithmic fits for the NBTI degradation of different constant scenarios to analytically model the degradation that is caused by varying mission scenarios. This approach may thus be suitable for the degradation analysis that is proposed within this thesis. Finally, the main concepts of the modeling of hot-carrier degradation are introduced and the possibility of a HCD simulation based on the CET approach is highlighted.

In order to introduce a simulation procedure for the timing degradation of a complete digital circuit, the occurrence of NBTI and HCD stress conditions is first specified. Afterwards, the main assumptions to characterize the transistors' individual stress scenarios for long time simulations are introduced. Different transistor level as well as gate level procedures already exist that simulate the delay degradation of a complete integrated circuit based on the efficient simulation methods of the degradation effects and the individual stress conditions for each transistor. Thus, the main concepts of the conversion between individual degradation states of several transistors and the delay degradation of a circuit are summarized. These different concepts like SPICE simulations for the degraded timings of various gates and modifications of the liberty library form the starting point for the development of the gate level degradation analysis that is proposed within this thesis. Finally, a classification of the published procedures is carried out to identify the main weaknesses of each procedure related to a mission scenario aware analysis.

CHAPTER 4

Mission Scenario Aware Gate Level Degradation Analysis

In order to develop a precise procedure for the simulation of an integrated circuit's timing degradation, the basic concept of the proposed timing degradation analysis of Section 1.4 is elaborated within this chapter. Hence, individual parts of the basic concept are realized based on the presented fundamentals (see Chapter 2) as well as the state-of-the-art in delay degradation modeling that is described in Chapter 3. In order to classify the potential application of the proposed procedure, essentials considerations for the supported mission scenarios are stated. Afterwards, the usage of industrial tools with corresponding interfaces is defined more precisely and requirements for missing methodologies are specified to finally expose the entire degradation analysis procedure.

4.1 Mission and Stress Scenarios

As stated in Chapter 1 and Figure 1.2, idle phases and variations of influencing factors (e.g. temperature) greatly influence a device's degradation. Variations of these parameters over time have to thus be regarded within degradation simulations and a standardized specification of the changing stress state is required for the tool flow as well as the degradation models. Since the process to determine the actual changing stress states of the integrated circuit's devices is split into two parts, the different specifications "mission scenario" and "stress scenario" are used within this thesis.

At first, the actual application of an integrated circuit within a product directly influences the degradation of each device within the circuit in the same way. Relevant parameters within this context are ambient temperature, supply voltage, overall off times and different system states (e.g. active/idle). Variations of these parameters have to be specified for an appropriate lifetime of the circuit based on a reasonable time step like minutes or hours. This order of magnitude is appropriate for a scenario's time step, as variations for very short periods of time have normally little impact on the overall degradation after several years. In addition, the major parameter ambient temperature can only be changed slowly in reality. Specification of the described variations is hereafter referred to as mission scenario and a XML file format defined by [55] is used for this purpose. As shown in Figure 4.1, mission scenarios have to be provided by the producer of the product, since relevant information of application of the integrated circuit and predicted usage of the product are needed.

In addition to the mission scenario specification, transformations of individual stress states are needed for each device within an integrated circuit to precisely simulate the timing degradation

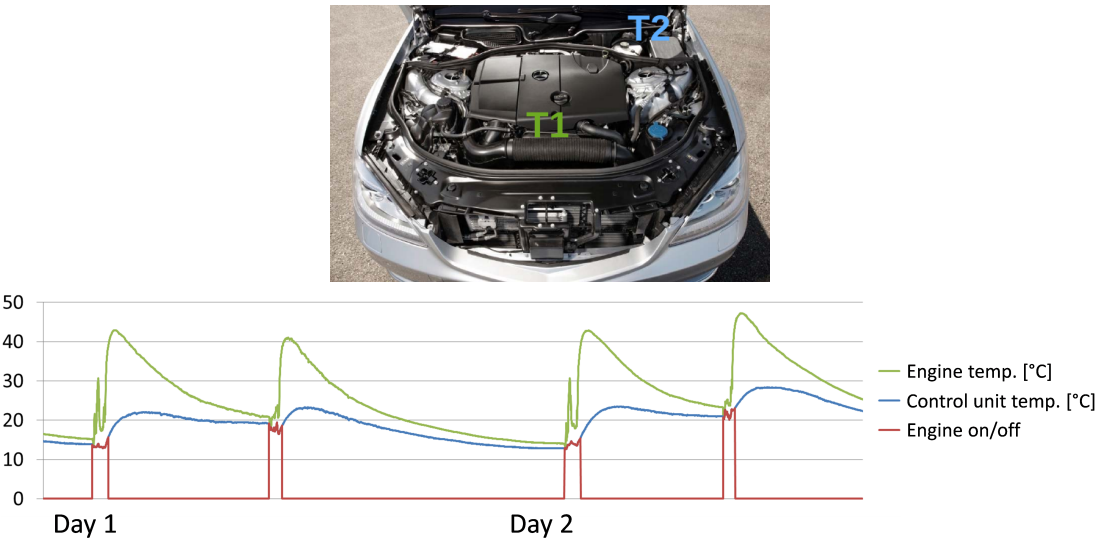


Figure 4.1: Overall usage of a product and different applications of integrated circuits (engine T1 or control unit T2) determine the mission scenario (bottom panel) for each circuit (Source: Provided by the author of [55])

[13]. As this information cannot be provided by the producer of a product, the manufacturer of the integrated circuit has to specify these stress states, which are hereafter referred to as stress scenarios, based on the circuit's mission scenario. Level of detail within the stress scenario may vary depending on utilization of transistor, gate or RT level abstraction [56]. Most relevant parameters of the individual stress scenarios on gate level are signal and transition probabilities of each gate's input pins, which may even be specified individually for each system state that is characterized within the mission scenario. Furthermore, self-heating and the IR-drop effect cause individual alterations of overall ambient temperature and supply voltage. Hence, mission and stress scenarios are clearly related to each other and required tools and procedures to calculate all information of the stress scenario are described within [12], at which computation of signal and transition probabilities based on a logic simulation (see Figure 2.5) is the most important part. In order to avoid the parsing of different file formats within the degradation models, [55] proposes an extension of the mission scenario XML file format concerning relevant information for each gate's stress scenario.

Stress scenarios may be highly diverse in the course of several years, but a worst case stress state is normally considered for the final degradation phase. This type of scenario is defined with the intention that an analysis should estimate whether a circuit still operates correctly if a worst case state is utilized for a realistic duration after several years of typical stress. Thus, varying stress states over the course of several years result in realistic predictions of the corresponding threshold voltage degradations and the final worst case state only slightly affects the threshold voltage degradations due to a very short duration of this stress phase in comparison to the overall simulated lifetime. This final worst case state however greatly influences the estimation of the timing degradations, as a circuit's timing is clearly directly affected by some of the stress scenario's parameters like supply voltage and temperature. Depending on how well the stress scenario is determined it may though be meaningful to perform the computation of the timing degradations using a more realistic stress state instead of a worst case assumption.

4.2 Gate Level Degradation Analysis within an Industrial Tool Flow

Main objective of this thesis is the development of a methodology that facilitates an accurate, efficient and reliable computation of the degraded timing of RT level components. Thus, Figure 4.2 proposes an appropriate tool flow for the computation of degraded timings, which utilizes the gate level abstraction as this abstraction level facilitates efficient timing estimations without introducing clear limitations of the methodology's accuracy. By comparing Figure 4.2 and 2.4, it is apparent that the proposed methodology is integrated as an additional component within an industrial tool flow to utilize industrial tools for timing estimations of complete signal path. Here, standard delay format (SDF) files (see Section 2.2.1) are used as main interface between the proposed methodology for the estimation of degraded timings of individual gates and well-known procedures for the timing estimation of complete signal paths [55]. Since degraded and non degraded (dashed lines in Figure 4.2) timings are calculated similarly and extensively verified timing estimation procedures are utilized, the industrial relevance of the proposed methodology may be fostered.

As stated in the last section, overall temperature variations and individual stress scenarios of each gate greatly affect the degradation. Hence, mission scenarios and derived stress scenarios that are generated based on tools provided by [57] are included within the gate level analysis flow. In this way, individual signal and transition probabilities that are computed by logic simulations are respected within the degradation calculations. Another consideration that has to be taken into account in this regard is that degradation stress states of each transistor within a gate may depend on the input activities of a single or several input pins (see Section 3.3.1). Thus, realistic threshold voltage shifts are calculated within the proposed gate level degradation analysis flow.

Analysis of the degraded timings may be performed for realistic conditions or worst case assumptions (see Section 4.1). As stated in Section 3.3.3, regression parameters of a gate's timing degradation as function of threshold voltage shift are not incorporated within publications

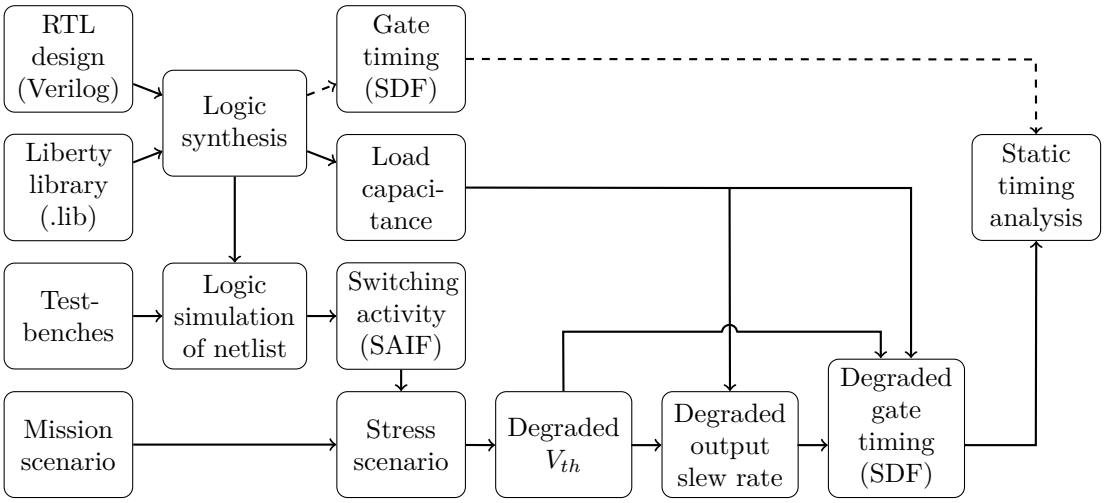


Figure 4.2: Proposed gate level degradation analysis flow. Dashed lines depict the standard timing analysis flow of the non degraded case.

concerning gate level degradation analysis techniques thus requiring SPICE measurements for the computation of degraded output slew rates and gate delays. These SPICE based timing estimations of each gate clearly require knowledge of respective input slew rates and load capacitances. While realistic input slew rates have to be chosen for initial gates within various signal paths, computed degraded output slew rates of preceding gates are directly used for the subsequent gates. Hence, impact of degraded output slew rates on the overall delay of a complete signal path is directly incorporated in the proposed gate level degradation analysis. Load capacitances, which are also a major determining factor of each gate's degraded timing, are estimated by industrial tools thus further fostering the industrial relevance of the methodology. Finally, a SDF file containing each gate's degraded delay is generated to enable timing estimations of complete signal path with industrial tools.

4.3 Summary

A tool flow for a gate level degradation analysis is proposed that relies on industrial tools for the timing estimations of complete signal path by usage of SDF files as main interface [55]. Thus, industrial relevance of the proposed methodology may be fostered as degraded and non degraded timings are calculated similarly and verified timing estimation procedures are utilized. Mission scenarios that contain degradation relevant information about the application of an integrated circuit like ambient temperature variations in the course of several years and predicted off times, are utilized as an additional input argument of the proposed tool flow. The deduction of the devices' varying stress states, which are specified within a defined stress scenario file, permits individual simulations of each device's degraded threshold voltage under realistic conditions. Finally, each device's degraded delay is estimated respecting the effect of output slew rate degradation and inserted within the SDF file to facilitate a static timing analysis of the complete circuit.

In order to efficiently simulate degraded timings based on this proposed gate level degradation analysis flow, high-performance methodologies of threshold voltage simulation and timing estimation are thus required. Hence, Chapter 5 proposes an efficient and varying stress scenario aware method to calculate a transistor's degraded threshold voltage and Chapter 6 elaborates on timing estimations that do not require SPICE simulations within the gate level degradation analysis.

CHAPTER 5

Mission Scenario Aware Analysis of Threshold Voltage Degradation

As stated in Section 4.2, the assessment of the degraded timing of complete RT components shall be based on SDF files that characterize the degraded delays of all gates for specific mission scenarios. First step in the process of generating these SDF files is the calculation of realistic threshold voltage shifts of the transistors within the CMOS logic gates using the transistors' unique stress scenarios as characterized by the corresponding XML files (see Section 4.1). Thus, an efficient method to calculate a transistor's degraded threshold voltage is required within the gate level degradation analysis. Within this chapter, an appropriate modeling technique for this purpose is proposed and an extensive evaluation with state-of-the-art reference models is performed. Besides, this chapter focuses on the threshold voltage degradation effect NBTI, as it is a very important degradation effect (see Figure 1.2) and physical NBTI models are very complex due to regeneration potentials offering prospects for considerable performance improvements.

5.1 Reference Models

5.1.1 Switching Trap Model

The switching trap model shall be used for the evaluation of the new NBTI modeling approach, that is presented within this chapter. Unfortunately, a working model has not been available, as the switching trap model, as implemented by other research groups, is integrated within proprietary transistor computer-aided design software [58]. Thus, the switching trap model has to be specifically implemented for this thesis based on the trap lists that are already described in Section 3.1.2. NBTI simulations of the switching trap model rely on the occupation probability of each trap that has to be tracked individually for the complete stress scenario. This is implemented using each trap's capture and emission times for the respective stress states instead of the fundamental two and four state models. In this way, each physical state of the four state approach (see Figure 3.4) need not be considered and detailed knowledge of the technology's physical properties is not required. The capture and emission times are calculated for each occurring combination of temperature and gate voltage by using activation energy based conversions and interpolations between tabularized values within the trap list. Here, traps normally have significantly different values of capture and emission time, but also very similar capture and emission times can occur for particular combinations of temperature and gate voltage. Since the switching model is required to precisely simulate the expectation value of the threshold voltage degradation, the statistical occupation probability of each trap is being considered during estimation of threshold voltage degradation. In this way, the expectation value is directly simulated instead of statistical nature of the degradation process (see Figure 2.2).

The transformation of the occupation probability is estimated for a very short time step by first calculating the probability increase using a trap's current capture time and Equation 3.1 and then calculating the probability decrease starting from the intermediate result using the trap's current emission time and Equation 3.2 for the same time step. In this way, the simultaneous capture and emission process is divided into a capture process that is followed by an emission process of the same duration. If the time step is sufficiently smaller than the duration of each short-time phase with constant gate voltage and temperature, this approach results in very similar occupation probabilities as the simultaneous capture and emission process. Alternatively, a sequence of emission phases followed by capture phases can of course also be used. Summation of the product of each trap's occupation probability and the trap's constant impact on the threshold voltage, which is also specified within the trap list, finally results in the overall expectation value of the threshold voltage degradation.

Within this thesis two different trap lists are used for the switching trap model. The first list is based on NBTI measurements of a 130 nm technology, which are published by [59]. The second trap list relies on NBTI measurements of a 22 nm FinFET technology and is already used for precise NBTI simulations by [58]. Both trap lists were directly provided by the authors named above. The trap lists consist of 7247 and 12997 individual defects. A time step of 1 ms is used within the switching trap model throughout the evaluation. This time step is considerably smaller than the minimal duration of phases with constant gate voltage in the evaluation examples. In this way, the time step should be small enough to meet the requirement

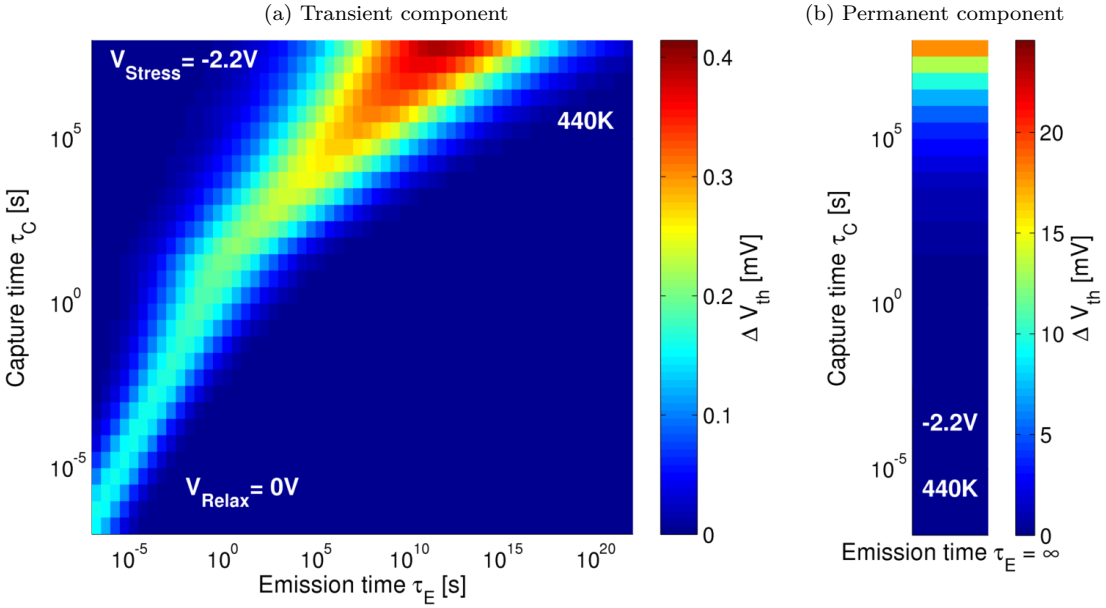


Figure 5.1: Transient and permanent component of the CET map of the 130 nm technology for 440 K and stress voltage -2.2 V. The CET map is directly provided by the author of [38] and corresponds to a scenario of accelerated aging.

of the division into capture process and following emission process.

5.1.2 Capture-Emission Time Model (CET)

The capture-emission time model is also used as a reference model and must therefore also be implemented for this thesis. In contrast to the switching trap model, this implementation requires no additional assumptions and can directly be based on the model description of Section 3.1.3. By using different CET maps within the CET model, threshold voltage degradations can be simulated for different technologies, temperatures or stress voltages. Here, the respective CET maps can directly be measured or can be generated based on trap lists of the switching trap model.

Before the trap lists, that are mentioned above, have been available, a CET map for an accelerated aging condition of the 130 nm technology has been provided by the author of [38]. In particular, the CET map is specified for 440 K and a stress voltage of -2.2 V. As shown in Figure 5.1, the CET map is divided into a transient and a permanent component. Whereas the transient component is conventionally processed within the CET model as described in Section 3.1.3, calculations for the permanent component have to be slightly adapted as the relaxation process has to be omitted. In this way, the permanent threshold shift is modeled similar to Equation 3.1 and 3.3 with the only difference that the capture time histogram $\Delta V_{th\,Perm}(\tau_C)$ and the respective occupation probability are only one-dimensional.

A script that is provided by the author of [58] is used to generate CET maps based on trap lists of the switching trap model. In this way, CET maps for arbitrary temperatures, stress and relaxation voltages can directly be created. Two examples for CET maps that are generated based on the provided trap lists of 22 and 130 nm technology are shown in Figures 5.2 and 5.3, respectively. The chosen CET map resolution of 80 capture time and 64 emission time bins as well as the specified capture and emission time intervals of $\left[10^{-15}, 10^{35}\right]$s and $\left[10^{-15}, 10^{25}\right]$s are used throughout the evaluation. In this way, CET maps have the same time resolution within capture and emission time dimensions. Furthermore, positions of defects within τ_C and τ_E dimensions should not exceed the CET maps' axis intercepts even for extreme scenarios.

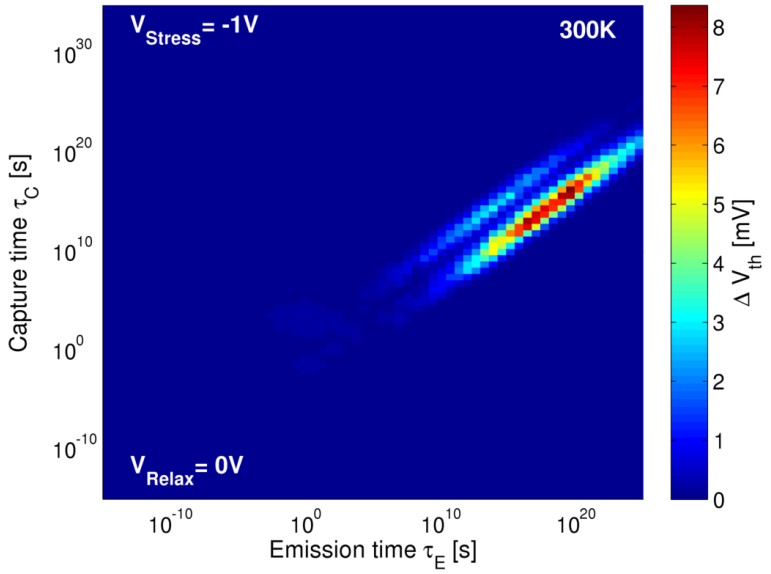


Figure 5.2: CET map of the 22 nm technology for 300 K and stress voltage -1 V with a resolution of 80 capture time and 64 emission time bins. CET data is calculated from a provided trap list of the switching trap model (see Section 5.1.1).

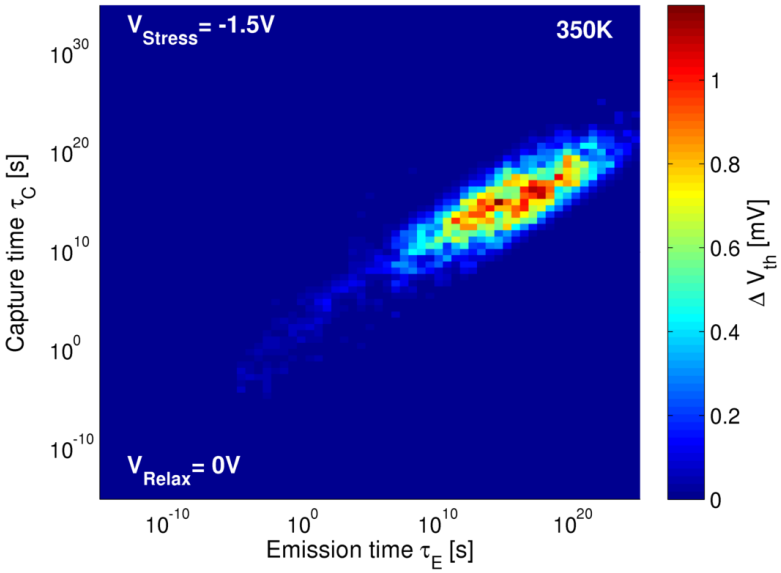


Figure 5.3: CET map of the 130 nm technology for 350 K and stress voltage -1.5 V with a resolution of 80 capture time and 64 emission time bins. CET data is calculated from a provided trap list of the switching trap model (see Section 5.1.1).

5.1.3 Mission Scenario Aware Analytical Approach

The mission scenario aware analytical approach (MSA) as published by [42] shall be used to asses the phase space model by comparing both models' simulation results for various degradation scenarios. Again, the model has to be implemented based on a published model description. As stated in Section 3.1.5, the model relies on logarithmic fits of the high side envelopes of ΔV_{th} curves for different constant stress scenarios and a transition technique between different logarithmic curves for varying scenarios. The impact of specific stress and relaxation phases is handled using the scenarios' overall stress and relaxation times, averaged capture and emission times of numerous defects and a single fitted parameter (see Equation 3.4). On the other hand, the effects of different temperatures and gate voltages are integrated based on a multiplication factor, which depends on temperature, gate voltage, averaged activation energy and two fitted parameters (see Equation 3.5).

Instead of specifying the process to calculate the averaged characteristics of numerous defects and the exact parameter values, only tabularized values for a single technology are denoted in [42]. However, these tabularized values cannot be used, as different technologies are analyzed within this thesis. Therefore, it was decided to not directly implement Equations 3.4 and 3.5 and to develop processes for averaging numerous defects and estimating the parameters, but to implement the main concepts of the mission scenario aware analytical approach in a different way.

The overall stress and relaxation times as well as the parameters of Equations 3.4 and 3.5 are only needed to greatly improve the model's performance. However, the concepts of logarithmic curves and transitions between different curves for varying scenarios can also directly be implemented. In this way, the characteristics of a logarithmic curve are estimated separately for each occurring constant stress scenario with specific values for temperature, gate voltage, frequency or duty cycle. Basic degradation and relaxation curves are simulated by the CET model for each constant scenarios and the parameters of each logarithmic function are calculated by MATLAB's data fitting procedures. In consequence, the logarithmic functions are more precisely tuned for each constant scenario than within the original mission scenario aware analytical approach as published by [42]. Main disadvantage of this approach is the significant deterioration of the MSA model's performance, which is fortunately of no importance within this thesis, as the model shall only be used to assess the phase space model's accuracy. Finally, the transitions between different logarithmic degradation and degradation curves within varying scenarios are directly implemented as stated in Section 3.1.5. The MSA model's overall accuracy should thus be improved in comparison to the published version of the modeling technique at the cost of a significant performance deterioration.

The only remaining degree of freedom within the model implementation is the degradation time that is used to characterize the logarithmic curve based on a CET simulation for each constant scenario. This variable differs from the actual degradation time that is afterwards used for the NBTI simulation of a varying mission scenario. The MSA model's overall accuracy is affected by this parameter, as very large or very short degradation times within the characterization induce deviations between the logarithmic curve and the actual ΔV_{th} curve in the short or long time region, respectively. These deviations may cause significant overall inaccuracies for corresponding varying scenarios, since only the short or long time region of a constant scenario may be required for the overall NBTI degradation simulation. Furthermore, the degradation state at maximum characterization degradation time is used as initial state to specify the logarithmic relaxation curve for each constant scenario. As stated in Section 3.1.5, this relaxation curve is scaled to the particular ΔV_{th} values within the simulation of varying scenarios. Due to activation of different kinds of defects, very large or very short characterization degradation times thus induce shallow or steep relaxation profiles, respectively. As, to the best of my knowledge, this set of problems is not discussed in the literature, an additional procedure has been developed to select characterization degradation times for different varying scenarios. In particular, degradation times between 1 hour and 1 day are used in the way that the final ΔV_{th} in each constant characterization scenario is neither significantly smaller or larger than the maximal ΔV_{th} of the corresponding varying scenario. This alignment of varying scenario and characterization degradation time should also slightly improve the MSA model's accuracy.

5.2 Phase Space Model of Single Transistor NBTI Degradation

Within this section a new NBTI modeling technique is presented. It is specifically designed for the use case of the mission scenario aware gate level degradation analysis as described in Chapter 4. Basic objective for the development of this NBTI modeling approach is the support of varying stress scenarios as described in Section 4.1. In this way, realistic degradation simulations without worst case assumptions can be introduced. Second and most important objective is a considerable improvement in simulation performance, since degradation simulations of a high number of transistors with different stress scenarios have to be performed within the gate level degradation analysis. Therefore, simulation performance of the NBTI modeling technique is crucial to achieve a reasonable duration of computation for the gate level analysis.

Basic concept of the performance improvements relies on the specific stress scenarios within gate level analysis. Instead of arbitrary analogue stress scenarios, the NBTI modeling technique has to support only square wave signals, which can be described with frequency, duty cycle and stress voltage (see Figure 3.12). This square wave assumption is a common simplification in gate level degradation simulations and is widely used in other publications [13, 39, 16]. However, there is a small impact of the square wave assumption on the degradation simulation's accuracy due to slightly longer stress times, since the signal changes instantly without ramps. In an additional study, which is not part of this thesis, it is shown that the square wave assumption and the effect of ignoring the drain bias cause up to 10% overestimation of the NBTI effect [60]. The effect of ignoring the drain bias is another second-order effect caused by usage of NBTI characterization data, that is measured without drain bias, within simulations of real systems where the drain bias is not always negligible.

After utilizing the relationship between actual voltage signal and the corresponding NBTI stress condition (see Section 3.3.1), frequency f and duty cycle dc of the corresponding square wave NBTI stress signal can be calculated based on clock cycle time Δt and up- P_{01} and down-switching P_{10} probabilities of the NBTI stress.

$$\begin{aligned} f &= \frac{2\mathrm{P}_{01}\mathrm{P}_{10}}{\mathrm{P}_{01}+\mathrm{P}_{10}}\frac{1}{\Delta t} \\ dc &= \mathrm{P}_{01}/(\mathrm{P}_{01}+\mathrm{P}_{10}) \end{aligned} \tag{5.1}$$

Alternatively, definition of the NBTI stress signal's frequency and duty cycle can be based on signal probability and transition probability that are calculated by standard logic simulation tools (see Figure 2.5). While the stated signal probability SP can be translated to duty cycle dc using the simple formula $dc = 1 - SP$ (see Section 3.3.1), the transition probability has to just be multiplied with $\frac{1}{\Delta t}$ in order to estimate the frequency f. Furthermore, for some kinds of gates within the gate level degradation analysis (e.g. NOR), signal probabilities of all input pins have to be considered and the probability of NBTI stress conditions is calculated using Equation 3.6.

The stress voltage of the square wave signal is also fairly restricted within the use case of mission scenario aware gate level degradation analysis. Due to dynamic voltage and frequency scaling (DVFS) only a few different voltage values need to be considered. Furthermore, transistors placed at different locations within the circuit may be subjected to slightly reduced voltages by reason of IR drop effect [48]. However, the IR drop effect does not increase the number of possible voltage values for a single transistor, but causes a slight reduction of all possible voltages. Therefore, the NBTI modeling technique has to support only a small number possible voltage values during the degradation simulation of a single transistor. In order to fully support the IR drop effect, different transistors can be classified by IR drop strength (see Section 5.2.7) or the IR drop effect can even be neglected due to a fine-grained power grid [48].

Based on these specific stress scenarios of the desired use case, a performance oriented yet mission scenario aware NBTI model is introduced. The basic concept of the modeling approach is described in Section 5.2.1. Different implementations of this concept and improvements eliminating minor shortcomings of the modeling approaches are presented in the following subsections. Each state and improvement of the model is evaluated using the respective reference model described in Section 5.1.1 and 5.1.2. Furthermore, simulation results of the improved

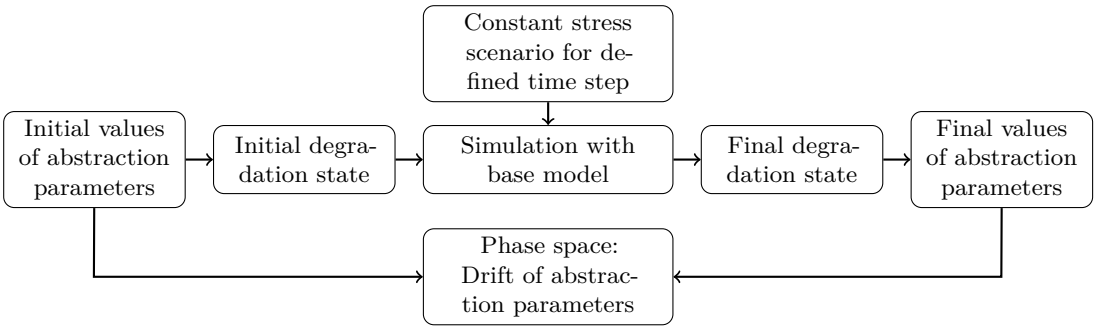


Figure 5.4: Flowchart of the calculation of the phase space (LUT). A large number of combinations of initial parameter values as well as different constant stress scenarios are used during phase space calculation.

model are compared to the implementation of the modeling approach of [42] as described in Section 5.1.3.

5.2.1 Basic Phase Space Concept

Main idea of the new performance oriented NBTI modeling technique is a lookup table (LUT) approach. For different stress scenarios, a precise physical NBTI model is used to calculate the transformations of various initial degradation states within a small time step. If initial and resulting degradation state can sufficiently be described with only a few parameters, these transformation can be stored within a LUT. The huge number of NBTI simulations for different stress scenarios within the gate level degradation analysis can then all be based on the same LUT, resulting in a huge speedup of a single simulation. However, this huge performance improvement comes at the cost of pre-calculating the LUT.

If all abstraction parameters are restricted to certain ranges, the LUT incorporates all possible degradation states of a transistor and the LUT can be used like a phase space in physics. In this way, a LUT based NBTI simulation can't leave the area of precomputed values. Therefore, the new NBTI modeling technique and LUT are hereafter referred to as phase space model and phase space, respectively.

The principle flow to calculate the phase space is presented in Figure 5.4. Pre-selected initial values of the abstraction parameters are used to generate initial degradation states of a physical NBTI model. Based on these initial degradation states, physical NBTI simulations are performed for small time steps with constant stress scenarios. Here, different constant stress scenarios need to be considered in order to support all the varying stress scenarios of the gate level degradation analysis (see Section 4.1). Finally, the resulting degradation states of the physical NBTI simulations are characterized by the abstraction parameters. The drifts between initial and final abstraction parameter values are then stored within the phase space for the different constant stress scenarios. The dimension of the phase space is given by the number of parameters for the abstraction plus the number of parameters that are required to characterize the constant stress scenario. An example of a calculated phase space with reduced number of dimensions for improved visualization is shown in Figure 5.5. The phase space construction can be processed in parallel and has to be done only once for each transistor technology and geometry. Construction duration is defined by the type of the physical NBTI model, the number of abstraction parameters, the step size of parameter values and supported stress scenarios.

It is not sufficient to abstract the degradation state of a transistor using a single parameter, since the transistor's degradation history must also be considered. This degradation history determines the transistor's ability to regenerate or to further degrade. In Figure 5.6 two different transistors are presented which have a totally different degradation history but share the same value of shift in threshold voltage at a single point in time. A NBTI model that is based on shift in threshold voltage as single abstraction parameter could not differentiate between these two transistors and would predict the same degradation behavior of these devices in a subsequent time step. However as shown by the dashed lines in Figure 5.6, these transistors will surely

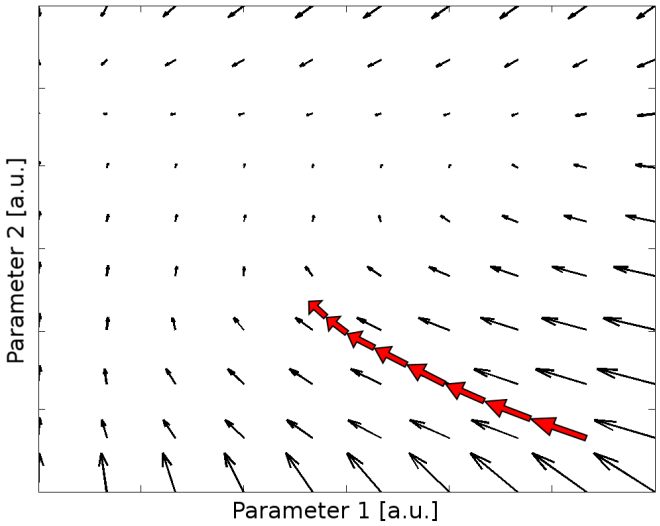


Figure 5.5: Quiver plot of an exemplified phase space with reduced number of dimensions for improved visualization. A possible trajectory during a phase space simulation is shown with red arrows.

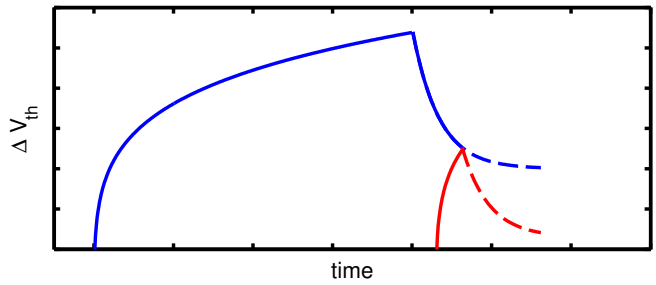


Figure 5.6: Shift in threshold voltage of two different transistors (blue and red) as function of time. Although both transistors share the same ΔV_{th} at a single point in time, their degradation curves will be different in the future (dashed lines) due to different degradation histories.

have differing degradation characteristics within this time step. Since the mission scenario aware analytical approach (see Section 3.1.5) uses shift in threshold voltage as single abstraction parameter of a device's degradation state, the described behavior is a major drawback of this modeling technique.

In order to take the degradation history of a transistor into account, the phase space approach uses more than one abstraction parameter to characterize the degradation state. At the same time, the number of abstraction parameters greatly affects size and computation time of the phase space, since each abstraction parameter represents a dimension within the phase space. Therefore, degradation state of a transistor should be abstracted using as few parameters as possible and the different parameters should be independent of one another. The limitation of possible stress scenarios within the use case of gate level degradation analysis can be utilized to increase the precision of the abstraction. In this way, the abstraction method can be constrained to only support digital scenarios and doesn't need to maintain arbitrary stress scenarios like analogue stress signals. Due to this confinement, a precise abstraction of degradation state, which takes the degradation history into account and uses only a few abstraction parameters, may be facilitated.

As indicated in Figure 5.4, the phase space approach can be implemented based on different physical NBTI models. Of course, each physical NBTI model has a unique way to characterize the degradation state resulting in different abstraction approaches and abstraction parameters for the different physical models. In addition, the physical model's unique way to characterize the degradation state may be more or less suited for an abstraction that relies on a few parameters.

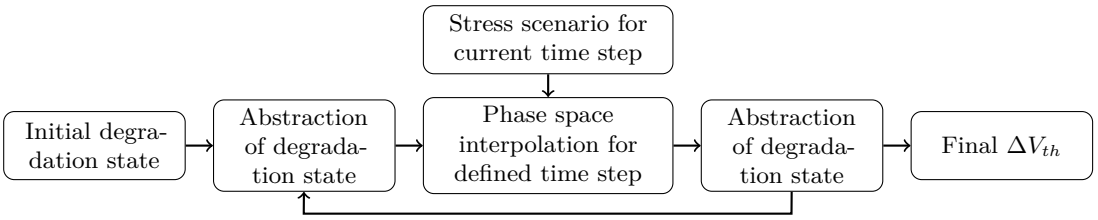


Figure 5.7: Basic NBTI simulation flowchart of the phase space model. A single phase space interpolation covers the degradation due to a constant stress scenario of defined time step. Repetitive interpolations with the same or different constant stress scenarios are performed until overall simulation time is attained.

Instead of a NBTI simulation that is based on the precise transformation of a physical model's complex degradation state, the phase space simulation tracks the transformation of a small number of abstraction parameters by performing repetitive interpolations within the pre-calculated phase space. A flowchart of this simulation concept is shown in Figure 5.7. Here, a single phase space interpolation covers the degradation due to a constant stress scenario of defined time step. Resulting parameter values of a preceding phase space interpolation are used in conjunction with parameters of the particular constant stress scenario to interpolate the degradation in the subsequent time step. As defined by the overall stress scenario of the concerned transistor, each interpolation may be based on a different constant stress scenario and the consecutive phase space interpolations are repeated until the overall simulation time is attained. In this way, the NBTI simulations results in a trajectory within the quiver plot representation of the phase space (see Figure 5.5).

The defined time step has a major effect on the accuracy of the phase space approach, since every time step on the phase space trajectory will induce a small interpolation failure, which depends on the phase space's parameter step size, and a failure due to the abstraction approach. Therefore, the phase space approach facilitates the usage of longer simulation time steps (e.g. 10min). However, the maximum time step is defined by the time resolution of the mission scenario, which only incorporates changes on similar macroscopic time steps (see Section 4.1).

The proposed phase space NBTI simulation should have a vastly increased performance in comparison to physical NBTI models. The improvement is based to a small extent on the interpolation being faster than a direct calculation of a physical model and to a great extent on the usage of a macroscopic time step. In this way, the number of computational steps of a long time degradation simulation is vastly reduced in comparison to the physical NBTI models.

In conclusion, the phase space NBTI model utilizes digital stress scenarios for a simple abstraction of the degradation state, should have a vastly increased performance due to usage of macroscopic time steps and supports varying stress scenarios on the basis of this macroscopic time step. Consequently, the phase space NBTI model is perfectly adapted for the use case of the gate level degradation analysis.

5.2.2 Reaction-Diffusion Based Model

The first implementation of the phase space concept is based on the reaction-diffusion model as described in Section 3.1.1. Since I had no access to reaction-diffusion parameters of a current transistor technology, model parameters are taken from [31] and correspond to accelerated transistor aging. The reaction-diffusion based phase space model is published in detail [61] and is also incorporated within a publication of a full chip design methodology at system level [62].

Two parameters are used for the abstraction of the NBTI degradation state. The first parameter is the transient component of the shift in threshold voltage. This is equivalent to the number of hydrogen molecules within the gate oxide in the case of the reaction-diffusion model. The second parameter should characterize the ability of the system to regenerate. It is therefore called "healability" within this section and is linked to the probability that the transient component of the shift in threshold voltage becomes permanent. The healability definition is based upon a solution of the system of differential equations (Equations 5.2 - 5.5)

that characterize the reaction-diffusion model. These equations are presented in [31] and are slightly adapted in [63] to comprise the conversion to molecular hydrogen.

$$\frac{dN_{IT}}{dt} = k_F(N_0 - N_{IT}) - 2k_R N_H N_{IT} \qquad x = 0 \tag{5.2}$$

$$\frac{dN_{IT}}{dt} = 2D_H \frac{dN_H}{dx} + \delta \frac{dN_H}{dt} \qquad 0 < x < \delta \tag{5.3}$$

$$D_H \frac{d^2 N_H}{dx^2} = \frac{dN_H}{dt} \qquad \delta < x < T_{ox} \tag{5.4}$$

$$D_H \frac{dN_H}{dx} = -k_P N_H \qquad T_{ox} < x \tag{5.5}$$

where N_{IT} is the number of interface traps, N_0 is the initial number of unbroken Si-H bonds, N_H is the hydrogen concentration, k_F is the forward dissociation rate constant, k_R is the annealing rate constant, k_P is the surface recombination velocity, D_H is the hydrogen diffusion coefficient, $x = 0$ denotes the Si-SiO_2 interface, δ is the interface thickness and T_{ox} is the oxide thickness.

The following approach is used to solve the system of differential equations.

$$N_H(x,t) = \begin{cases} -m(t)x + b(t) + \kappa(x,t) \\ A(t)e^{-\beta(t)x} \end{cases} \tag{5.6}$$

$\kappa(x,t)$ stands for the deviation between linear solution and current degradation state. Inserting the approach in Equation 5.5 gives

$$\beta(t) = \frac{k_P}{D_H} \tag{5.7}$$

Continuity in T_{ox} yields

$$A(t) = (-m(t)T_{ox} + b(t))\, e^{\frac{k_P}{D_H} T_{ox}} \tag{5.8}$$

Continuity of the derivative in T_{ox} yields

$$b(t) = m(t)\left(\frac{D_H}{k_P} + T_{ox}\right) \tag{5.9}$$

Integration of $N_{IT}(t) = \int_0^\infty 2N_H(x,t)dx$ delivers

$$m(t) = \frac{N_{IT}(t) - 2\int_0^\infty \kappa(x,t)dx}{\left(T_{ox}^2 + 2\frac{D_H T_{ox}}{k_P} + 2\frac{D_H^2}{k_P^2}\right)} \tag{5.10}$$

The linear solution $\kappa(x,t) = 0$ of the concentration of hydrogen molecules $N_H(x,t)$ is therefore thoroughly defined by the number of interface traps N_{IT}. In this way, there is a certain linear solution for every shift in threshold voltage. This solution is shown as a blue line in Figure 5.8 as an example of possible hydrogen concentration for a given value of threshold voltage shift.

The calculation of the healability value for a given hydrogen concentration starts with the linear solution of the corresponding threshold voltage shift. Next step is a weighted difference between hydrogen concentration and linear solution. The weight function is a linear function with the value 1 at Si-SiO_2 interface ($x = 0$) and the value -1 at the end of the oxide ($x = T_{ox}$). Thus, in the first half of the oxide a hydrogen concentration that is higher than the linear solution produces positive values of the weighted difference in this region. Likewise, a hydrogen concentration that is lower than the linear solution produces positive values in the second half of the oxide. This is illustrated by the arrows in Figure 5.8. As shown in Equation 5.11, the healability is then defined as the spatial integral of the weighted difference divided by the spatial integral of the hydrogen concentration.

$$\text{Healability} = \frac{\int_0^{T_{ox}} (N_H - LS) \cdot WF \; dx}{\int_0^{T_{ox}} N_H \; dx} \tag{5.11}$$

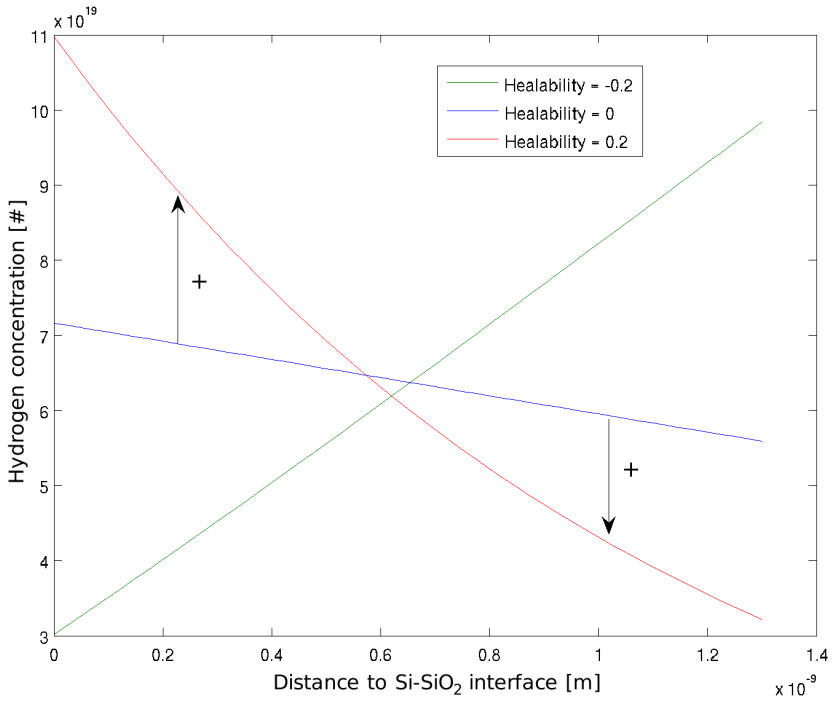


Figure 5.8: Three examples of concentration of hydrogen molecules within the gate oxide. The examples have the same shift in threshold voltage (integral) but different healability values. The blue curve (Healability = 0) is the linear solution of the system of differential equations. Difference between hydrogen concentration and linear solution defines the healability value.

where LS represents the linear solution of the system of differential equations (see Equations 5.6, 5.9 and 5.10) and WF is the weight function. In this way, healabilty is restricted to the region $[-1, 1]$. Examples of hydrogen concentrations with various healability values are given in Figure 5.8. A transistor's regeneration ability that is higher or lower than the regeneration ability of the linear solution is thus indicated by positive or negative healability values, respectively.

In order to calculate the phase space it is needed to generate a hydrogen concentration based upon a given shift in threshold voltage and healability. This is done starting with the linear solution of the shift in threshold voltage. Afterwards, the hydrogen concentration is changed repeatedly in minor steps in the first and second half of the oxide as long as the desired healability is reached. At each position, the change of the hydrogen concentration is a product of the appropriate hydrogen concentration value of the previous iteration and a predefined multiplier. Thereby, the spatial integral of the hydrogen concentration is kept constant.

The permanent component of the shift in threshold voltage is treated specially. During a simulation it can only increase over time. As long as the permanent shift in threshold voltage is small in comparison to the maximum possible shift in threshold voltage, the calculations of the reaction-diffusion model are barely affected by a permanent voltage shift. Therefore, it was decided to not use the permanent shift in threshold voltage as a third abstraction parameter. Instead, the permanent voltage shift is calculated for every phase space entry. During a phase space simulation the permanent shift in threshold voltage that occurs in each time step is then summed up. This is a small enhancement to the basic phase space simulation flow as presented in Figure 5.7.

The phase space is calculated for a wide range of initial values of transient shift in threshold voltage and healability as well as different constant stress scenarios on the basis of a time step of one minute. The different constant stress scenarios facilitate varying duty cycles, temperatures and gate voltages in a long time NBTI simulation. Here, the duty cycle parameter of the phase space can be used to simulate the impact of different signal probabilities as well as power gating and off-times on NBTI degradation. The reaction-diffusion simulation results are independent of duty frequency in a wide frequency range [20]. Therefore, duty frequency is irrelevant as long

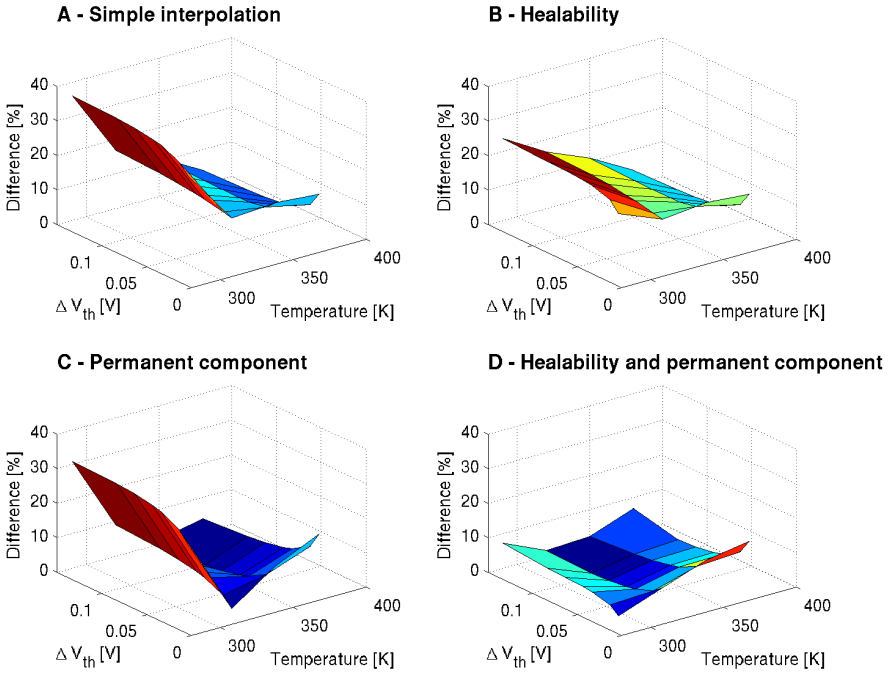


Figure 5.9: Percental differences between direct reaction-diffusion calculation and phase space simulation of one hour of permanent stress. Panel A shows the differences for a simulation without healability parameter and summation of permanent component of the shift in threshold voltage. Simulation methods used in Panel B and D make use of the healability parameter while the summation of the permanent component of the shift in threshold voltage is used in Panel C and D.

as it is within this range and it is not used as a phase space parameter.

5.2.2.1 Evaluation

Degradation curves of a transistor during one hour of permanent stress are simulated using the phase space model with a time step of one minute. Hence, the simulations are composed of 60 interpolations within the phase space. In order to evaluate the phase space method, direct reaction-diffusion simulations are also performed for all conditions. Simulations with the phase space model were more than 600 times faster than direct reaction-diffusion simulations. In this way, simulations that would need about a month with the reaction-diffusion model can be calculated in less than 75 minutes. The short time step of the phase space was chosen, since the parameters of the reaction-diffusion model correspond to accelerated transistor aging. With access to reaction-diffusion parameters based upon a current technology, a phase space with a greater time step could have been used. In this case, the benefit in simulation speed would have been even greater.

The simulations are performed for a wide range of initial shift in threshold voltage and temperature. Percental differences between the direct simulations and the phase space simulations are calculated for all those conditions. These results are presented in Figure 5.9. The different panels of the figure illustrate the benefit of the different features of the reaction-diffusion implementation of the phase space approach. In Panel A the phase space simulation is performed without using the healability abstraction parameter and the summation of the permanent component. A difference up to 40% occurs in this situation. This difference is reduced when the healability parameter (Panel B) or the summation of the permanent component of the shift in threshold voltage (Panel C) is used during the simulation. As expected, best phase space simulations are achieved when both healability and summation of the permanent component are used (Panel D). In this case, the percental difference between direct calculation and phase space simulation is always smaller than 10%.

5.2.2.2 Discussion

The evaluation results indicate that the abstraction parameter "healability" is well defined with respect to the requirements. Furthermore, healability and summation of the permanent shift in threshold voltage are crucial for the success of the phase space model. The model is well suitable in performance oriented use cases, since a small deterioration of the simulation results comes with a vastly improved simulation speed. The additional phase space dimensions of temperature, supply voltage and duty cycle permit a performance efficient way to simulate the dependence on various stress scenarios. In this way, NBTI degradation can be calculated efficiently without disregarding power gating, temperature profiles and the IR drop.

During development of the reaction-diffusion based phase space model, a paradigm shift in understanding NBTI took place in the community of physical NBTI modeling [22]. It became evident that the reaction-diffusion model isn't able to correctly simulate several degradation measurements [19] and the switching trap model has been proposed [4]. Therefore, the phase space model has been revised in order to be based on the new switching trap approach. As described in Section 5.2.1, the main concept of the phase space model can be transfered to rely on a different physical model by using an abstraction approach that is adapted to the new model.

5.2.3 Switching Trap Based Model

Efficient abstraction of the switching trap model's representation of the degradation state is the main task during development of a switching trap based phase space model. Within the fundamental switching trap model the degradation state is characterized by strongly diverse occupation probabilities of several thousand different traps. An apparent abstraction that can reduce the number of required parameters for the degradation state characterization is grouping of several traps based on the traps' timing characteristics (see Section 3.1.2). Unfortunately, it is not possible to establish this grouping, since each trap's timing characteristic is dependent on temperature and gate voltage in a strongly individual way. Therefore, grouping of traps would introduce a severe inaccuracy due to averaged dependencies on temperature and gate voltage for the complete group of traps.

Instead of a phase space model that directly relies on the switching trap model, the phase space approach can also be based on the capture-emission time (CET) model. As described in Section 3.1.3, the CET model is already a first abstraction of the fundamental switching trap model and introduces a form of grouping of different traps. Hence, the CET model is a perfect intermediate step in order to abstract the switching trap model's representation of the degradation state. However, some additional considerations need to be made in order to support varying temperatures and gate voltages within the CET concept (see Section 5.2.5 and 5.2.7). The number of parameters that characterize the degradation state has to be further reduced for the usage within the phase space model. An apparent way is a considerable reduction of the CET map's resolution. As published in [64], this technique causes a huge deterioration of simulation results and is thus not part of the phase space model. A different approach that further reduces the number of parameters, which is part of a student's graduation work, characterizes the activated region within the CET map with three parameters [65]. Since the approaches' accuracy is not as good as expected, it was decided to not continue with this abstraction method. Finally, a third abstraction method is developed that relies on the occupation probability $P(\tau_C, \tau_E)$ (see Equation 3.3) and utilizes the specific stress scenarios of the gate level degradation analysis [64]. Different aspects of this method are described in the following paragraphs.

NBTI simulations with the CET model calculate the occupation probability $P(\tau_C, \tau_E)$ for each pixel of the CET map and the shift in threshold voltage is estimated by integrating the multiplication of occupation probability and CET map (see Equation 3.3). Hence, impact of duty cycle and frequency of a square wave signal are only incorporated within the occupation probability. Typical occupation probabilities for fundamentally different stress scenarios are shown in Figure 5.10. As can be seen, occupied and unoccupied traps are always separated by a monotone function, since trap charging and discharging always starts at low values of τ_C and τ_E, respectively. In this way, constant stress triggers an occupation that is uniformly distributed in the τ_E dimension and varying stress scenarios produce occupation profiles where the highest

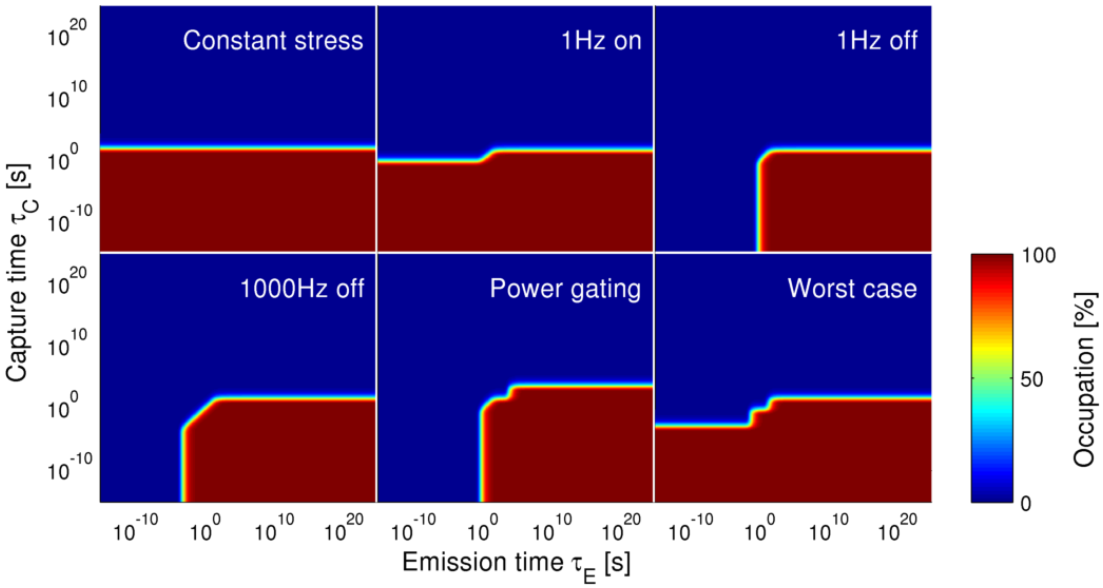


Figure 5.10: Occupation probabilities $P(\tau_C, \tau_E)$ for different stress scenarios

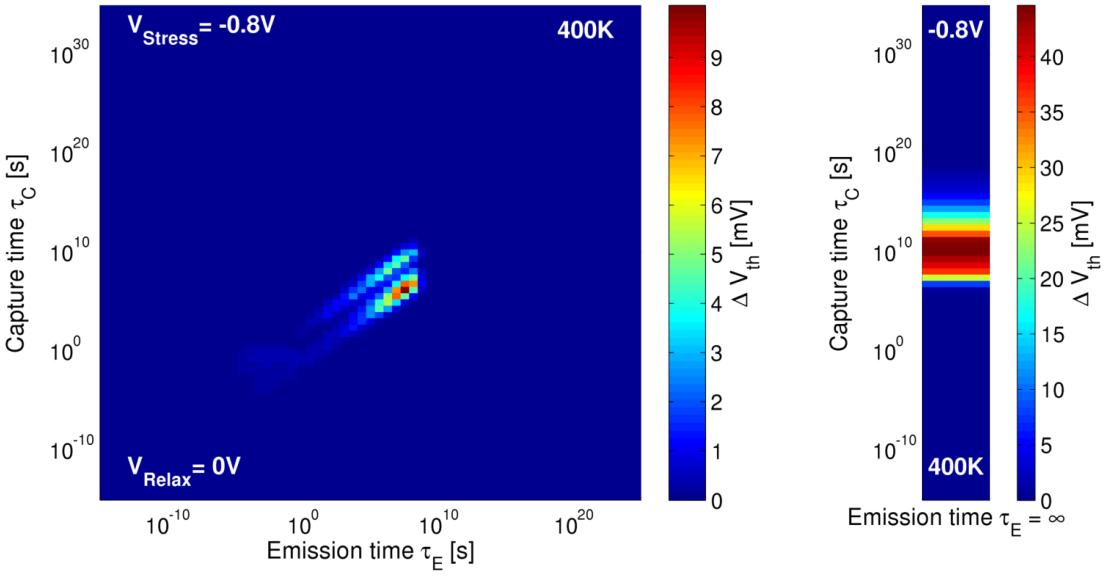
occupation in the τ_C dimension is always located in last τ_E column. Scenarios finishing with a stress state (see "1 Hz on" or "Worst case" in Figure 5.10) have non zero occupation probabilities for very small τ_E values. On the other hand, occupations of scenarios that finish with an off state are typically almost rectangularly shaped. Furthermore, a square wave signal produces a region with increasing occupations at higher capture time for increasing emission time. In the double logarithmic representation this slope has always of gradient of 45° and the width of the slope region depends on the signal's frequency and duty cycle (see "1 Hz off" and "1000 Hz off"). This correlation between properties of a square wave signal and corresponding occupation probabilities is also described in [66]. Finally, long regeneration times within the power gating scenario or the worst case scenario that consists of a very generic minutes stress - minutes relax - seconds stress - seconds relax - milliseconds stress sequence, produce steps within the occupation probability.

Similar to the CET map for accelerated aging of the 130nm technology that was directly provided by the author of [38] (see Figure 5.1), a separation of CET maps into transient and permanent components is utilized within the CET based phase space model. In order to perform this separation for the CET maps that rely on trap lists of the switching trap model, the script of the author of [58] is adapted. All traps having emission times larger than 10^8 s in a scenario with worst case temperature are considered to be part of the permanent component. As an example, the separation in transient and permanent component is shown for a worst case temperature of 400 K in Panel a and b of Figure 5.11. This worst case temperature is used throughout the thesis for the separation of transient and permanent component. The separation criterion is only breached if there are continuous regeneration times larger than 10^7 s, which roughly equals 4 months, at worst case temperature within the stress scenario of a transistor. Since emission times of all traps are considerably larger at lower temperatures, the separation assumption also supports much longer continuous regeneration times, if the regeneration of a transistor occurs at a lower temperature. This is shown by the occurrence of considerably larger emission times at 300 K in Panel c of Figure 5.11.

Main goals for the abstraction with only a few parameters are to accurately abstract occupations within long capture and emission time regions for long time degradation scenarios and to support the impact of square wave frequency and duty cycle. Three different abstraction parameters that are independent from each other have been selected for the CET based phase space model. They are characterized in Figure 5.12 and rely on the possible occupation probabilities of Figure 5.10. First and most important abstraction parameter is the transient shift in threshold voltage $\Delta V_{th\ Trans}$. In order to calculate this parameter, the current trap occupation (e.g. Figure 5.12) has to be multiplied with the relevant CET map (e.g. Figure 5.11a)

(a) Transient component at 400 K

(b) Permanent component at 400K



(c) Transient component at 300 K

(d) Permanent component at 300K

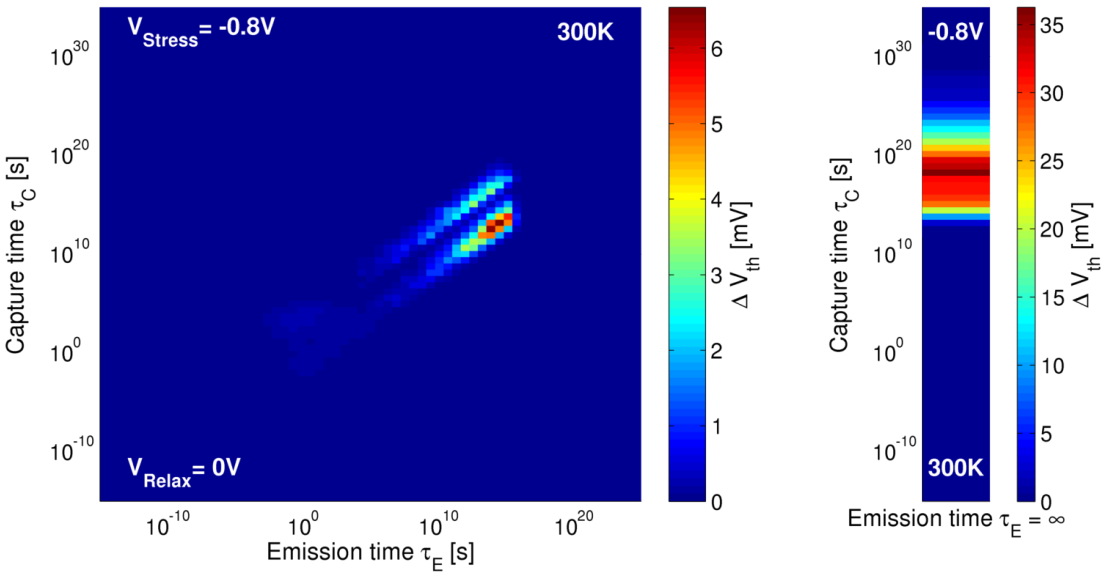


Figure 5.11: Transient and permanent component of the CET maps of the 22 nm technology for 300 and 400 K. CET data is calculated from a provided trap list of the switching trap model (see Section 5.1.1).

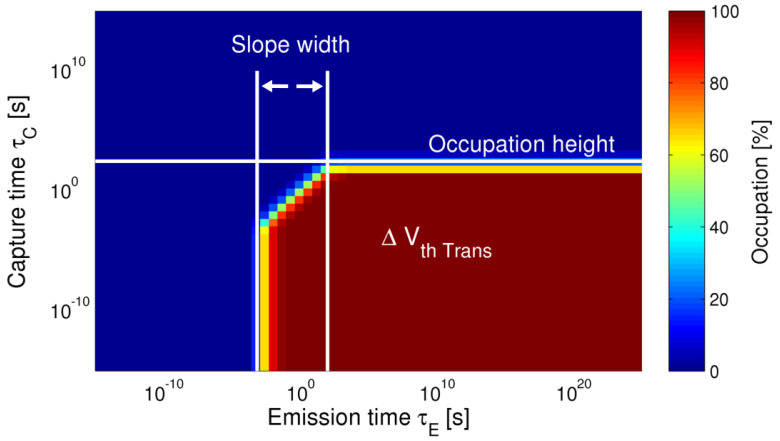


Figure 5.12: Abstraction parameters of the CET based phase space model

and a summation on both τ_C and τ_E dimension has to be performed (see Equation 3.3). Due to this parameter, it is impossible that an immediate shift in simulated threshold voltage is caused solely by the abstraction of a CET degradation state. Furthermore, this parameter is also the primary NBTI model output to be used in an adjacent delay degradation estimation.

The next abstraction parameter is supposed to encompass the occupation height in τ_C dimension for the last τ_E column. Although there might not be a single defect within this column, this occupation height characterizes the overall time that the transistor is exposed to stress since last full recovery of the transient shift in threshold voltage. If highest emission time within the transient component of the CET map is considerably larger than the maximum regeneration time of any reasonable mission scenario, occupation in the last τ_E column cannot regenerate and the corresponding occupation height is equivalent to the permanent shift in threshold voltage $\Delta V_{th\ Perm}$. In this way, $\Delta V_{th\ Perm}$ is used as abstraction parameter of the transient component's occupation height and a summation of the permanent component during a NBTI simulation, which is part of the reaction-diffusion based phase space model (see Section 5.2.2), is not required.

The third and last abstraction parameter is used to characterize the slope within the occupation due to periodic stress by specifying the logarithmic temporal width of the slope region. A precise definition of the slope region is given by the interval specification of Equation 5.12.

$$\frac{\int d\tau_C P\left(\tau_C, \tau_E\right)}{\int d\tau_C P\left(\tau_C, \tau_{E\ Max}\right)} \in [0.01, 0.99] \tag{5.12}$$

First the sum of the occupation is calculated within the τ_C dimension for each emission time value and the resulting values are normalized with the occupation summation for the maximum emission time. The slope region is defined as the emission time range where the result is greater than 0.01 (non zero) and smaller than 0.99 (smaller than the occupation of maximum emission time). This region is also marked within Figure 5.12. If there is still a distinct occupation within the short time region, the slope region extends to minimal emission time. As specified by Equation 5.13, the abstraction parameter slope width SW is then defined as logarithm of the difference between the emission time values of upper and lower boundary of the slope region.

$$SW = log\left(\tau_E(\text{upper boundary}) - \tau_E(\text{lower boundary})\right) \tag{5.13}$$

As shown in Figure 5.10, the slope region expands with increasing frequency of the square wave stress signal. In this way, the slope width parameter becomes more important for high frequency scenarios. However, the importance of the slope width is still limited within these scenarios as the parameter also abstracts the occupation within very short emission time regions where the trap density of the CET map is minimal (see Figure 5.11a). A clear deviation between occupation and abstraction within this region has only a small impact on the phase space model's accuracy.

A conversion between the original occupation of the CET model and the abstract representation is mandatory for the phase space approach. The calculation of the abstraction parameters based on the trap occupation is a straight forward process using Equations 3.3 and 5.13. However, the generation of a trap occupation based on the abstraction parameters is a complex operation with various degrees of freedom. The following procedure is used within the thesis to perform this conversion. Main goal of the determination of the procedure is to generate precise occupation probabilities in the region of long capture and emission times, as inaccuracies in the region of short capture and emission times may dissolve due to short relaxation times within the transistor's stress scenario. First a virtual capture time is increased and the appropriate occupations are calculated as long as $\Delta V_{th\,Perm}$ of the resulting occupation matches the desired value of $\Delta V_{th\,Perm}$. This technique determines the generation of the occupation within the permanent component and the occupation height within the transient component (see Figure 5.12). In this phase, the generated occupation of the transient component resembles the constant stress scenario of Figure 5.10. A slope region based on the desired slope width is produced in the next step of the procedure. Besides having the correct slope width, the slope region of the generated occupation is always located at minimal emission times in this phase. Finally, the value of $\Delta V_{th\,Trans}$ is used in order to shift the slope region to the desired emission time region. This is done by increasing a virtual emission time in combination with repetitive restoration of the slope width as long as $\Delta V_{th\,Trans}$ of the resulting occupation matches the desired value of $\Delta V_{th\,Trans}$.

An additional procedure is used when this algorithm cannot generate an occupation that fulfills all abstraction parameters. This may occur when the occupation as specified by just $\Delta V_{th\,Perm}$ and slope width results in a value of $\Delta V_{th\,Trans}$ that is smaller than the desired value. Due to the technique of increasing a virtual emission time, the original algorithm can only handle $\Delta V_{th\,Trans}$ values larger than the desired value as intermediary result. Therefore, the desired slope width is reduced as long as the intermediary $\Delta V_{th\,Trans}$ is larger than the desired value in this case.

Since $\Delta V_{th\ Trans}$, $\Delta V_{th\ Perm}$ and slope width are restricted to certain ranges, the phase space incorporates all degradation states and a trajectory (see Figure 5.5) can't leave the area of precomputed values. The presented abstraction based on the CET model directly facilitates degradation simulations of stress scenarios with varying signal probabilities as induced by power gating or different system states. Since a single CET map is associated with precise values of temperature and gate voltage, varying temperatures and gate voltages cannot be directly supported. In Section 5.2.5 and 5.2.7 possible improvements of the CET approach are examined to overcome these restrictions.

5.2.3.1 Evaluation

The CET occupation examples of Figure 5.10 are abstracted using the parameters $\Delta V_{th\ Trans}$, $\Delta V_{th\ Perm}$ and slope width as described in the last section. Afterwards, the occupations are recreated exclusively based on the abstraction parameters. These generated occupations are shown in Figure 5.13. The recreation of the scenarios "Constant stress", "1 Hz off" and "1000 Hz off" is almost perfect. Here, the abstraction parameter slope width is able to characterize the increased slope range due to a higher signal frequency. Scenarios with periodic stress and no relaxation period at the end of the stress signal ("1 Hz on" and "Worst case") result in an occupied short time region and a non homogeneous occupation in τ_E dimension. Generated occupations of these scenarios differ significantly from the original occupation profiles. The slope region does not extend to the minimal emission time and the parameter slope width is considerably smaller. These two scenarios are the only scenarios with slightly deviating abstraction parameters between original and recreated occupations. Furthermore, the abstraction does not characterize the occupation steps of the scenarios "Power gating" and "Worst case". In this ways, the recreated occupation contains a broad slope region instead of these occupation steps.

Impact of the separation criterion for transient and permanent CET component on the component's overall ΔV_{th} is shown in Table 5.1. Although the permanent component only contains traps with very large emission times, this component always represents more than 60% of total possible NBTI degradation. Of course, this partitioning only occurs after infinite degradation time. Since traps with very large emission times tend to also have long capture

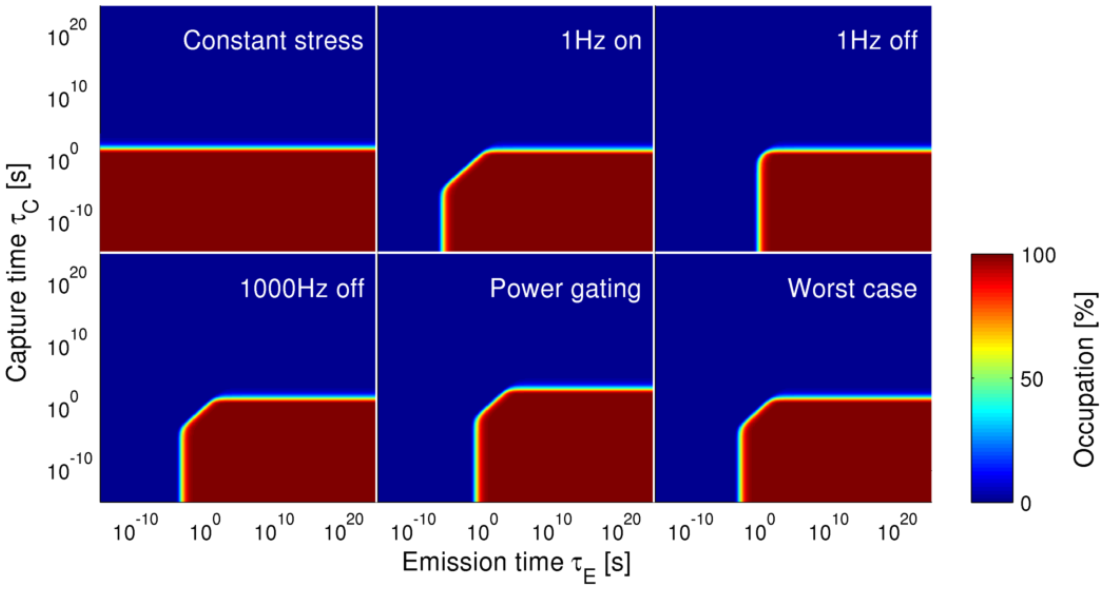


Figure 5.13: Generation of CET occupation based on three parameter abstraction for the different stress examples of Figure 5.10

	Stress voltage [V]	Temperature [K]	ΔV_{th} [mV] Transient	Permanent	Permanent part of ΔV_{th} [%]
22 nm technology	-0.8	300	241.4	429.3	64.0
		350	240.9	427.0	63.9
		400	240.3	424.8	63.9
	-1	300	243.3	429.5	63.8
		350	243.2	427.3	63.7
		400	243.0	425.4	63.6
130 nm technology	-1.5	300	28.9	115.8	80.0
		350	29.5	115.5	79.7
		400	30.2	115.1	79.2
	-2	300	76.9	130.0	62.8
		350	77.4	129.4	62.6
		400	77.8	128.7	62.3

Table 5.1: Overall ΔV_{th} of transient and permanent CET component for both 22 nm and 130 nm technology

times (see Section 3.1.3), importance of the transient component increases with decreasing degradation time. In Table 5.1 a small impact of temperature on the components' overall ΔV_{th} is visible. The effect is not based on the separation criterion but is induced by slightly different values of the traps' equilibrium occupancy difference for different temperatures. However, this effect is very small and can be disregarded.

In order to evaluate the CET based phase space model, the full procedure with consecutive abstractions and occupation recreations, as fundamental part of the phase space generation, has to be applied. Figure 5.14 shows simulated threshold voltage degradations of CET and phase space model for several scenarios with different duty cycles, stress durations and frequencies. These results are also published in [64]. Both models rely on the provided CET map for 130 nm technology and a scenario of accelerated aging (see Figure 5.1). Phase space simulations are based on a time step of 1 minute, resulting in 60, 1440 and 10080 conversions between abstraction and occupation for stress durations of 1 hour, 1 day and 1 week, respectively. The square wave

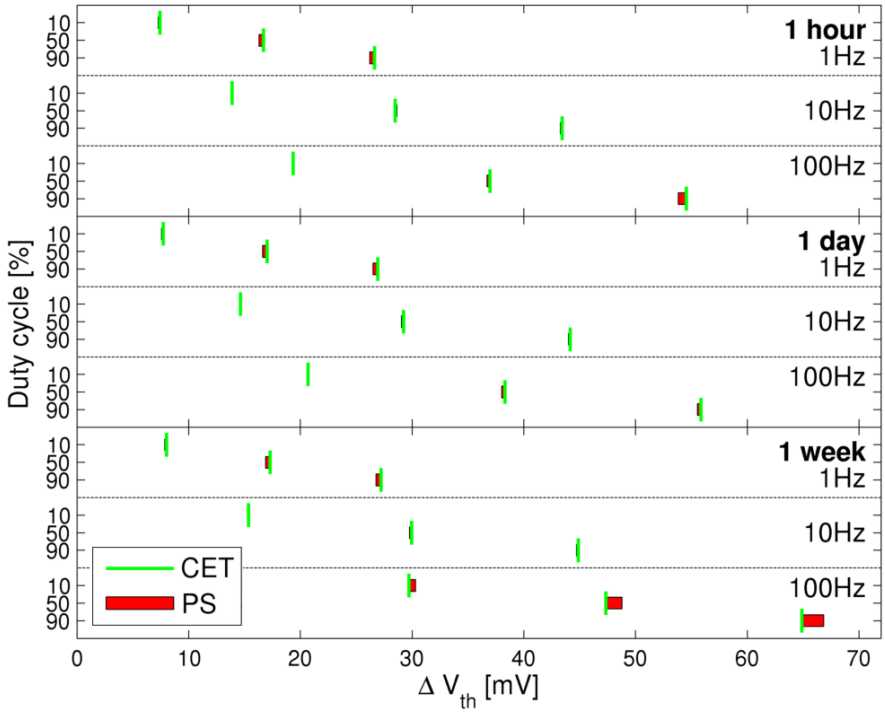


Figure 5.14: Simulated threshold voltage degradations for 130 nm technology and scenarios with 440 K, -2.2 V and different duty cycles (y-axis), stress durations (panels) and frequencies (sub-panels). Panels and sub-panels are divided by solid and dashed horizontal lines, respectively. Short green vertical lines depict simulation results of the CET model based on a provided CET map for accelerated aging (see Figure 5.1) as reference. Deviations between CET and phase space results (PS) are presented using red bars.

signal is applied in the way that the 1 minute intervals always end with a relaxation phase. Hence, occupation profiles that have to be abstracted and recreated are similar to "1 Hz off" of Figure 5.10 and 5.13. Recreation of the occupation is generated directly for the emerging values of abstraction parameters. Inaccuracies of the phase space simulations thus only depend on the consecutive abstraction and occupation recreation procedure and do not incorporate phase space interpolation inaccuracies.

Simulated threshold voltage degradations of CET and phase space model differ only marginally in Figure 5.14. In the different scenarios, threshold voltage degradations are overrated as well as underrated by the phase space approach and neither variation of stress time, frequency nor duty cycle has a distinct impact on relative accuracy of phase space results. Since the scenarios correspond to accelerated aging, the permanent component of ΔV_{th} already accounts for 20 to 51% of overall threshold voltage degradation in these examples.

Threshold voltage degradations for the 22 nm technology are shown in Figure 5.15. Here, a CET map that is based on the provided trap lists is used (see Section 5.1.2). The scenarios correspond to Figure 5.14 with sole stress duration of 1 day and a considerably lower temperature of only 300 K. Instead of marginally differences, simulated threshold voltages of the phase space model are considerably smaller than the reference results of direct CET approach for this technology. Furthermore, contribution of permanent component $\Delta V_{th\,Perm}$ on overall threshold voltage degradation can be neglected.

5.2.3.2 Discussion

The proposed algorithm that generates a CET occupation based on the abstraction parameters $V_{th\,Trans}$, $V_{th\,Perm}$ and slope width is able to precisely recreate the occupations of scenarios with constant stress or periodic stress ending with a relaxation period. Effects of different square wave frequencies and duty cycles are also comprised by the abstraction.

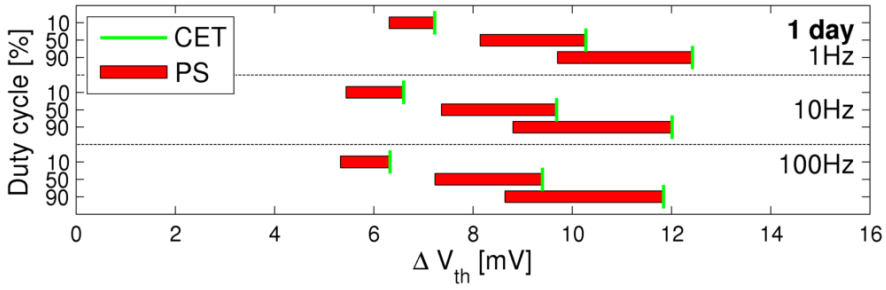


Figure 5.15: Simulated threshold voltage degradations for 22 nm technology and scenarios with 1 day stress duration, 300 K, -0.8 V and different duty cycles (y-axis) and frequencies (sub-panels). Representation corresponds to Figure 5.14 and CET simulations are based on provided trap lists of the switching trap model.

Periodic stress scenarios ending with a stress period provoke the procedure of slope width reduction during generation of occupation. The flat occupation in the short time region within the original occupation (see Figure 5.10 "1 Hz on") results in a large value of slope width, which in turn produces a large region with increasing occupation within the recreation. Traps that are located at very low emission times and slightly higher capture times are thus not activated within the recreation. Therefore, $V_{th\,Trans}$ of the recreated occupation is too small and the parameter slope width has to be reduced in order to decrease the upper τ_E boundary of the recreated slope region which results in an increase of $V_{th\,Trans}$. In this way, occupations of periodic stress scenarios ending with a stress period can be recreated but feature only poor accuracy (see Figure 5.13 "1 Hz on").

This drawback is resolved by using a slightly adapted simulation technique within the phase space approach. Instead of the repetitive abstraction of occupations that are produced by scenarios ending with a stress period, the phase space model always uses equivalent scenarios that end with relaxation periods. Here, the abstraction features a much higher accuracy. The impact of the stress ending has to be added within the model only for the final time step. This is done using a direct CET simulation for a single stress period as a follow up of the phase space simulation. This procedure can also be described by first simulating the low side envelope of a periodic ΔV_{th} curve. If the desired degradation time is attained, the CET simulation of a single stress period induces the transition to the high side envelope of the ΔV_{th} curve.

Occupation steps of the power gating scenario are not characterized by the abstraction. The impact of this restriction on long time phase space simulation has to be examined. This is done after some model improvements in Section 5.2.6. Largest deviation between original and generated occupation examples occurs for the worst case scenario. Since the stress signal is characterized by square wave frequency and duty cycle within the use case of gate level degradation analysis, the phase space simulation technique doesn't need to support such an arbitrary digital stress signal. However, the worst case example shows that abstraction technique and occupation generation procedure are very robust and can be applied for any occupation profile. Main difference between original and generated occupation is always located in the short time region and generated occupation within long time region is quite precise. In this way, the inaccuracy of a phase space simulation may even be eliminated by an extended phase of relaxation or constant stress.

As shown in Table 5.1, the permanent component's impact on overall ΔV_{th} degradation is even higher than the impact of the transient component for very long degradation times. Due to the simple alteration of occupation probability $P(\tau_C)$, the permanent component's occupation is very accurately abstracted with the single abstraction parameter $\Delta V_{th\,Perm}$. In this way, discrepancies in simulated permanent threshold shift between phase space and direct CET approach can only arise because of interpolation deviations within the phase space calculations. These discrepancies are only minimal, when using a high-resolution phase space. Consequently, the overall accuracy of the phase space model increases when the permanent component starts to be a significant share of the overall ΔV_{th} degradation. Thus, accuracy of the phase space model increases for very long degradation times.

The evaluation of the phase space model is supposed to only assess the consecutive abstraction and occupation recreation procedure, since phase space interpolation accuracy just depends on the number of initial characterization values of each parameter. Therefore, interpolation accuracy only depends on the phase space's calculation duration and is not a fundamental part of the model. The basic method of consecutive abstractions and occupation recreations within the phase space model is able to accurately simulate long time threshold voltage degradation as shown in Figure 5.14. These results rely partially on over 10000 conversions between CET occupation and abstraction without introducing clear inaccuracies. However, simulation results for the 22 nm technology with a considerably lower temperature indicate clear deviations to the reference results. These deviations may be justified by the negligible contribution of the permanent threshold voltage component and the described relation between impact of the permanent component and the model's accuracy. A different explanation of the deviations is infeasibility of the abstraction parameter $\Delta V_{th\,Perm}$ for the phase space approach. The characterization of the occupation height within the transient component may be too inaccurate, if there is practically no permanent shift in threshold voltage after a single phase space time interval. In this way, accuracy of the phase space model may be increased by using a different abstraction parameter for the characterization of the occupation height within the transient component.

5.2.4 First Iteration of Model Improvements: Maximum Capture Time

As stated in the last section, occupation height within the transient component shall be abstracted differently in order to avoid possible inaccuracies due to hardly existing permanent shifts in threshold voltage. Thus, a new abstraction parameter, called "maximum capture time", is used for the abstraction of the transient component. This parameter is a direct substitution of parameter $\Delta V_{th\,Perm}$ within the phase space model.

Maximum capture time is defined as capture time of the 50% occupation value within last τ_E column of the occupation probability P. In this way, the transient component's occupation height (see Figure 5.12) is characterized directly without usage of the permanent component. Hence, the described effect, that hardly existing permanent shifts may cause the inaccuracies in Figure 5.15, cannot occur with the abstraction parameter maximum capture time. As long as there is still activation in the very high emission time region of the occupation probability P, which may never decline under realistic stress conditions depending on the emission time scale of the utilized CET map, the maximum capture time value cannot be diminished due to arbitrary stress scenarios. However, main restriction of possible maximum capture time values is the capture time scale of the utilized CET map. In this way, it can again be ensured that the phase space incorporates all degradation states and a trajectory can't leave the area of precomputed values.

Final ΔV_{th} degradation of the permanent component that has to be calculated during phase space simulations can rely on the interpolated maximum capture time of the final phase space time step. This potential method of calculation rests upon the direct relation between maximum capture time and the overall time that a transistor is exposed to stress. In this way, the basic simulation flow of Figure 5.7 can still be used with a maximum capture time based phase space and a conversion of final maximum capture time to $\Delta V_{th\,Perm}$.

However, final permanent ΔV_{th} degradation can also be calculated based on a direct CET approach without a distinct performance impact. Frequency and duty cycle profiles of the stress scenario account for the transistor's overall stress time. Since there is no relaxation process for the permanent component, the actual stress pattern is irrelevant and final permanent ΔV_{th} degradation can be calculated using Equations 3.1 and 3.3 within a single computational step for the whole stress time. Computational overhead of this operation is thus negligible. Accuracy of the phase space approach is theoretically improved by this method, because the permanent component is calculated in the same way as within the reference CET model without being affected by interpolation inaccuracies due to the phase space's resolution. Therefore, this method is used for the calculation of the permanent ΔV_{th} degradation within the improved phase space approach as shown in Figure 5.16. The phase space interpolation process is thus only used for the transient component and accuracy of the overall simulation method generally improves with increasing share of the permanent component on the entire threshold voltage degradation.

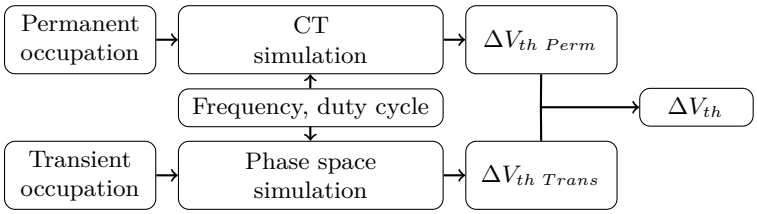


Figure 5.16: Improved NBTI simulation flowchart of the phase space model with detached calculation of transient and permanent ΔV_{th} component. Capture time (CT) simulation refers to a direct simulation of the permanent component with the reference CET model and phase space interpolation relegates to the simulation method of Figure 5.7.

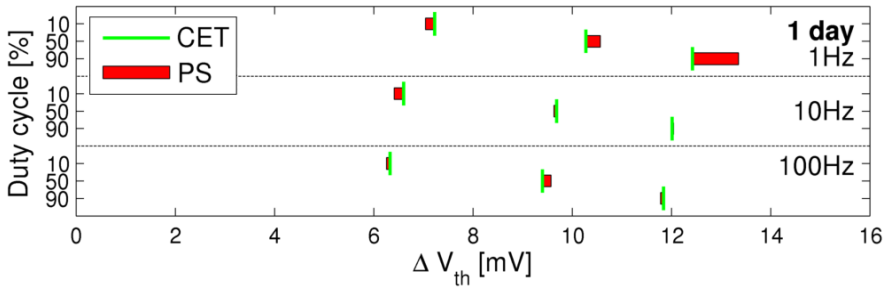


Figure 5.17: Simulated threshold voltage degradations for 22 nm technology as shown in Figure 5.15. Phase space results are calculated using the abstraction parameter maximum capture time.

An additional improvement of the phase space technique can be utilized, due to the detached calculation of transient and permanent ΔV_{th} component. Very long degradation times may cause a total annealing of transient ΔV_{th} while there is still activation in the very high emission time region of the occupation probability P. This scenario can thus only occur when there are no defects within the activated very high emission time region of the CET map. As defined, the abstraction parameter maximum capture time still characterizes the occupation height within the activated region. In this way, a subsequent stress phase causes a major inaccuracy within the phase space model, since the occupation is abstracted using always the same maximum capture time value until the overall stress time after the relaxation phase exceeds the stress time before relaxation phase. In order to avoid this problematic abstraction state, the improved phase space model resets the maximum capture time parameter whenever there is a total annealing of the transient threshold voltage shift.

5.2.4.1 Evaluation

The improved phase space model is used to simulate the threshold voltage degradations for the scenarios that are utilized in Figure 5.15. As shown in Figure 5.17, the differences between phase space simulations and the results of the CET reference model are considerably reduced by the improved modeling approach. The improvements can only originate from the usage of the abstraction parameter maximum capture time instead of $\Delta V_{th\,Perm}$, as Figure 5.15 and 5.17 feature the same contribution of the permanent threshold voltage component. Furthermore, the utilized scenarios cannot provoke the described method to reset the maximum capture time after long relaxation phases.

5.2.4.2 Discussion

The improved phase space model with detached calculation of transient and permanent ΔV_{th} component and abstraction parameter “maximum capture time” is able to considerably improve the simulation’s accuracy for scenarios that are critical for the basic phase space model with negligible computational overhead. The phase space model can already simulate stress scenarios with varying activity states and power gating intervals as these scenarios are directly supported

		ΔV_{th} deviation [%]	
		Transient	Permanent
22 nm technology	Stress voltage	1.12	0.14
	Temperature	0.46	1.06
130 nm technology	Stress voltage	166.09	12.26
	Temperature	4.50	1.01

Table 5.2: Maximal overall ΔV_{th} deviation [%] between maps for different stress voltages and maps for different temperatures

by the CET approach, which is used as base model for the generation of the phase space. Evaluation results for these scenarios as well as additional evaluation results with considerably longer degradation times are presented in Section 5.2.6. However, additional considerations are required to execute phase space simulations for stress scenarios that are not directly supported by the CET model.

5.2.5 Second Iteration of Model Improvements: Temperature Variations

The phase space model does not support mission scenarios with varying temperatures or gate voltages so far. Within this section, extensions of the phase space approach are presented that facilitate these temperature variations. The CET model, which is used as reference model within the phase space approach, relies on CET maps that each represent a single temperature and gate voltage. In this way, the CET modeling technique has to be extended to support transitions between different CET maps in order to support varying temperatures or voltages within the phase space approach.

Table 5.2 shows the maximal overall ΔV_{th} deviation between maps for different stress voltages and maps for different temperatures. The maximal overall ΔV_{th} deviation of 166.09% for the transient component of 130 nm technology denotes for instance that the comparison of the CET maps for different stress voltages results in maximal overall ΔV_{th} of these CET maps being 166.09% larger than minimal overall ΔV_{th}. As indicated by this example, there might be a huge difference in overall ΔV_{th} for CET maps of different stress voltages. This is due to some switching trap defects being practically only activated at high stress voltages (see Section 3.1.2). On the other hand, the overall ΔV_{th} for CET maps of different temperatures is nearly constant. Only minor deviations arise by reason of the traps' equilibrium occupancy differences (EOD) being slightly dependent on temperature (see Section 3.1.3). Since capture and emission time intervals of the used CET maps are chosen generously, possible displacements of the defects within τ_C and τ_E dimension do not exceed the CET maps' axis intercepts and thus do not affect the overall ΔV_{th} in Table 5.2. Figure 5.18 shows an example of temperature induced defect displacements by using a superposition of the CET maps for 300 and 400 K. Since defect activation is a temperature activated process, locations of the defects are shifted to lower capture and emission times within a CET map for higher temperatures. Hence, a NBTI simulation with the CET model that just adopts the occupation probability P within the CET map of previous temperature for the occupation probability within the CET map of subsequent temperature would cause an abrupt change in threshold voltage. As a result of nearly constant overall ΔV_{th} of different CET maps and predictable displacements of the defects within the CET maps, CET based simulations for mission scenarios with varying temperatures may instead be feasible by using a conversion technique to estimate the occupation probability P within the CET map of subsequent temperature based on the occupation probability within the CET map of previous temperature. Two different approaches to realize such a conversion technique are presented in Section 5.2.5.1 and 5.2.5.2.

The usage of a CET conversion technique to support varying temperatures provokes some enhancements within the phase space simulation flowchart as shown in Figure 5.19. The initial occupations within transient and permanent component may be converted based on the current temperature. As described in Section 5.2.4, threshold voltage degradations within the

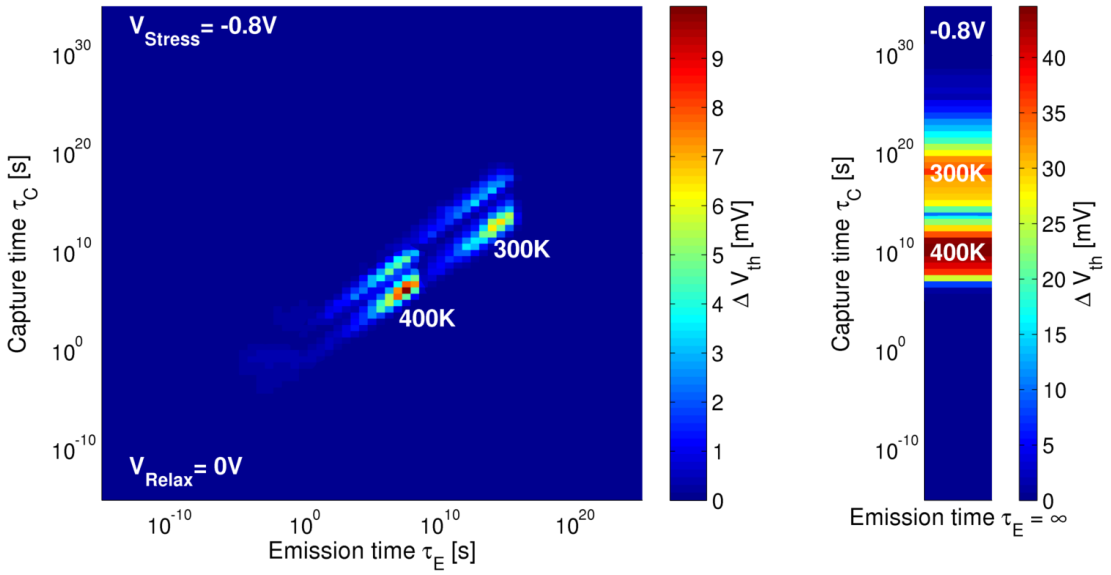


Figure 5.18: Superposition of the CET maps of the 22 nm technology for 300 and 400 K. The CET maps are presented separately in Figure 5.11

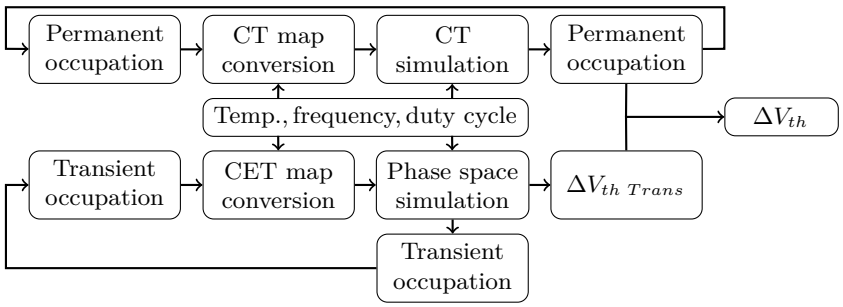


Figure 5.19: Improved NBTI simulation flowchart of the phase space model in order to support temperature variations by integrating occupation conversions between CET maps for different temperatures. Phase space and CT simulation refer to simulation method of Figure 5.7 and direct simulation of the permanent component with reference CET model, respectively.

transient and permanent component are calculated with phase space interpolations and direct CET approach, respectively. For this reason, temperature is introduced as a new dimension within the underlying phase space. The resulting value of $\Delta V_{th\,Trans}$ and the occupation of the permanent component are used to estimate the overall threshold voltage degradation. Furthermore, occupation of the permanent component and occupation of the transient component, which is generated based on the resulting abstraction parameter values, can afterwards be used as initial occupations for an adjacent time interval with another temperature.

An additional modification is introduced within the phase space approach in order to enable reasonable comparisons between phase space and MSA simulation results. As stated in Section 3.1.5, the MSA approach simulates the high side envelope of the threshold voltage degradation curve. Therefore, the stress scenario has to end with a stress phase instead of a relaxation phase as used in Section 5.2.3.1 and 5.2.4.1. Nevertheless, the phase space approach has to use a square wave signal that is applied in the way that the phase space time intervals always end with a relaxation phase due to accuracy of the abstraction procedure (see Figure 5.13). A direct CET simulation for a single stress period is thus always used as a follow up of the phase space simulation (see Section 5.2.3.2). The stress times of reference switching trap model, CET model and MSA approach are extended accordingly so as to always simulate threshold voltages of high side envelope for the same degradation time.

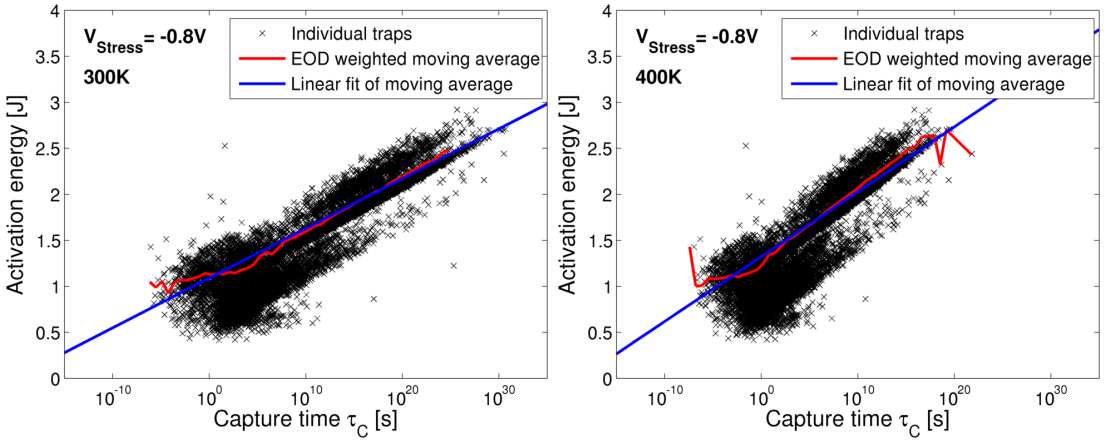


Figure 5.20: Correlations between capture time and capture activation energy of individual traps of the 22 nm technology for 300 K (left panel) and 400 K (right panel). A weighted moving average that uses the individual traps' EOD as weight function and the linear fit of the moving average are represented by red and blue lines, respectively.

5.2.5.1 Activation Energy Based Transformation

The temperature conversion of the permanent component can be realized easily by using the methodology that is already disposed for the abstraction parameter $\Delta V_{th\,Perm}$. As the transition between CET maps for different temperatures may not trigger an abrupt change of permanent shift in threshold voltage, calculations of $\Delta V_{th\,Perm}$ before and after the transition must result in the same value. Consequently, the occupation of the permanent component after a temperature transition can be generated based on $\Delta V_{th\,Perm}$ of the previous CET map. As described in Section 5.2.3, this can be done by increasing a virtual capture time as long as $\Delta V_{th\,Perm}$ of the resulting occupation matches the desired value of $\Delta V_{th\,Perm}$.

The temperature conversion of the transient component must account for a transition in capture as well as emission time. Basic approach of this conversion is to use the three abstraction parameters described above. In that case, maximum capture time and slope width have to be adjusted for the new CET map based upon the parameter values of the old map and the corresponding temperature values.

The activation energy based transformation that is presented in this section uses correlations between capture time and capture activation energy to calculate the adjusted value of maximum capture time. Capture times and capture activation energies of individual traps are shown for two different temperatures in Figure 5.20. Although there is a clear variation in the relation between capture time and capture activation energy for individual traps, an overall correlation is recognizable. This is more evident within the weighted moving average that uses the individual traps' EOD as weight function and the linear fit of the moving average, which are also visible in Figure 5.20. Here, the weighted moving average breaks off in the region of very high capture times for 300 K, as the EOD of all traps with higher capture time is equal to zero for this temperature. In the final conversion, the value of maximum capture time is used to estimate the corresponding capture activation energy for the previous temperature based on the linear fit of the weighted moving average. The capture activation energy is in turn used to estimate the maximum capture time for the subsequent temperature in the same way. Thus, the abstraction parameter maximum capture time is converted for the CET map of the subsequent temperature. In order to apply this conversion technique within CET and phase space model, the linear fits of the weighted moving averages of the relation between capture time and capture activation energy have to be calculated for every CET map that may be used within the NBTI simulations. Furthermore, the parameters of these linear fits have to be stored as an additional characteristic of each CET map.

The slope width of the occupation within the new CET map is estimated similarly. The linear fit of the weighted moving average of the relation between emission time and emission activation energy is calculated for every CET map and stored as an additional characteristic of

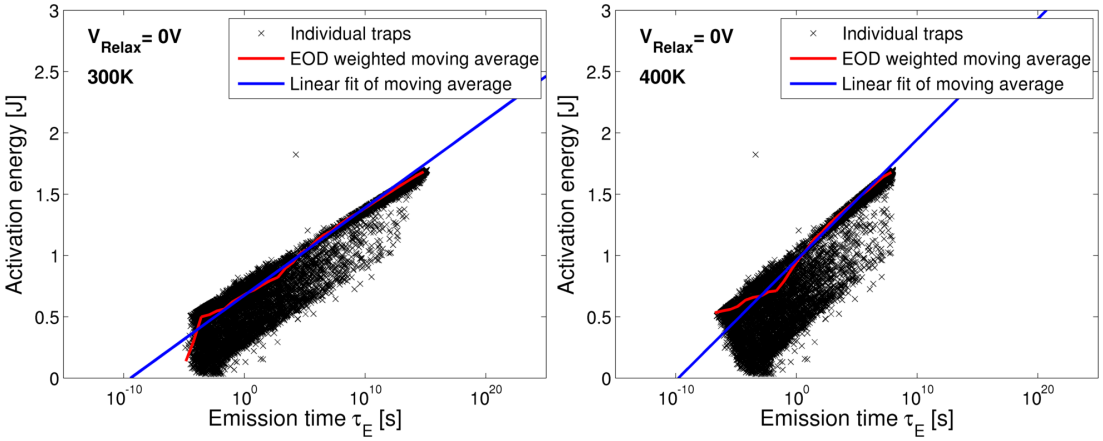


Figure 5.21: Correlations between emission time and emission activation energy of individual traps of the 22 nm technology for 300 K (left panel) and 400 K (right panel). A weighted moving average that uses the individual traps' EOD as weight function and the linear fit of the moving average are represented by red and blue lines, respectively.

each CET map (see Figure 5.21). Emission time values of the old CET map can consequently be transferred to emission time values of the new CET map as specified by Equation 5.14.

$$\begin{aligned} & E_{a\,E} = s \cdot \tau_E + b \\ \Rightarrow \qquad & \tau_{E\,new} = \frac{s_{old} \cdot \tau_{E\,old} + b_{old} - b_{new}}{s_{new}} \end{aligned} \tag{5.14}$$

Inserting this emission time conversion in the slope width SW definition of Equation 5.13 yields

$$\begin{aligned} SW_{new} &= log\left(\frac{s_{old} \cdot \tau_{E\,upper\,old} + b_{old} - b_{new}}{s_{new}} - \frac{s_{old} \cdot \tau_{E\,lower\,old} + b_{old} - b_{new}}{s_{new}}\right) \\ &= log\left(\frac{s_{old}}{s_{new}}\left(\tau_{E\,upper\,old} - \tau_{E\,lower\,old}\right)\right) \\ &= SW_{old} + log\left(\frac{s_{old}}{s_{new}}\right) \end{aligned} \tag{5.15}$$

The abstraction parameter slope width is thus converted for the CET map of the subsequent temperature by just using the previous value of the slope width and pre-calculated gradients of the correlation between emission time and emission activation energy. As can be seen in Figure 5.21, the slope of the linear fit increases with increasing temperature. In this way, the parameter slope width is reduced or increased at a transition point with increasing or decreasing temperature, respectively.

5.2.5.2 Transformation Based on Permanent Shift in Threshold Voltage

The temperature transformation technique that is presented in this section uses a very similar approach. The occupation conversion relies on the abstraction parameters and the methodologies for the permanent component as well as slope width conversion are adopted for this transformation technique. However, the parameter maximum capture time is converted slightly different. The maximum capture time of the permanent component within the new CET map is estimated by identifying the 50% occupation value within the permanent occupation probability that is already converted. Afterwards, this maximum capture time value is used as abstraction parameter of the transient component for the new CET map. A basic assumption of this transformation approach is that defects of transient and permanent component are affected similarly by the changing temperature.

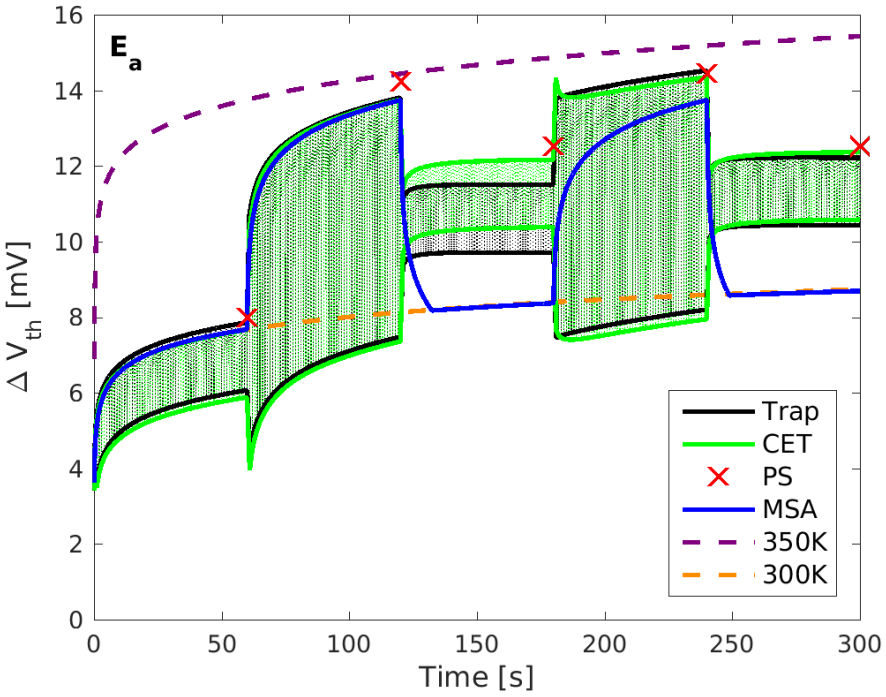


Figure 5.22: Time traces of simulated threshold voltage degradations for 22 nm technology and short time scenarios with -0.8 V, 1 Hz, 90% duty cycle and transitions between 300 and 350 K based on 1 minute intervals. Black lines indicate simulation results of the switching trap model (Trap) that serves as a reference. Simulation results of CET model, phase space model (PS) and mission scenario aware analytical approach (MSA) are represented by green, red and blue, respectively. CET and phase space models use the activation energy based transformation method of Section 5.2.5.1. Traces of switching trap and CET model contain a fine structure of the individual stress and relaxation phases and the corresponding high and low side envelopes are emphasized. MSA approach simulates the high side envelope with high time resolution, whereas the resolution of the PS model is linked to the used phase space time step of 1 minute. High side envelopes of CET simulation traces for similar scenarios with constant temperature of 300 and 350 K are represented by orange and purple dashed lines, respectively.

5.2.5.3 Evaluation

Time traces of simulated threshold voltage degradations for 22 nm technology and short time scenarios with temperature transitions are presented within the Figures 5.22 to 5.25. The activation energy based transformation approach is used in Figure 5.22 and 5.24, whereas Figure 5.23 and 5.25 rely on the transformation method that utilizes the permanent shift in threshold voltage. Furthermore, temperature transformations of 300 - 350 K are covered in Figure 5.22 and 5.23, while Figure 5.24 and 5.25 examine temperature transformations of 300 - 400 K. Simulation results of the switching trap model (black) always serve as a reference, since the model directly supports these scenarios.

Comparisons between simulation results of switching trap model (black) and temperature transformation enhanced CET model (green) indicate that the transformation approaches are able to support the increasing and decreasing threshold voltage shifts due to temperature variations. The activation energy based transformation approach seems to be better suited for the conversions, as the relaxation in the 300 K phases of Figure 5.23 are clearly overrated. However, impact of the transformation methods within long time scenarios has to be also examined. Minor deviations between switching trap and CET results already occur in the first stress phase. These deviations cannot be provoked by the temperature transformations but originate from the intrinsic error of the CET approach.

Results of the mission scenario aware analytical approach (MSA) demonstrate the time behavior that is already described in Section 3.1.5. Within the low temperature phases of the

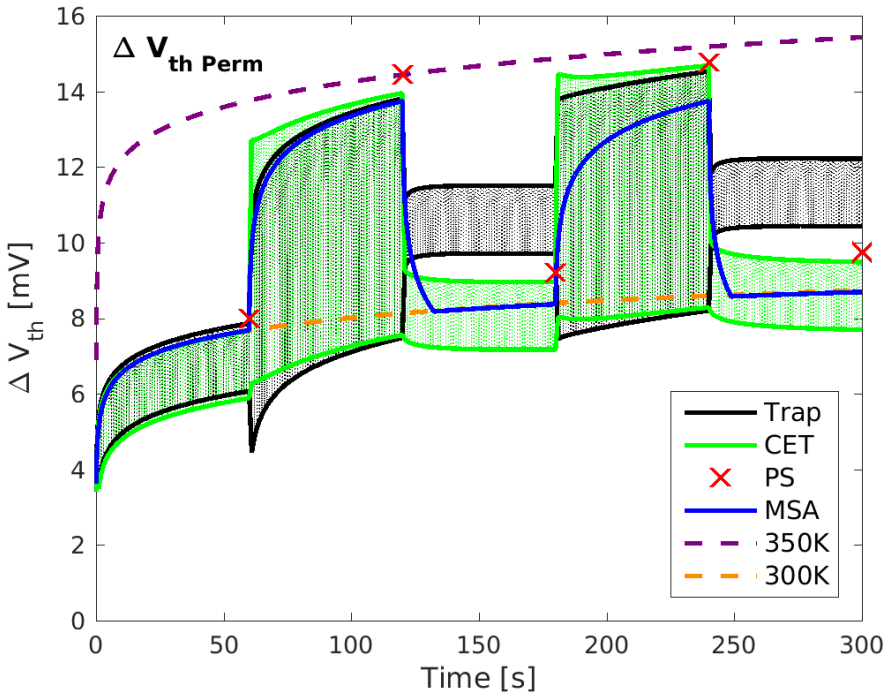


Figure 5.23: Time traces of simulated threshold voltage degradations for 22 nm technology and short time scenarios with -0.8 V, 1 Hz, 90% duty cycle and transitions between 300 and 350 K based on 1 minute intervals. CET and phase space models use the transformation based on permanent shift in threshold voltage of Section 5.2.5.2. Representation corresponds to Figure 5.22.

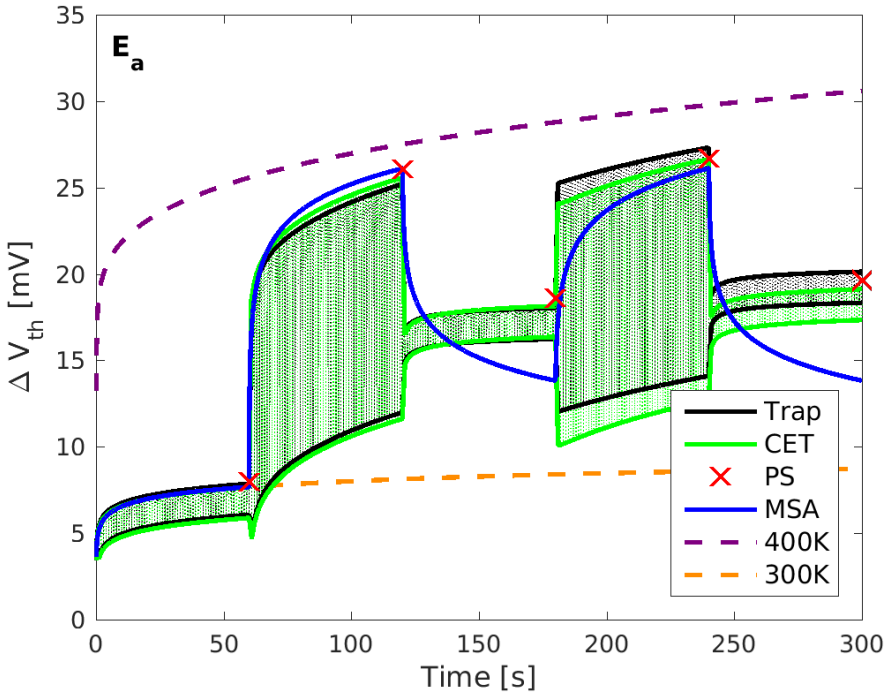


Figure 5.24: Time traces of simulated threshold voltage degradations for 22 nm technology and short time scenarios with -0.8 V, 1 Hz, 90% duty cycle and transitions between 300 and 400 K based on 1 minute intervals. CET and phase space models use the activation energy based transformation method of Section 5.2.5.1. Representation corresponds to Figure 5.22.

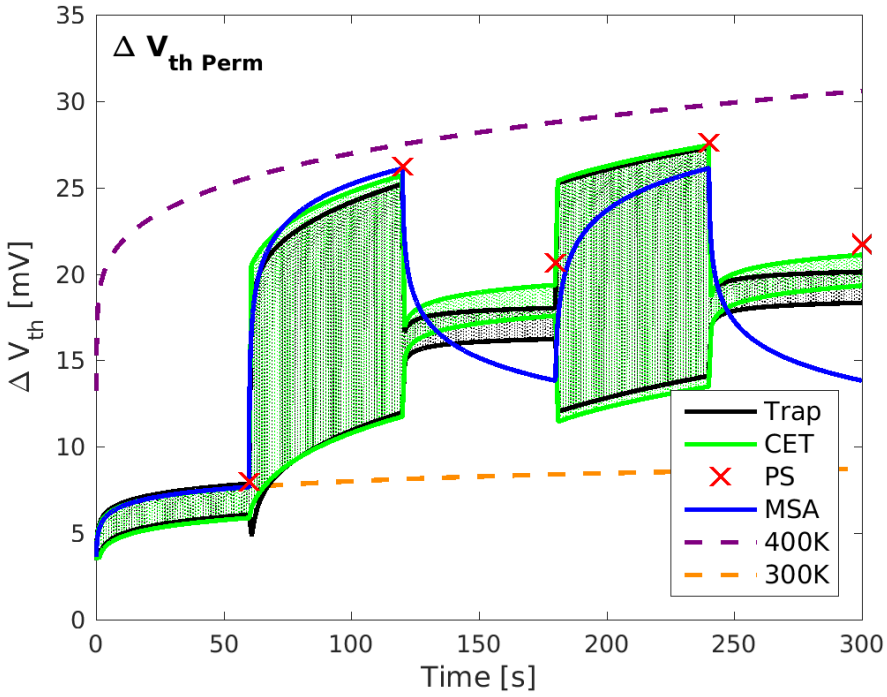


Figure 5.25: Time traces of simulated threshold voltage degradations for 22 nm technology and short time scenarios with -0.8 V, 1 Hz, 90% duty cycle and transitions between 300 and 400 K based on 1 minute intervals. CET and phase space models use the transformation based on permanent shift in threshold voltage of Section 5.2.5.2. Representation corresponds to Figure 5.22.

scenario with transitions between 300 and 350 K, the MSA simulation results change from a relaxation characteristic to the degradation behavior of the scenario with constant temperature of 300 K. Furthermore, MSA results show a long lasting relaxation behavior in the low temperature phases of the scenario with transitions between 300 and 400 K. In contrast, switching trap and CET results feature additional degradation after an abrupt relaxation in these phases. In addition, the MSA threshold voltage degradation is at the end of the second high temperature phase always hardly larger than the value at the end of the first phase. Consequently, the simulation results of the temperature transformation enhanced CET model are always in a better agreement with the switching trap model.

The phase space approach simulates the high side envelope of the CET degradation curve only at the end of each temperature phase with little additional error. These examples also illustrate each model's temporal resolution. Whereas switching trap and CET model simulate the fine structure of the individual stress and relaxation phases and the MSA approach tracks the high side envelope with high temporal resolution, the phase space model is linked to the used phase space time step of 1 minute.

Evaluation results for longer degradation times and both temperature transformation methods are shown in Figures 5.26, 5.27 and 5.28. These figures have a new form of representation that is similar to Figure 5.14 but uses simulation results of the switching trap model as reference (black vertical lines) for the deviations of CET, phase space and MSA model. Both CET and phase space models use the same temperature transformation methods, however only the phase space approach comprises the parameter abstractions. Simulation results of switching trap and MSA model slightly deviate for the different transformation methods, since the temperature transitions are defined using a probability of high temperature state based on 10 minute intervals, which results in slightly different versions of the actual temperature variation characteristic. However, within a single scenario all four models are always served with the same temperature variation characteristic. Tables 5.3, 5.4 and 5.5 specify the root mean square of the relative errors in relation to the simulation results of switching trap model for each

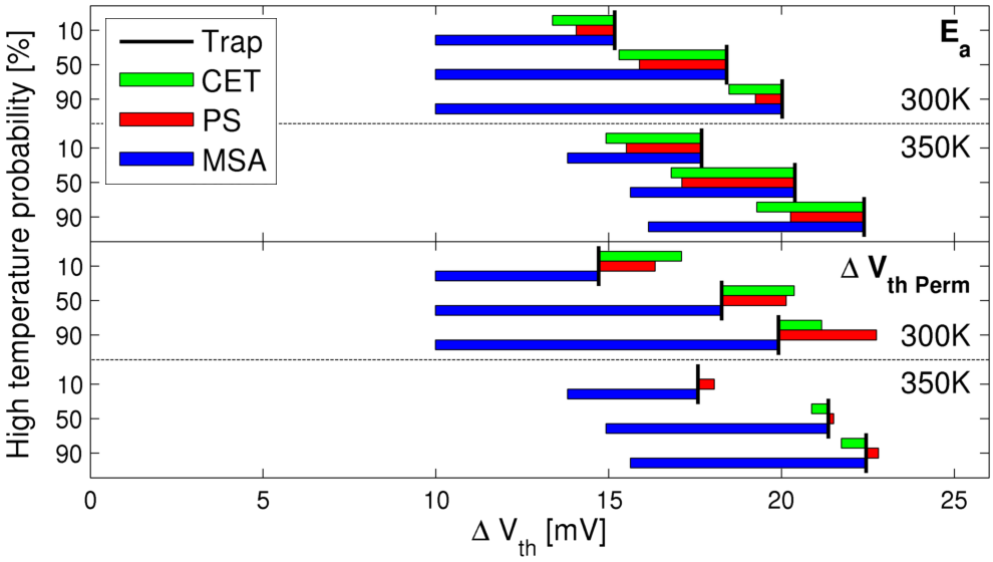


Figure 5.26: Simulated threshold voltage degradations for 22 nm technology, different temperature transformation methods and scenarios with 1 day stress duration, -0.8 V, 1 Hz, 50% duty cycle and temperature conversions between 300 and 350 K based on 10 minute intervals. Different high temperature probabilities, transformation methods and final temperature values are represented by y-axis, panels (divided by solid horizontal lines) and sub-panels (divided by dashed horizontal lines), respectively. Short black vertical lines depict simulation results of the switching trap model (Trap) based on provided trap lists (see Section 5.1.1) as reference. Deviations between switching trap simulation results and results of CET model, phase space model (PS) and mission scenario aware analytical approach (MSA) are presented using green, red and blue bars, respectively.

		Method	
		E_a	ΔV_{thPerm}
CET	RMS of rel. error [%]	16.94	7.68
	Underestimation [%]	100	50
PS	RMS of rel. error [%]	11.88	12.03
	Underestimation [%]	100	0
MSA	RMS of rel. error [%]	62.12	62.62
	Underestimation [%]	100	100

Table 5.3: Root mean square of relative errors and underestimation percentages of the simulation results presented in Figure 5.26. Stated data of CET model, phase space model (PS) and mission scenario aware analytical approach (MSA) refer to the corresponding simulation results of the switching trap model (Trap) and are divided by the different temperature transformation methods of Figure 5.26.

model and temperature transformation approach. Furthermore, the respective underestimation percentages are also presented. These tabulated numbers provide an overview of the quality of each transformation technique by specifying mean deviations and mean underestimation tendencies for each transformation technique.

The activation energy based transformation approach always underestimates the threshold voltage shifts for these longer degradation times, whereas the transformation based on permanent shift in threshold voltage over- and underestimates the actual degradation for the 22 nm technology. In addition, the overall deviation of the simulation results is significantly smaller for the $\Delta V_{th\,Perm}$ approach. Interestingly, results of the short time scenario with transitions between 300 and 350 K have a tendency of overrating the relaxation in the 300 K phases for this

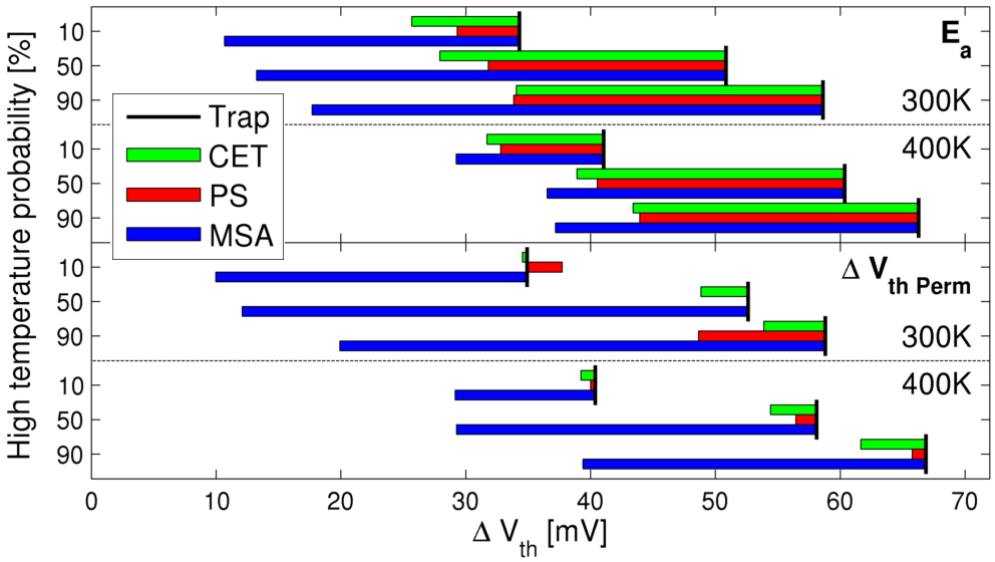


Figure 5.27: Simulated threshold voltage degradations for 22 nm technology, different temperature transformation methods and scenarios with 1 day stress duration, -0.8 V, 1 Hz, 50% duty cycle and temperature conversions between 300 and 400 K. Representation corresponds to Figure 5.26.

		Method	
		E_a	ΔV_{thPerm}
CET	RMS of rel. error [%]	57.39	6.73
	Underestimation [%]	100	100
PS	RMS of rel. error [%]	44.43	5.63
	Underestimation [%]	100	67
MSA	RMS of rel. error [%]	180.2	195.4
	Underestimation [%]	100	100

Table 5.4: Root mean square of relative errors and underestimation percentages of the simulation results presented in Figure 5.27. Data representation corresponds to Table 5.3.

		Method	
		E_a	ΔV_{thPerm}
CET	RMS of rel. error [%]	20.47	15.06
	Underestimation [%]	100	100
PS	RMS of rel. error [%]	19.96	15.5
	Underestimation [%]	100	100
MSA	RMS of rel. error [%]	32.43	32.7
	Underestimation [%]	100	100

Table 5.5: Root mean square of relative errors and underestimation percentages of the simulation results presented in Figure 5.28. Data representation corresponds to Table 5.3.

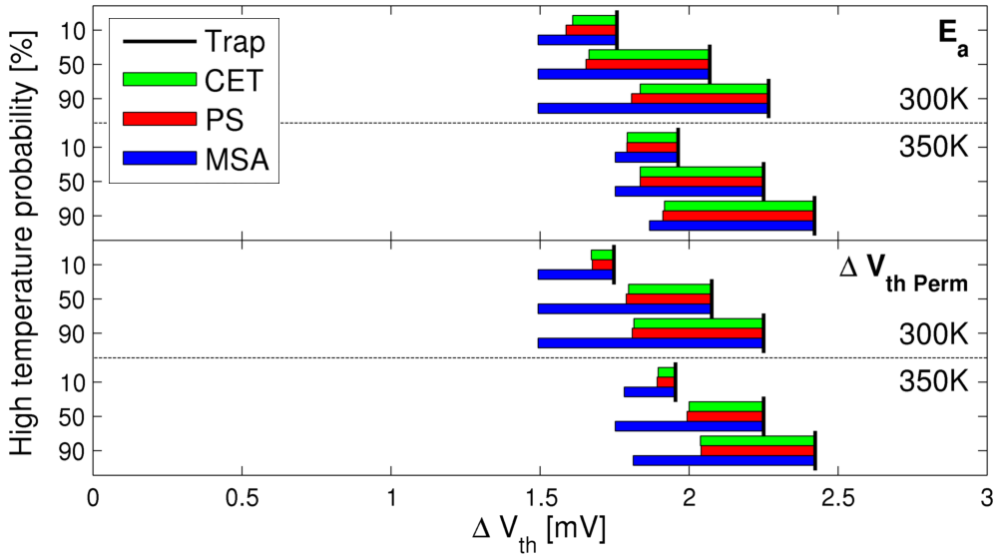


Figure 5.28: Simulated threshold voltage degradations for 130 nm technology, different temperature transformation methods and scenarios with 1 day stress duration, -1.5 V, 1 Hz, 50% duty cycle and temperature conversions between 300 and 350 K. Representation corresponds to Figure 5.26.

transformation technique (see Figure 5.23), whereas the corresponding longer time scenarios with final 300 K phases show an overestimation of the degradation (see Figure 5.26). For 130 nm technology, both temperature transformation methods result in constant underestimations of the degradation effect and larger relative deviations to the simulation results of the switching trap model (see Figure 5.28 and Table 5.5). The overall deviation of the simulation results is again slightly smaller for the transformation based on permanent shift in threshold voltage. Furthermore, the MSA model always clearly underestimates the actual shift in threshold voltage with mean relative deviations up to 195% (see Table 5.4).

The larger relative deviations between simulation results of switching trap model and CET as well as phase space approach for the 130 nm technology are further analyzed. Figure 5.29 presents simulation results for 130 nm technology and scenarios with constant temperatures of 350 and 400 K. Especially in the 400 K scenarios, large relative deviations between switching trap and CET results occur. Since there are no temperature transformations within the presented scenarios, these deviations don't rely on modeling techniques that are developed within this thesis but originate directly from the reference switching trap and the reference CET model. As shown in Figure 5.30, these large deviations for 400 K scenarios cannot be reduced by using a CET map with a significantly higher resolution of 320 τ_C and 256 τ_E bins or by using a significantly higher temporal resolution of the switching trap model via an internal switching trap time step of 0.1 ms (see Section 5.1.1). On the other hand, Figures 5.29 and 5.30 only exhibit minor deviations between CET and phase space results.

5.2.5.4 Discussion

Time traces of simulated threshold voltage degradations for short time scenarios with temperature transitions show that the CET approach with temperature transformation technique obtains simulation results with higher accuracy than the MSA model. In particular, the activation energy based transformation yields the most accurate results. On the contrary, MSA results indicate that the simulated threshold voltage at the end of the 300 K phases is practically independent of number and duration of phases with 350 K for the scenario with transitions between 300 and 350 K. Furthermore, the MSA threshold voltage degradation at the end of the high temperature phases hardly increases with each additional phase which may cause a clear underestimation of the degradation within long time scenarios.

Performed simulations of scenarios with longer degradation time confirm that the MSA

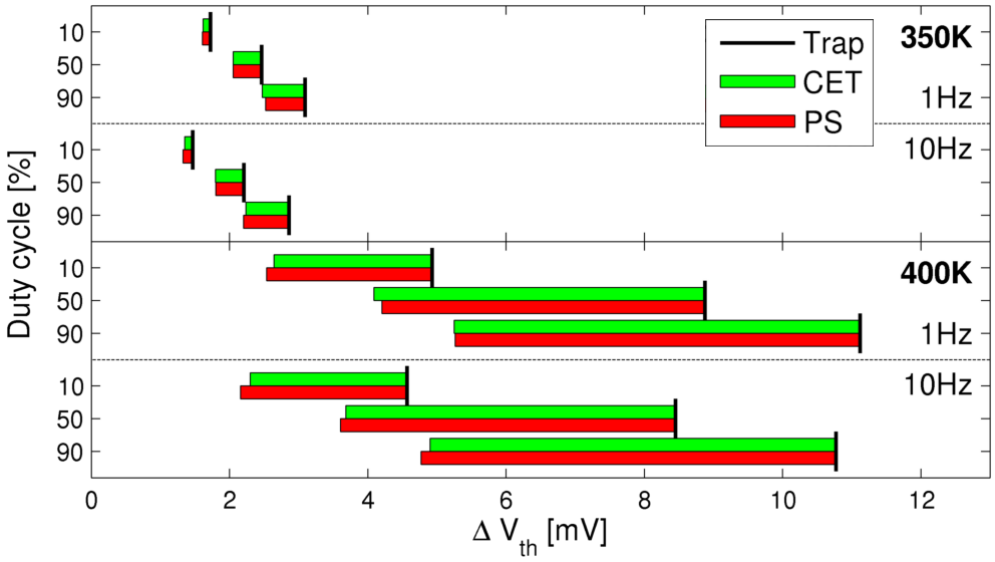


Figure 5.29: Simulated threshold voltage degradations for 130 nm technology and scenarios with 1 day stress duration, -1.5 V and different duty cycles (y-axis), temperatures (panels) and frequencies (sub-panels). Short black vertical lines depict simulation results of the switching trap model (Trap) based on provided trap lists (see Section 5.1.1) as reference. Deviations between switching trap simulation results and results of CET model and phase space model (PS) are presented using green and red bars, respectively.

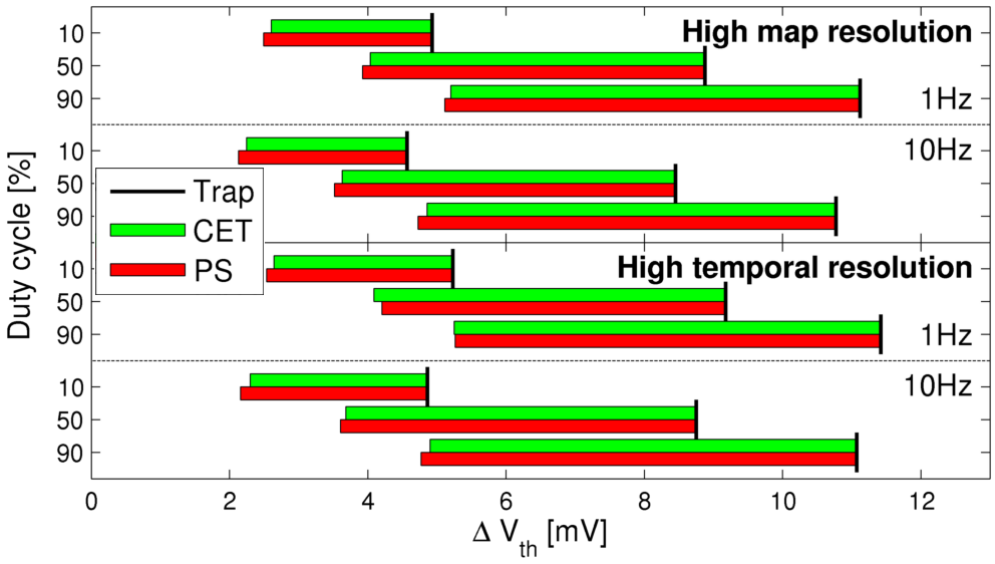


Figure 5.30: Simulated threshold voltage degradations for 130 nm technology, changed parameters of the reference models and scenarios with 1 day stress duration, 400 K, -1.5 V and different duty cycles (y-axis) and frequencies (sub-panels). Top and bottom panel depict results based on a significantly higher resolution of the CET map and a significantly higher temporal resolution of the switching trap model, respectively. Representation of simulation results corresponds to Figure 5.29.

model clearly underestimates the degradation effect for longer degradation times. Phase space and CET models with temperature transformation are able to simulate the long time threshold voltage degradation with minor deviations. Here, deviations of simulated threshold voltages are considerably smaller for the temperature transformation based on permanent shift in threshold voltage. Furthermore, minor drawbacks of this transformation approach, which occur within the simulated short time traces, cannot be observed within long time scenarios. Therefore, transformation based on permanent shift in threshold voltage as described in Section 5.2.5.2 is utilized as temperature transformation method in the following evaluation examples.

Large relative deviations between simulation results of switching trap model and CET as well as phase space approach for the 130 nm technology and scenarios with temperature conversions between 300 and 350 K presumably rely on inaccuracies between reference switching trap and CET model for elevated temperatures. To the best of my knowledge, such inaccuracies between switching trap and CET models that are both based on the same trap list are not discussed in the literature. Simple modifications within the modeling technique, like CET resolution and time resolution within the model, cause no improvement in the CET model's accuracy. However, methods that are used to generate CET maps from a provided trap list may also heavily affect the accuracy. In particular, the effects of parameters within EOD calculation or Gaussian convolution as described in Section 3.1.3 may be analyzed. On the other hand, even within the presented 400 K scenarios the absolute threshold voltage degradation is small and may be determined by a few traps. Significantly longer degradation times may enhance the model's accuracy in this case.

Nevertheless, simulation results of the phase space model approximately match the CET model's results within these scenarios. Since the phase space modeling approaches that are introduced within this thesis rely on the CET model and are only affected indirectly by the inaccuracies between switching trap and CET model, possible reasons of these inaccuracies are not further analyzed. Therefore, simulations for 130 nm technology and scenarios with 400 K are omitted in the following evaluation examples and inaccuracies in scenarios with 350 K in Section 5.2.6 may partly be attributed to an imprecision between CET and switching trap models. In order to better asses these imprecisions, threshold voltage degradations are always calculated for both CET and switching trap models within the following sections even if the CET model fully supports the particular scenario and switching trap results are not required to evaluate the phase space approach. The imprecisions within these scenarios may be used to draw conclusions regarding the proportion between phase space inaccuracies and inaccuracies directly caused by the CET approach for scenarios that are only supported by the switching trap model.

5.2.6 Third Iteration of Model Improvements: Reference Map

The method to support varying temperatures within NBTI simulations of Section 5.2.5 is directly tied to the CET model. As presented in Figures 5.22 to 5.28, the method can also be used within the phase space model by using temperature as a new dimension within the underlying phase space. Consequently, the effort of space construction is increased considerably and the number of supported temperature values has to be restricted. Main objection of this iteration of phase space model improvements is to overcome these drawbacks by using a temperature independent phase space.

A temperature independent phase space may be generated by only using abstraction parameters that characterize the degradation state in the exact same way for different temperatures. This is equivalent to using abstraction parameters which do not rely on the actual CET map. Both maximum capture time and slope width parameters fulfill this requirement. However, $\Delta V_{th\,Trans}$ certainly depends on the used CET map and is thus bound to a particular temperature value. Since $\Delta V_{th\,Trans}$ is used to ensure that there are no abrupt changes in the simulated threshold voltage degradation during consecutive phase space interpolations, it is very unlikely to design a different phase space approach with CET map independent parameters which provides a similar accuracy.

A phase space model with a temperature independent phase space can also be designed in a completely different way. The degradation state of a transistor is fully characterized with the occupation probability P and the threshold voltage degradation is calculated using a

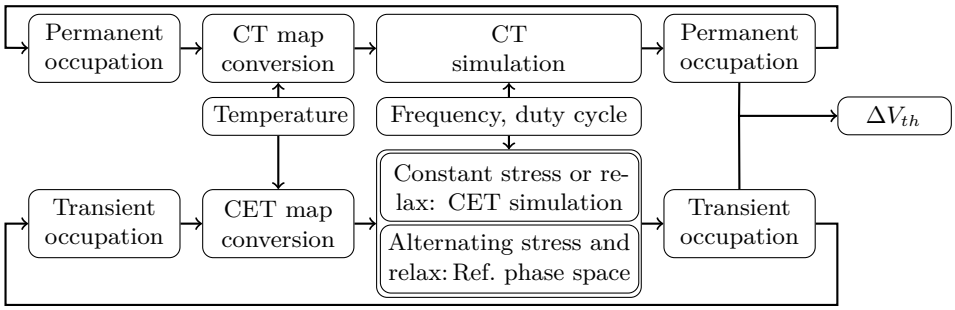


Figure 5.31: Improved NBTI simulation flowchart of the phase space model in order to utilize a temperature independent reference phase space. Degradation simulation of the transient component is based on the CET approach for scenarios with constant stress or relaxation or the phase space method of Figure 5.7 with reference phase space for scenarios with alternating stress and relaxation phases.

multiplication of occupation and CET map (see Equation 3.3). The three abstraction parameters are used to abstract the occupation probability P. Thereby, the parameter $\Delta V_{th\,Trans}$ relies on a CET map multiplication. However, the values of $\Delta V_{th\,Trans}$ that are used within the phase space may always rely on the same CET map. In this way, the final $\Delta V_{th\,Trans}$ of a phase space interpolation does not correspond to the actual threshold voltage degradation, but the final values of all three abstraction parameters can be used to generate an equivalent occupation (see Section 5.2.3) which in turn is used to calculate the actual threshold voltage degradation via multiplication with the CET map of actual temperature. This procedure is very similar to the NBTI simulation flowchart of Figure 5.19 with the main difference that a transient occupation has to be generated based on the phase space values not only for each change in temperature but also for the final time step in order to estimate the actual threshold voltage degradation (see Figure 5.31). This additional occupation generation has only a marginal impact on the phase space model's accuracy and performance. Due to the temperature independent phase space, the parameters of the stress scenario are processed separately in Figure 5.31. While the temperature variations determine the conversions of occupation probabilities within transient and permanent components, impact of frequency and duty cycle is solely processed within direct CET simulations and phase space interpolations.

The described method of always using the same CET map within the phase space may trigger an inaccuracy if the region with high trap density of actual CET map differs significantly from the region with high trap density of the CET map that is used within the phase space. For instance, an occupation that causes a major shift in threshold voltage in a scenario with high temperature may practically provoke a non-existent threshold voltage degradation when applied to a low temperature CET map. If the low temperature CET map is used within the phase space, occupations with very low capture times cannot be abstracted accurately, since $\Delta V_{th\,Trans}$ is practically always zero. Similar problems arise when an occupation that is caused by a very long relaxation time shall be abstracted with a high temperature CET map, where $\Delta V_{th\,Trans}$ is already fully recovered. To avoid this problem the CET map that is used within the phase space requires non zero trap densities in all relevant capture and emission time regions. Thus, a superposition of CET maps for various combinations of temperature and gate voltage, referred to as reference map, is used within the phase space. The reference map of the 22 nm technology that is used within the following evaluation is presented in Figure 5.32.

In order to further reduce the size of the phase space and to improve accuracy, constant stress or relaxation scenarios are not processed within the phase space (see Figure 5.31). Since these scenarios are calculated effectively with the CET model, there is a negligible impact on phase space model's performance. In fact, the performance may sometimes even be improved, as long time stress or relaxation states are processed with a single operation instead of consecutive operations on the basis of the phase space's time step.

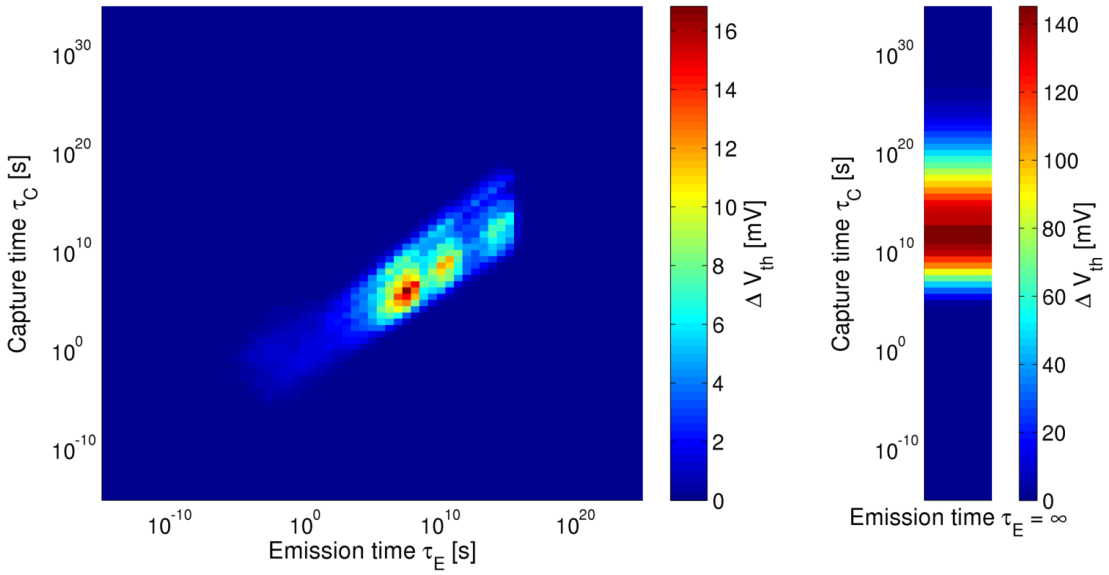


Figure 5.32: Reference CET map of the 22 nm technology as superposition of CET maps for various combinations of temperature and gate voltage.

5.2.6.1 Evaluation

The phase space approach with reference map is the final version of the model for scenarios with constant gate voltage. Therefore, this section contains an extensive evaluation comprising degradation simulations for constant scenarios with different degradation times, frequencies and duty cycles and changing scenarios with varying temperatures, power gating phases and varying activity states. The impact of the permanent component clearly increases for very long degradation times until the permanent component finally represents more than 60% of the overall threshold voltage degradation (see Section 5.2.3.1). Since the permanent component within the phase space approach is calculated in the exact same way as within the reference model, the evaluation has to focus on the transient component. However, due to the usage of the switching trap model as a reference model especially for scenarios with varying temperatures or gate voltages, which are not supported by the published version of the CET model, the evaluation has to be based on the overall threshold voltage degradation. For this reason, a degradation time of 1 month is chosen as maximum degradation time in combination with a phase space time step of 1 minute for the following evaluation scenarios, as this degradation time may already induce a significant degradation, the abstraction procedure is already heavily used by the corresponding 43200 phase space interpolations and the permanent component should not have a significant impact on the overall threshold voltage degradation. In this way, the evaluation with a degradation time of 1 month may analyze the maximal deviation between simulation results of phase space approach and reference models. Furthermore, Figure 5.14 already presents highly accurate evaluation results for accelerated aging scenarios, where the permanent component accounts for 20 to 51% of overall threshold voltage degradation. Very long degradation times are also not feasible within an extensive evaluation due to the very long simulation durations of the reference models.

This evaluation approach is also fostered by the assessment of the physical switching trap NBTI model. Here, silicon measurements are performed for stress times up to approximately 10^6 s (see Figure 2.1) in combination with very high temperatures to trigger severe NBTI degradation. Since these measurement results are accurately simulated by the switching trap model, the model's simple and well-tested temperature dependence is used to accurately predict the NBTI degradation for scenarios with lower temperatures. Within these scenarios, the measured severe NBTI degradation may only occur after extremely long stress times. Hence, it is not expect that the physical NBTI characteristics change for long stress times and the described evaluation approach with a maximum stress time of 1 month is sufficient.

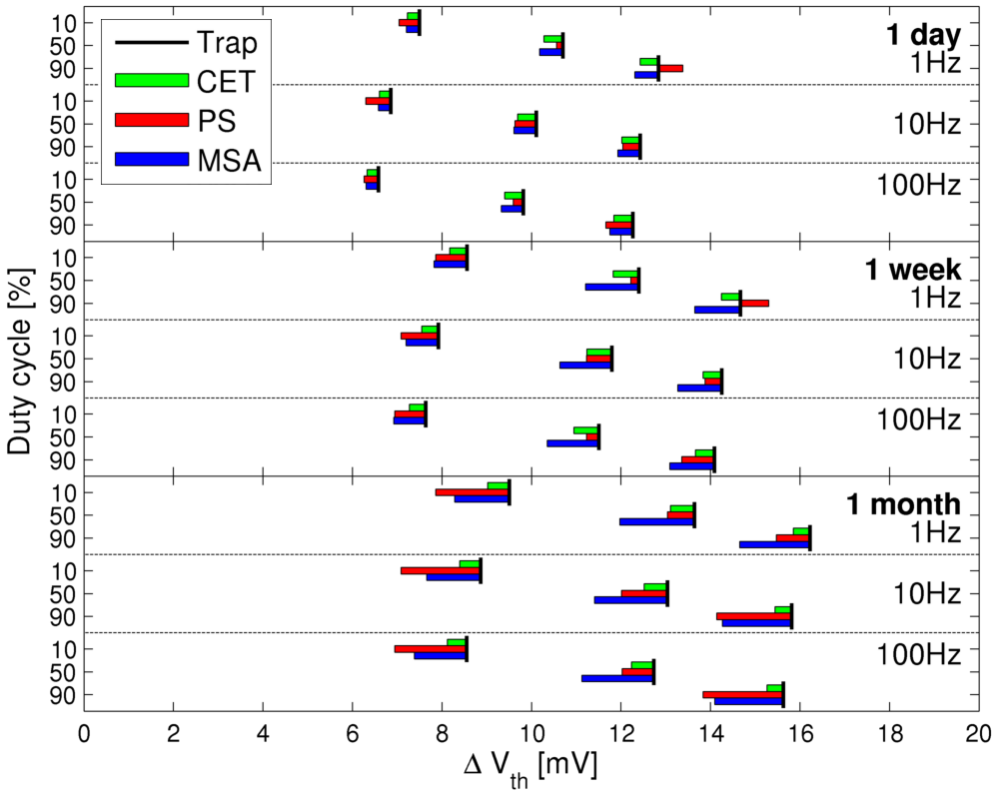


Figure 5.33: Simulated threshold voltage degradations for 22 nm technology and scenarios with 300 K, -0.8 V and different duty cycles (y-axis), stress durations (panels) and frequencies (sub-panels). Short black vertical lines depict simulation results of the switching trap model (Trap) based on provided trap lists (see Section 5.1.1) as reference. Deviations between switching trap simulation results and results of CET model, phase space model (PS) and mission scenario aware analytical approach (MSA) are presented using green, red and blue bars, respectively.

		Duration			Frequency			Duty cycle		
		1d	1w	1m	1Hz	10Hz	100Hz	10%	50%	90%
CET	RMS of rel. error [%]	3.88	4.38	4.16	4.04	4.18	4.21	4.75	4.47	2.99
	Underestimation [%]	100	100	100	100	100	100	100	100	100
PS	RMS of rel. error [%]	5.08	6.71	15.16	8.45	11.23	10.16	15.17	4.65	7
	Underestimation [%]	89	89	100	78	100	100	100	100	78
MSA	RMS of rel. error [%]	4.59	9.54	13.65	9.68	10.04	10.19	10.88	10.78	8
	Underestimation [%]	100	100	100	100	100	100	100	100	100

Table 5.6: Root mean square of relative errors and underestimation percentages of the simulation results presented in Figure 5.33. Data representation corresponds to Table 5.3 and is divided by the different stress scenarios of the associated figure.

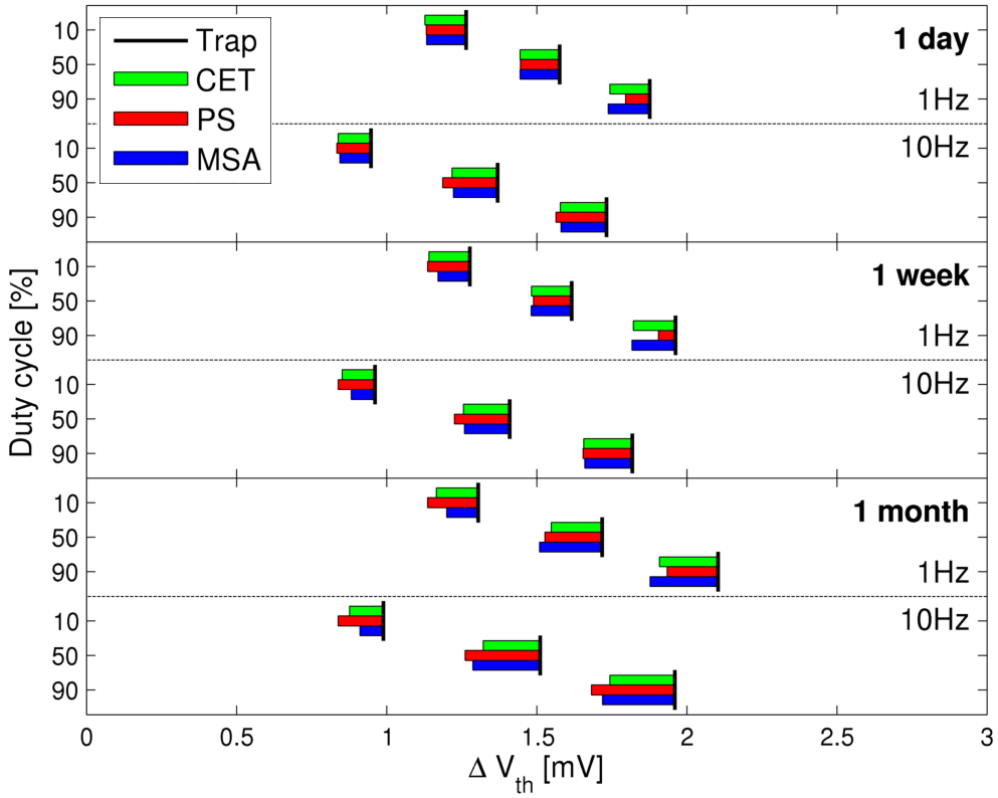


Figure 5.34: Simulated threshold voltage degradations for 130 nm technology and scenarios with 300 K, -1.5 V and different duty cycles (y-axis), stress durations (panels) and frequencies (sub-panels). Representation corresponds to Figure 5.33.

		Duration			Frequency		Duty cycle		
		1d	1w	1m	1Hz	10Hz	10%	50%	90%
CET	RMS of rel. error [%]	10.87	10.77	12.22	10.23	12.28	12.5	11.51	9.72
	Underestimation [%]	100	100	100	100	100	100	100	100
PS	RMS of rel. error [%]	11.41	11.38	15.53	10.17	15.17	14.39	13.95	9.95
	Underestimation [%]	100	100	100	100	100	100	100	100
MSA	RMS of rel. error [%]	10.58	9.5	12.84	10.13	11.93	10.01	12.57	10.45
	Underestimation [%]	100	100	100	100	100	100	100	100

Table 5.7: Root mean square of relative errors and underestimation percentages of the simulation results presented in Figure 5.34. Data representation corresponds to Table 5.6.

Figures 5.33 and 5.34 contain evaluation results for 22 nm and 130 nm technology and several constant scenarios with different duty cycles, stress durations and frequencies. The corresponding relative errors that are averaged for each duration, frequency and duty cycle are presented in Tables 5.6 and 5.7. The MSA approach always uses degradation curves of 1 day for the logarithmic curve fitting within each scenario. Mean relative errors of phase space as well as MSA models are in the range of 10% and the relative error of both models slightly increases with increasing degradation time. However, CET and switching trap models already deviate by approximately 4 and 10% for 22 and 130 nm technologies, respectively. Furthermore, the phase space approach is the only model that partly overestimates the threshold voltage degradation.

The specific power gating scenarios are chosen in a way that steps, as presented in Figure 5.10, should be produced within the occupation probability. The power gating phases are specified based on the probability that 10 minute intervals contain power gating states. In this way, the

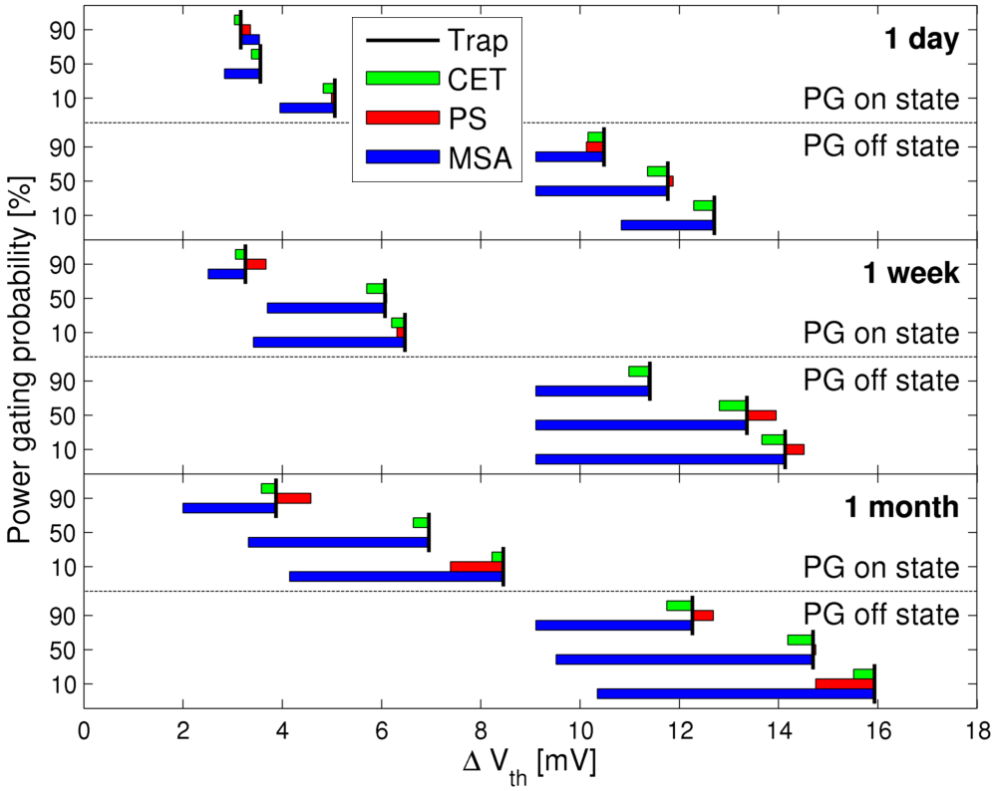


Figure 5.35: Simulated threshold voltage degradations for 22 nm technology and scenarios with 300 K, -0.8 V, 1 Hz, 90% duty cycle and possible power gating states within 10 minute intervals. Power gating probabilities of the intervals, stress durations and power gating states within the final interval are represented by y-axis, panels and sub-panels, respectively. Representation of simulation results corresponds to Figure 5.33.

power gating phases last at least 10 minutes and the number of these phases increases with increasing power gating probability, which partly also results in longer continuous power gating phases. Simulation results are presented in Figures 5.35 and 5.36 as well as Tables 5.8 and 5.9. CET and phase space models have comparable accuracies, but the simulation results of the MSA approach clearly deviate from the reference results. The power gating probability and thus number and durations of power gating phases do not have a distinct effect on the phase space model's accuracy. However, the accuracies of all models are clearly affected by the power gating state within the final 10 minute interval in the way that the relative error is increased for scenarios with final power gating phases. Furthermore, the MSA approach underestimates the degradation for scenarios with final stress phases, whereas the threshold voltage degradation is clearly overestimated for 130 nm technology and scenarios with final power gating phases.

Specific scenarios with varying activity states are chosen in a similar way as the maximum degradation time. The activity states are changed with a high frequency in order to evaluate a high number of state transformations that should normally be part of a mission scenario with much longer degradation times, while having a negligible permanent component at the same time. Therefore, activity states are changed based on the minimal time step for mission scenario transformations. As stated in Section 5.2.1, this corresponds to the phase space time step, which is set to 1 minute within the evaluation examples. The results are presented in Figures 5.37 and 5.38 as well as Tables 5.10 and 5.11. Again, phase space results are considerably more accurate than simulation results of the MSA model. The probability of low activity state and thus number of state transitions, which is considerably higher in the 50% probability state, does not have a distinct effect on the phase space model's accuracy.

The probability that 10 minute intervals contain a high temperature state is used as specification for the scenarios with varying temperatures. The interval of 10 minutes is chosen, since

		Duration			PG state		PG prob.		
		1d	1w	1m	off	on	10%	50%	90%
CET	RMS of rel. error [%]	4.18	4.98	4.81	3.64	5.5	3.67	4.77	5.4
	Underestimation [%]	100	100	100	100	100	100	100	100
PS	RMS of rel. error [%]	2.83	5.17	9.29	3.55	8.25	6.92	1.81	8.36
	Underestimation [%]	67	33	50	44	56	83	33	33
MSA	RMS of rel. error [%]	22.06	56.07	80.26	39.89	71.58	65.61	61.61	44.41
	Underestimation [%]	83	100	100	100	89	100	100	83

Table 5.8: Root mean square of relative errors and underestimation percentages of the simulation results presented in Figure 5.35. Data representation corresponds to Table 5.6.

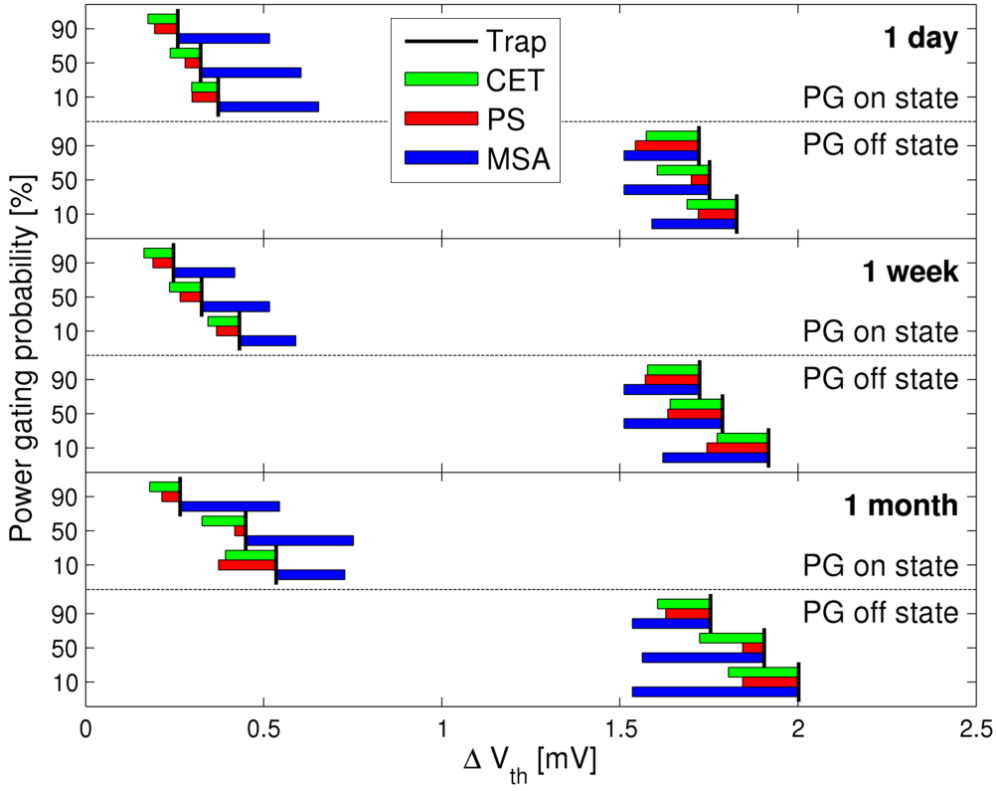


Figure 5.36: Simulated threshold voltage degradations for 130 nm technology and scenarios with 300 K, -1.5 V, 1 Hz, 90% duty cycle and possible power gating states within 10 minute intervals. Representation corresponds to Figure 5.35.

		Duration			PG state		PG prob.		
		1d	1w	1m	off	on	10%	50%	90%
CET	RMS of rel. error [%]	27.23	28.54	29.74	9.36	39.24	21.79	27.22	35.01
	Underestimation [%]	100	100	100	100	100	100	100	100
PS	RMS of rel. error [%]	19.08	18.55	20.96	8.2	26.42	22.29	12.47	22.26
	Underestimation [%]	100	100	100	100	100	100	100	100
MSA	RMS of rel. error [%]	34.57	27.75	33.02	18.69	41.08	28.14	32.17	35.06
	Underestimation [%]	50	50	50	100	0	50	50	50

Table 5.9: Root mean square of relative errors and underestimation percentages of the simulation results presented in Figure 5.36. Data representation corresponds to Table 5.6.

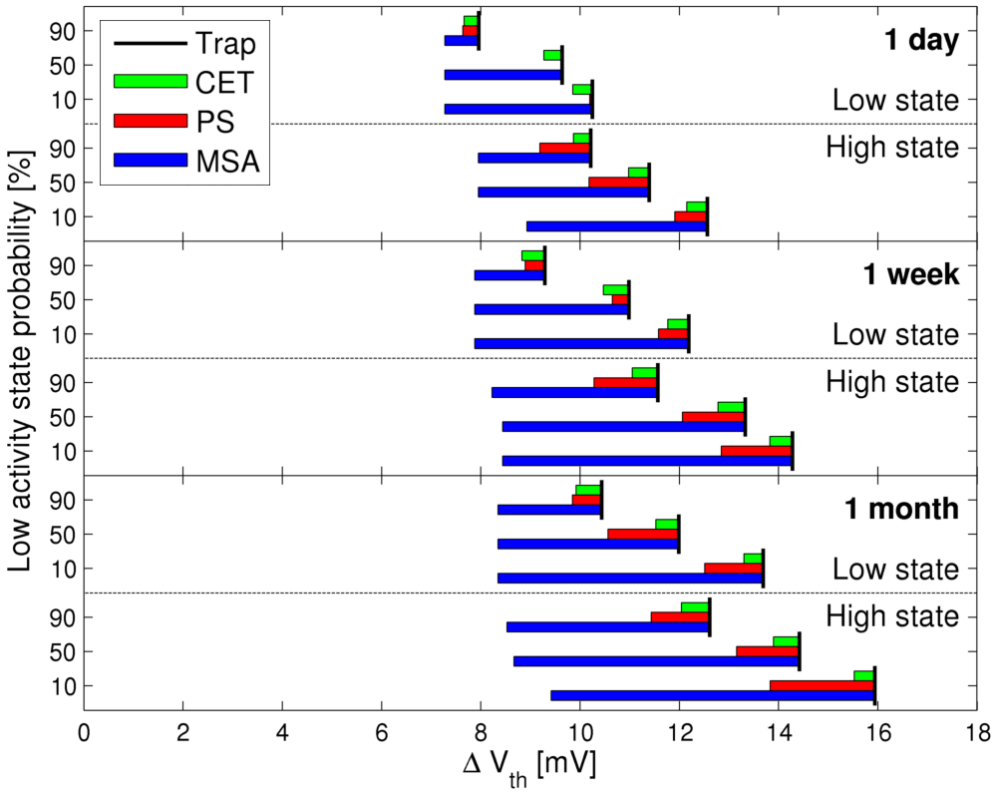


Figure 5.37: Simulated threshold voltage degradations for 22 nm technology and scenarios with 300 K, -0.8 V, 1 Hz and different activity states, covered by 10 and 90% duty cycle, based on 1 minute intervals. Probabilities of low (10% duty cycle) activity state, stress durations and activity state within the final interval are represented by y-axis, panels and sub-panels, respectively. Representation of simulation results corresponds to Figure 5.33.

		Duration			State		Low state prob.		
		1d	1w	1m	Low	High	10%	50%	90%
CET	RMS of rel. error [%]	3.78	4.39	3.98	4.26	3.84	3.34	4.15	4.59
	Underestimation [%]	100	100	100	100	100	100	100	100
PS	RMS of rel. error [%]	7.25	8.64	11.09	6.52	11.15	9.15	9.46	8.77
	Underestimation [%]	83	100	100	89	100	100	83	100
MSA	RMS of rel. error [%]	34.56	49.43	54.93	39.88	53.36	57.73	48.55	31.08
	Underestimation [%]	100	100	100	100	100	100	100	100

Table 5.10: Root mean square of relative errors and underestimation percentages of the simulation results presented in Figure 5.37. Data representation corresponds to Table 5.6.

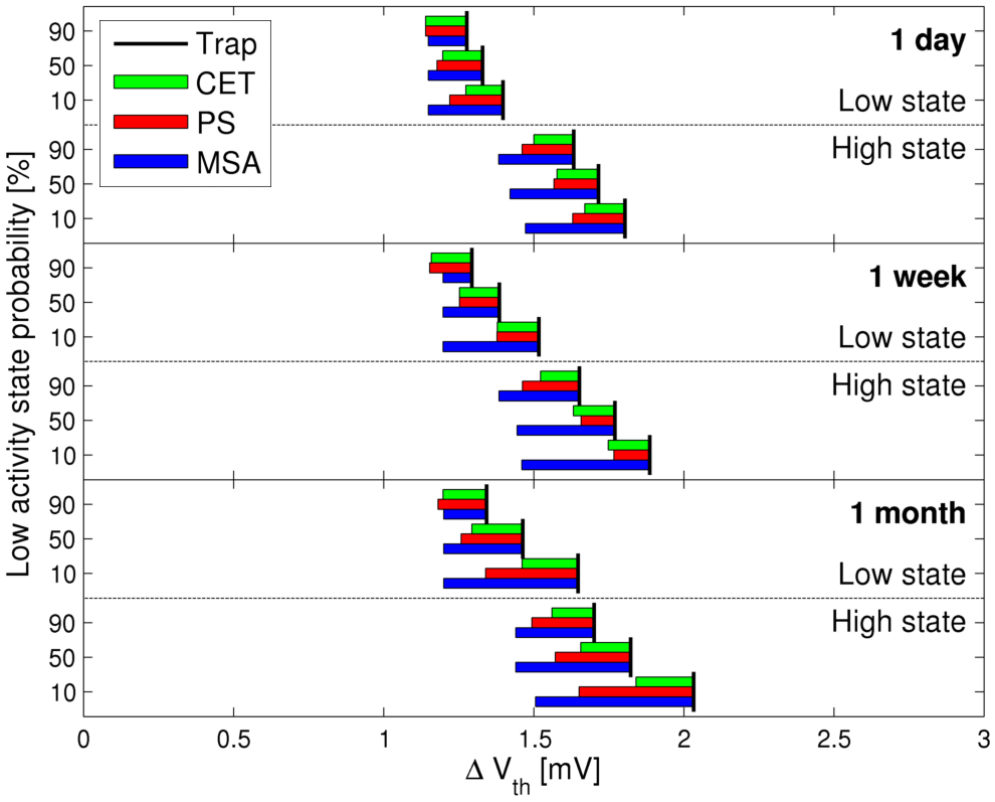


Figure 5.38: Simulated threshold voltage degradations for 130 nm technology and scenarios with 300 K, -1.5 V, 1 Hz and different activity states, covered by 10 and 90% duty cycle, based on 1 minute intervals. Representation corresponds to Figure 5.37.

		Duration			State		Low state prob.		
		1d	1w	1m	Low	High	10%	50%	90%
CET	RMS of rel. error [%]	9.84	9.63	11.36	11.53	8.92	9.95	10.47	10.49
	Underestimation [%]	100	100	100	100	100	100	100	100
PS	RMS of rel. error [%]	12.02	10.21	18.11	14.43	13.28	16.01	12.51	12.81
	Underestimation [%]	100	100	100	100	100	100	100	100
MSA	RMS of rel. error [%]	18.74	21.43	26.69	20.78	24.16	29.31	20.9	15.08
	Underestimation [%]	100	100	100	100	100	100	100	100

Table 5.11: Root mean square of relative errors and underestimation percentages of the simulation results presented in Figure 5.38. Data representation corresponds to Table 5.6.

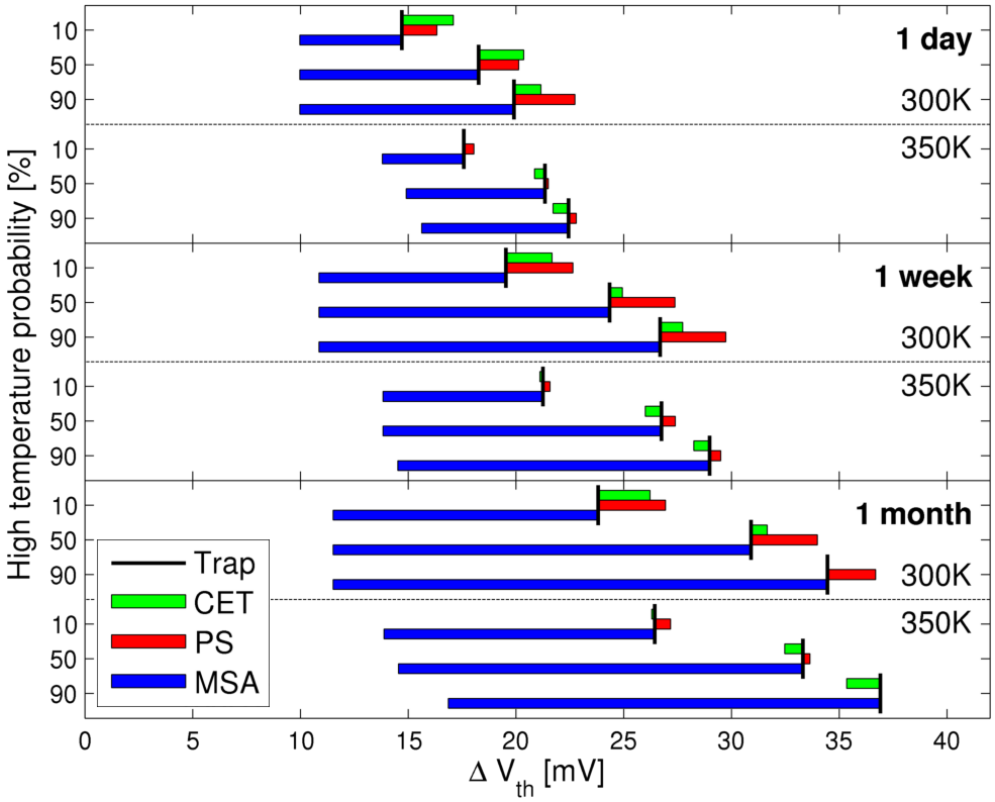


Figure 5.39: Simulated threshold voltage degradations for 22 nm technology and scenarios with -0.8 V, 1 Hz, 50% duty cycle and temperature variations between 300 and 350 K based on 10 minute intervals. Probabilities of high temperature state, stress durations and temperature state within the final interval are represented by y-axis, panels and sub-panels, respectively. Representation of simulation results corresponds to Figure 5.33.

it is unrealistic that distinct temperature transformations occur within short-time intervals. Threshold voltage degradations for the 22 nm technology, as presented in Figures 5.39 and 5.40 and Tables 5.12 and 5.13, show that the enhanced CET approach as well as the phase space model is able to accurately simulate the threshold voltages with mean relative errors in the range of 10% deviation in comparison to the switching trap model. The temperature transformation method slightly overestimates the degradations in the 350 K scenarios and causes small underestimations within scenarios of transitions between 300 and 400 K. The phase space interpolations introduce only minor additional inaccuracies and tend to further overestimate the degradation effect. The probability of high temperature state and thus number of state transitions, which is considerably higher in the 50% probability state, does not have a distinct effect on both models' accuracies. However, the accuracies are clearly affected by the temperature state within the final 10 minute interval in the way that the relative error is increased for scenarios with final 300 K intervals. The MSA approach strongly underestimates the threshold voltage degradations with mean relative errors up to 140 and 437% for the scenarios with high temperature of 350 and 400 K, respectively.

Threshold voltage degradations for the 130 nm technology and scenarios with varying temperatures are presented in Figure 5.41 and Table 5.14. All models clearly underestimate the degradation effect for this technology. Nevertheless, the phase space model is still more accurate than the MSA approach and simulation results of CET and phase space models are nearly identical. The accuracies of all models are affected by the high temperature probability and longer overall durations within the high temperature state cause larger relative errors.

		Duration			Temperature		High temp. prob.		
		1d	1w	1m	300K	350K	10%	50%	90%
CET	RMS of rel. error [%]	7.68	4.73	4.41	7.79	2.55	7.96	4.79	3.8
	Underestimation [%]	50	50	67	11	100	50	50	67
PS	RMS of rel. error [%]	7.65	8.44	6.61	10.6	1.79	8.56	7.03	7.11
	Underestimation [%]	0	0	17	0	11	0	0	17
MSA	RMS of rel. error [%]	62.62	103.7	140.4	125.3	84.97	72.78	113.9	127
	Underestimation [%]	100	100	100	100	100	100	100	100

Table 5.12: Root mean square of relative errors and underestimation percentages of the simulation results presented in Figure 5.39. Data representation corresponds to Table 5.6.

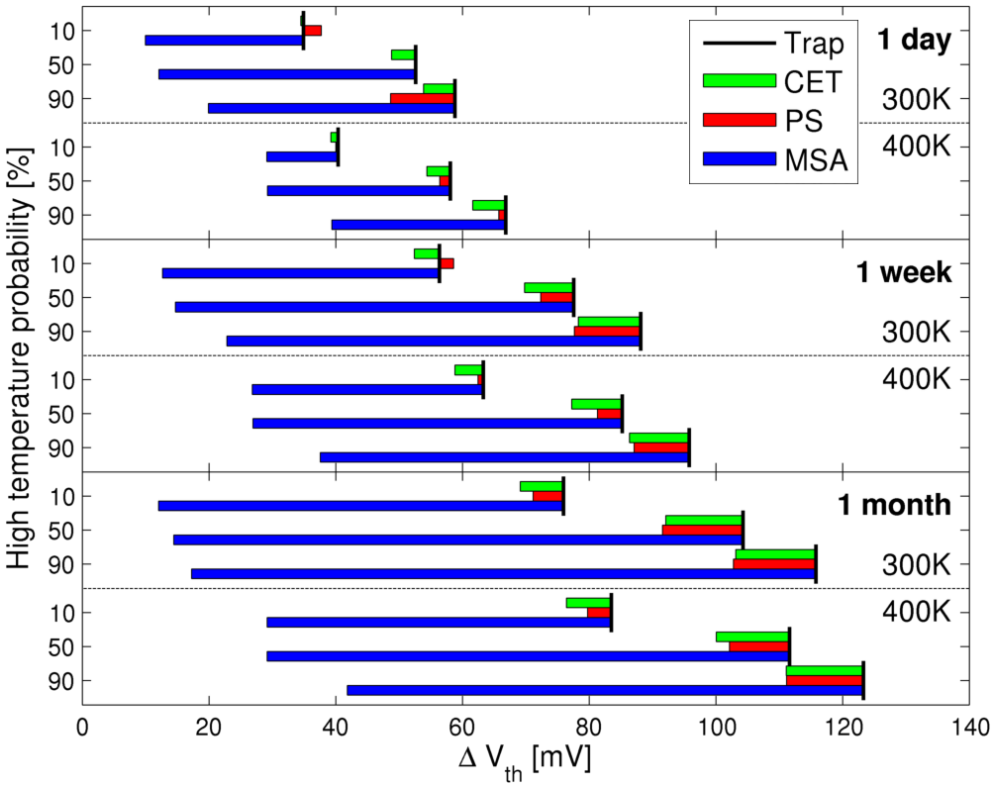


Figure 5.40: Simulated threshold voltage degradations for 22 nm technology and scenarios with -0.8 V, 1 Hz, 50% duty cycle and temperature variations between 300 and 400 K based on 10 minute intervals. Representation corresponds to Figure 5.39.

		Duration			Temperature		High temp. prob.		
		1d	1w	1m	300K	400K	10%	50%	90%
CET	RMS of rel. error [%]	6.73	10.16	11.27	10.03	9.12	7.15	10.37	10.81
	Underestimation [%]	100	100	100	100	100	100	100	100
PS	RMS of rel. error [%]	9.16	7.88	10.22	11.25	6.36	4.81	7.78	12.91
	Underestimation [%]	67	83	100	67	100	67	83	100
MSA	RMS of rel. error [%]	195.4	280.6	436.9	420.2	169.1	293.1	369.2	292.4
	Underestimation [%]	100	100	100	100	100	100	100	100

Table 5.13: Root mean square of relative errors and underestimation percentages of the simulation results presented in Figure 5.40. Data representation corresponds to Table 5.6.

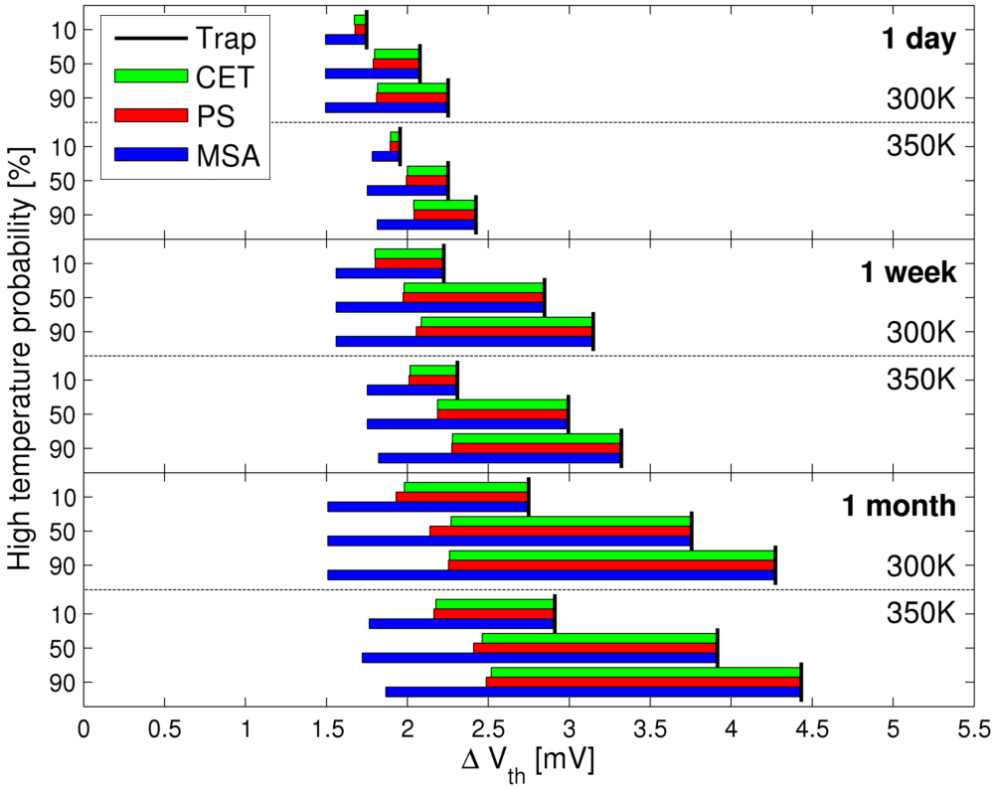


Figure 5.41: Simulated threshold voltage degradations for 130 nm technology and scenarios with -1.5 V, 1 Hz, 50% duty cycle and temperature variations between 300 and 350 K based on 10 minute intervals. Representation corresponds to Figure 5.39.

		Duration			Temperature		High temp. prob.		
		1d	1w	1m	300K	350K	10%	50%	90%
CET	RMS of rel. error [%]	15.06	38.16	63.39	46.68	40.27	23.93	43.7	56.73
	Underestimation [%]	100	100	100	100	100	100	100	100
PS	RMS of rel. error [%]	15.27	38.8	66.77	49.13	41.44	25.12	47.2	57.77
	Underestimation [%]	100	100	100	100	100	100	100	100
MSA	RMS of rel. error [%]	32.7	72.7	130.2	97.58	77.51	48.55	93.52	110.4
	Underestimation [%]	100	100	100	100	100	100	100	100

Table 5.14: Root mean square of relative errors and underestimation percentages of the simulation results presented in Figure 5.41. Data representation corresponds to Table 5.6.

5.2.6.2 Discussion

The enhanced phase space approach of using a reference map does not introduce noticeable additional inaccuracies. This can be seen by comparing the simulation results of the initial phase space approach as presented in Figure 5.17 with the respective results of the enhanced phase space approach (see top panel of Figure 5.33). The evaluation results of the constant long term scenarios indicate that the additional inaccuracy, which is introduced by the phase space approach on top of the CET model's inaccuracy, is approximately in the same range as the deviation between CET and switching trap model. The analytical MSA model that is adapted for each scenario by relying on the respective 1 day degradation curve obtains a similar accuracy as the phase space model. The application of the phase space approach with a time step of 10 minutes should offer degradation simulations of 1 year with superior accuracy in comparison to the evaluation examples of 1 month, since the number of phase space interpolations is roughly equal and the permanent component should have an increased share in overall threshold voltage degradation. As discussed in Section 5.2.3.2, an increased share of the permanent component in overall threshold voltage degradation enhances the accuracy of the phase space approach. Since the relative error of the phase space model's results does slightly increase with increasing degradation time, the maximal deviation between phase space and switching trap results may occur for a degradation time larger than one month. However, the relative error of the phase space model's results should not increase for very long degradation times due to the increased share of the permanent component. Figure 5.14 also supports this conclusion by demonstrating highly accurate phase space simulation results for scenarios that induce significant occupations within the permanent component. Furthermore, the MSA model should not become more accurate than the phase space model with increasing degradation time, since the relative error of the MSA model's results also increases with degradation time and there is no procedure within the MSA model that may change this trend for very long degradation times.

Power gating scenarios that produce steps within the occupation probability do not cause a decline in the phase space model's accuracy. Instead, CET and phase space models have comparable accuracies within these scenarios. While the MSA approach usually clearly underestimates the degradation effect in the power gating scenarios, the relaxation within a final power gating state is underrated, as the model uses a scaling method of a pre-characterized relaxation curve. Longer degradation times within the characterization process of the MSA approach should cause less overall underestimation but also less relaxation within a final power gating state. The high number of activity state transformations in the scenarios with varying activity does not affect the accuracy of the phase space model, which still introduces an additional inaccuracy in the same range as the deviation between CET and switching trap model. On the other hand, the MSA model clearly underestimates the threshold voltage degradation within these scenarios.

Threshold voltage degradations for the 22 nm technology and scenarios with varying temperatures are simulated accurately with the temperature aware CET approach as well as the phase space model. In contrast, the MSA approach considerably underestimates the degradations with mean simulated threshold voltages being over 4 times smaller than the actual values. This behavior of the MSA approach is already identified within short time scenarios and is discussed in Section 5.2.5.4. All models clearly underestimate the degradation effect for the 130 nm technology, whereas the phase space model is still more accurate than the MSA approach. Since longer overall durations within the high temperature state cause considerably larger relative errors, these results may be strongly impaired by the inaccuracies between reference switching trap and CET model for elevated temperatures as discussed in Section 5.2.5.4.

Simulation results of the phase space model are more accurate for scenarios with final high temperature intervals and scenarios without final power gating phases. Fortunately, these scenarios correspond to the use case of mission scenario aware gate level degradation analysis. As described in Section 4.1, the stress scenario may be highly diverse in the course of several years, but a worst case stress scenario is normally considered for the final degradation phase. These scenarios are defined with the intention that the gate level degradation analysis should estimate whether a circuit still operates correctly if the worst case scenario may be utilized for a realistic duration after several years of ordinary usage.

5.2.7 Fourth Iteration of Model Improvements: Voltage Variations

The phase space model, presented so far, does not support mission scenarios with varying gate voltages, which have to be supported for scenarios with dynamic voltage and frequency scaling (DVFS) within the gate level degradation analysis. Extensions of the phase space approach that facilitate simulations with gate voltage variations as well as the corresponding evaluation results are presented within the section. These model improvements constitute the final version of the phase space model that is developed within this thesis. Since scenarios with constant gate voltage are not affected by the fourth iteration of model improvements, the final evaluation for these scenarios is already presented in Section 5.2.6.1.

In contrast to the temperature dependence, the number of defects that contribute to the CET map increases with increasing gate voltage. In this way, CET maps for different stress voltages may have clearly different overall threshold voltage degradations. This is already shown in Table 5.2 on page 57. In particular, the stress voltages that are applied within this example cause a maximum overall $\Delta V_{th\,Trans}$ deviation of 1.12% and 166.09% for 22 nm and 130 nm technologies, respectively. Consequently, a similar method as the temperature conversion technique, which is based on constant overall $\Delta V_{th\,Trans}$ and two-dimensional defect displacements within CET maps for different temperatures, cannot be directly adopted for voltage conversions. Deviations of the maximum overall shift in permanent threshold voltage, that are also stated in Table 5.2, are not relevant in this context, as the already applied conversion technique for the permanent component (see Section 5.2.5.1) is not dependent on the CT maps' maximum overall threshold voltage shifts.

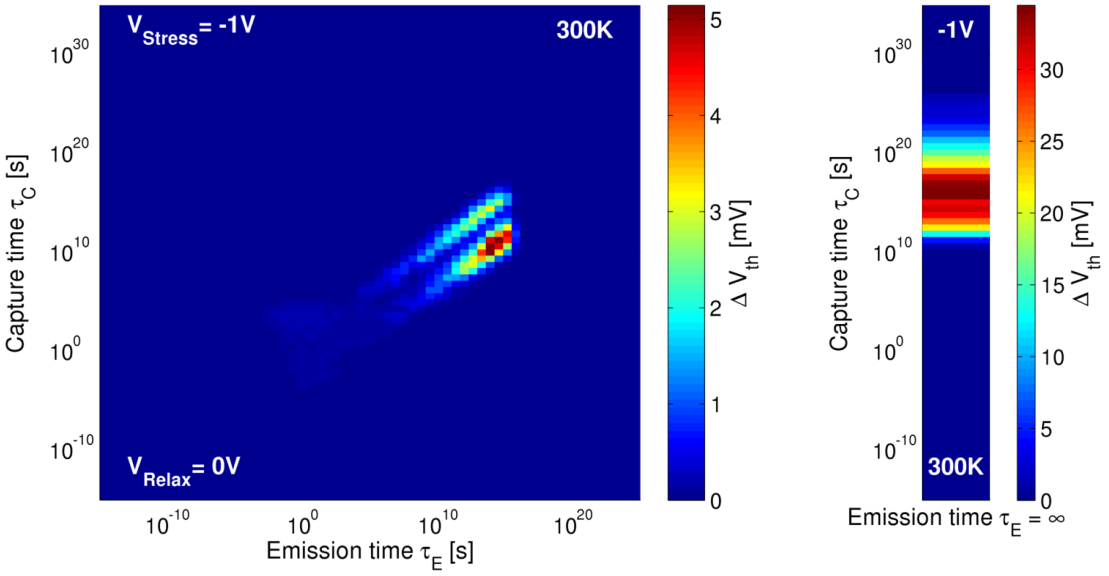
As described in Section 5.2, the phase space model only has to support a small number of different stress voltages due to the specific use case. Therefore, the CET map of a high voltage state may be divided into several CET maps each comprising only traps that are already activated in specific voltage ranges. Separation of the -1 V CET map of 22 nm technology into a CET map of traps that are already activated in the voltage range between 0 and -0.8 V and a CET map of traps that are only activated in the voltage range between -0.8 and -1 V is shown in Figure 5.42. This separation process is implemented within the adapted script of the author of [58] that generates CET maps based on trap lists of the switching trap model. The first CET map (top panels) characterizes all defects that may already be activated at gate voltage -0.8 V and the second CET map (bottom panels) only characterizes additional defects that may be activated at -1 V but cannot be activated at -0.8 V. In this way, a particular CET map is required for every possible gate voltage. This method does not scale with analogue conditions, as theoretically an infinite number of CET maps has to be generated. However, the method is well adapted to model the impact of DVFS conditions, where just a few different stress voltages can be applied. During a NBTI simulation the occupations within the maps for each voltage range have to be tracked separately and the final shift in threshold voltage is calculated by a summation of the threshold voltage shifts for each voltage range. The separation into several voltage ranges does not affect the model's accuracy, but certainly causes a deterioration of the model's performance, which is another reason that only a few different stress voltages can be supported.

As presented for varying temperatures, the individual capture and emission times of the different traps within CET maps for a specific voltage range may also be affected by the gate voltage. Here, a conversion technique for the occupation probability can be applied for transitions between maps for different gate voltages similar to the temperature conversion technique, as different CET maps for the same voltage range certainly share the same overall $\Delta V_{th\,Trans}$. Separation into different voltage ranges as well as gate voltage induced occupation conversions are incorporated within the final NBTI simulation flowchart in Figure 5.43. In order to simulate the overall threshold voltage degradation, a subflow for each voltage range has to be executed and the final degradation is calculated by summation of the threshold voltage degradation for each voltage range. Here, the reference map approach, as presented in Section 5.2.6, is used for the phase space interpolations of each voltage range. In this way, the phase space is independent of voltage ranges as well as gate voltages and the same precomputed phase space can be used for all phase space interpolations. Consequently, the effort of phase space construction does not increase due to support of varying gate voltages.

In order to develop a conversion technique for stress voltage transitions between CET maps

(a) Transient component of range [0 -0.8] V

(b) Permanent component of range [0 -0.8] V



(c) Transient component of range [-0.8 -1] V

(d) Permanent component of range [-0.8 -1] V

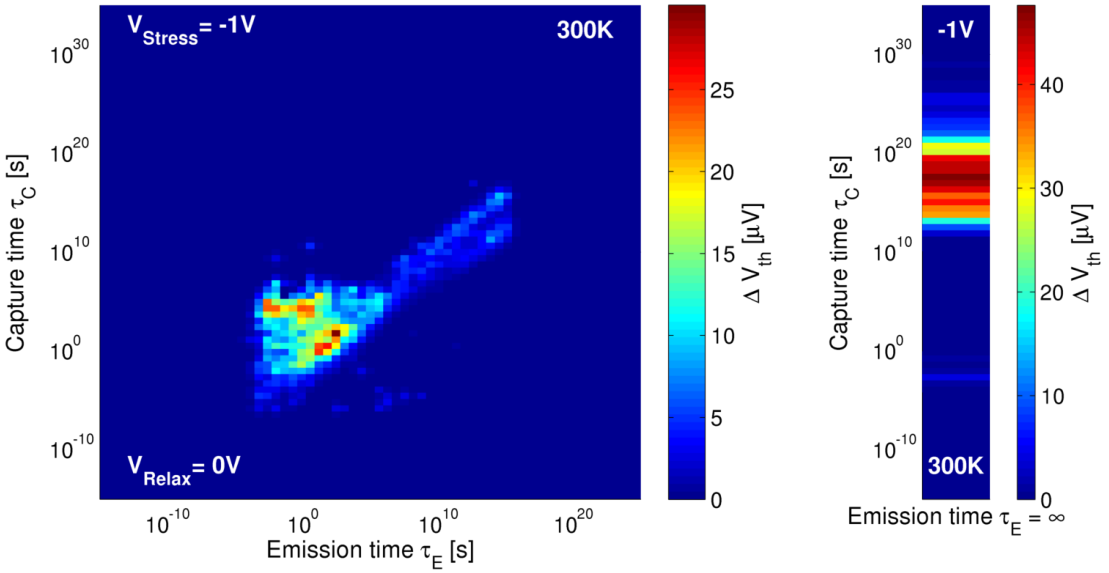


Figure 5.42: Separation of the CET map for 300 K and stress voltage -1 V (see Figure 5.2) into a CET map of traps that are already activated in the voltage range [0 -0.8] V (Panel a and b) and a CET map of traps that are only activated in the voltage range [-0.8 -1] V (Panel c and d).

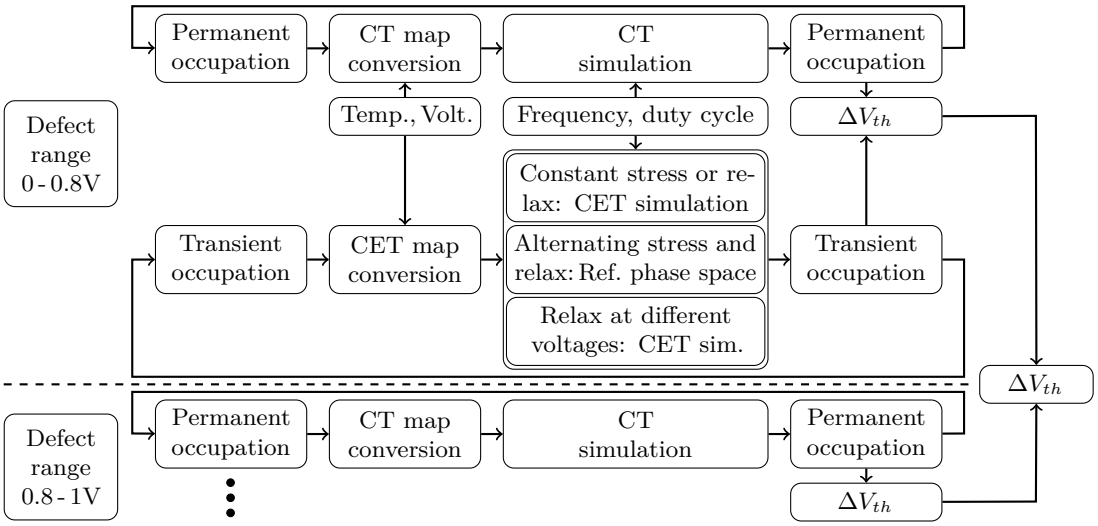


Figure 5.43: Final NBTI simulation flowchart of the phase space model. In order to overcome differences in overall ΔV_{th} of CET maps for different gate voltages due to some traps only being activated at higher voltages, the flowchart consists of subflows for different voltage ranges (divided by a dashed line) and a final summation of the threshold voltage degradations. Occupation conversions and reference map approach are extended in order to support varying temperatures as well as gate voltages. Degradation simulation of the transient component is based on the CET approach for scenarios with constant stress or relaxation, the phase space method of Figure 5.7 for scenarios with alternating stress and relaxation phases or a modified CET approach for voltage ranges that regenerate at both stress and relaxation voltage.

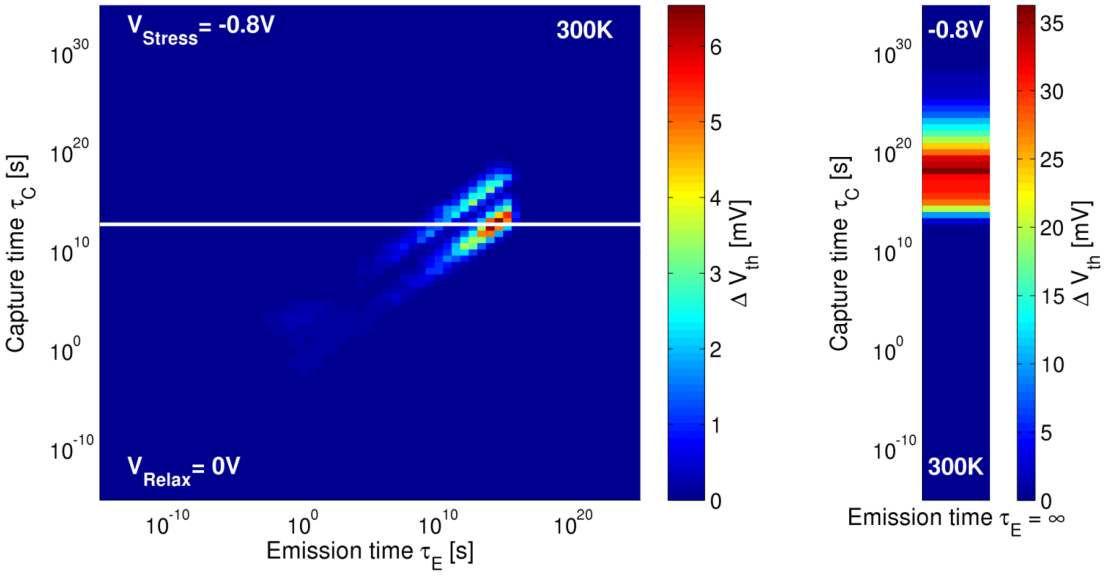
for the same voltage range, the impact of stress voltages on the corresponding CET maps have to be analyzed. Displacements of the defects within CET maps of different stress voltages but same voltage range are shown in Figure 5.44. Naturally, displacements only occur in τ_C dimension, as the relaxation voltage is kept constant. While the capture times of individual defects seem to shift apparently random, higher stress voltages cause lower capture times on average. This is illustrated in Figure 5.44 using horizontal lines that characterize the capture time of maximal threshold voltage shift within the transient components of the different CET maps. As can be seen, the horizontal line for stress voltage -1 V (bottom panel) is significantly lower than within the CET map for -0.8 V (top panel). Logically, the same effect also occurs within the permanent components of these CET maps.

The conversion method for stress voltage transitions within the permanent component is implemented in the same way as the temperature conversion technique. The requirement of unchanged $\Delta V_{th\,Perm}$ during alteration of the CET maps directly produces the occupation probability of the permanent component for the new stress voltage. However, the conversion method for the transient component can presumably not rely on permanent shift in threshold voltage and the three abstraction parameters. This assumption is based on the consideration that the new value of maximum capture time, which is determined based on the converted occupation of the permanent component, does not exactly match the actual mean displacement within the transient component. Due to the requirement of unchanged $\Delta V_{th\,Trans}$, the inaccuracy of maximum capture time produces a displacement of the low emission time border of the occupied region (also described as lower boundary of the slope region in Section 5.2.3). Such a displacement does not occur within the transformation of the CET maps (see Figure 5.44) and may thus cause simulation inaccuracies. In this way, an overestimation of maximum capture time induces a shift of the low emission time border to a higher emission time and the occupation region that is depleted by this displacement, can be reoccupied within the subsequent time step. The simulated additional degradation due to the reoccupation may be in conflict with the actual degradation or relaxation process of this time step. Likewise, an underestimation of maximum capture time triggers a similar behavior.

Instead of the abstraction parameters, the conversion method for stress voltage transitions

(a) Transient component of range [0 -0.8] V

(b) Permanent component of range [0 -0.8] V



(c) Transient component of range [0 -0.8] V

(d) Permanent component of range [0 -0.8] V

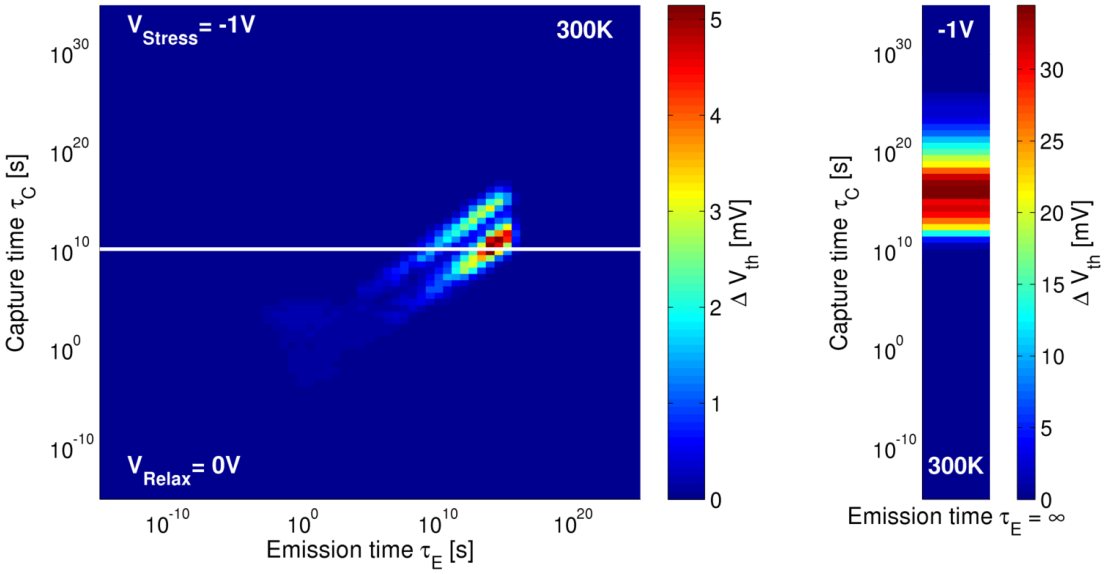


Figure 5.44: CET maps of traps that are already activated in the voltage range [0 -0.8] V (see Panel a and b of Figure 5.42) for different stress voltages. The horizontal line characterizes the capture time of maximal threshold voltage shift within the different CET maps.

within the transient component relies solely on displacements within the τ_C dimension. In order to respect the slope region of the occupation, the displacements are estimated independently for each τ_E value of the CET map. Here, the one-dimensional conversion approach that is used for the permanent component, can be utilized for each τ_E column. Displacements of the low emission time border, as discussed in the last paragraph, can thus not occur within the conversion process. This technique can directly be used for an enhanced CET approach that supports varying stress voltages. However, the converted occupation probability must be characterized by the abstraction parameters within the phase space approach. This may also induce inaccuracies as the converted occupation probability must not comply with the initial characterization of possible occupation probabilities (see Figure 5.10). In particular, the occupation height in the τ_C dimension may vary for different τ_E values in the region of high emission times. The parameter maximum capture time, which is defined as occupation height in τ_C dimension for the highest τ_E value, may thus not precisely characterize the occupation. This problem is solved by averaging the logarithmic occupation height between the emission time of maximal occupation height and the maximal τ_E value. This approach also tends to slightly overestimate the maximum capture time, which causes a displacement of the low emission time border during the abstraction. However, this approach should be more precise than the conversion based on the permanent component.

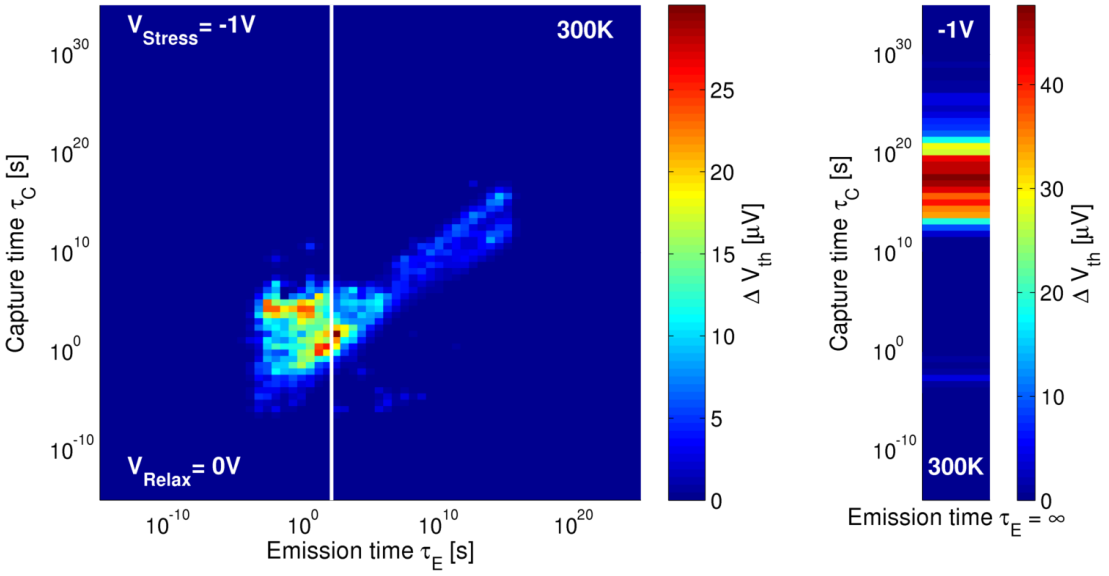
Within a stress scenario, defects of a high voltage range may already be partially occupied and may afterwards be subjected to alternating stress and relaxation phases with reduced stress voltage that cannot provoke additional degradation for this voltage range. Thus, only relaxation can occur during the corresponding interval, but emission times of the defects may be totally different for these two relaxation voltages. Emission times of defects of the voltage range between -0.8 and -1 V are presented for relaxation voltages of 0 and -0.8 V in Figure 5.45. Logically, displacements of defects within these CET maps only occur in τ_E dimension and the permanent component can of course not be affected by varying relaxation voltages. Analogous to the example with different stress voltages, the emission times of individual defects seem to shift apparently random, but higher relaxation voltages cause higher emission times on average, which is illustrated using vertical lines in Figure 5.45.

The described stress scenario with two different relaxation voltages cannot be simulated with a single CET map and can thus not be abstracted with the phase space parameters. Therefore, a different simulation technique has to be developed to support this scenario. Within the final NBTI simulation flowchart in Figure 5.43, this special scenario is specified as “Relax at different voltages”. Since there is only relaxation, the actual alternating high and low voltage stress pattern is irrelevant and the voltage trace can be rearranged to consist of two intervals at the respective voltage values without introducing additional simulation inaccuracies. The regeneration within each interval is directly simulated with the CET approach using CET maps with the corresponding relaxation voltage as presented in Figure 5.45. Each interval is processed within a CET computation of a single time step. In this way, the modified simulation technique has practically no impact on the overall performance of the phase space approach. In between both intervals, the occupation probability, which is based on the CET map of the first interval, has to be converted for the CET map of the latter interval. Additionally, another conversion of the occupation probability with regard to the relaxation voltage may be needed for the subsequent time step of the mission scenario. Furthermore, the order of the intervals for the two relaxation voltages is selected in the way to minimize the overall number of occupation conversions.

The actual conversion of the occupation probability with regard to the relaxation voltage is similar to the conversion technique described above. The displacements within the τ_E dimension are estimated independently for each capture time value of the CET map. In particular, threshold voltage shifts are calculated for each τ_C row using the CET map of the first interval. Starting with a fully occupied state, a virtual emission time is increased independently for each τ_C row as long as each resulting threshold voltage shift for the CET map of the latter interval matches the corresponding pre-computed value. Again, the converted occupation probability must not comply with the initial characterization of possible occupation probabilities. Clear discrepancies can arise for the low emission time border of the occupied region as shown in Figure 5.46a. The τ_E value of the low emission time border for the minimal capture time is larger than the τ_E value for higher capture times. This characteristic cannot be abstracted with the process

(a) Transient component of range [-0.8 -1] V

(b) Permanent component of range [-0.8 -1] V



(c) Transient component of range [-0.8 -1] V

(d) Permanent component of range [-0.8 -1] V

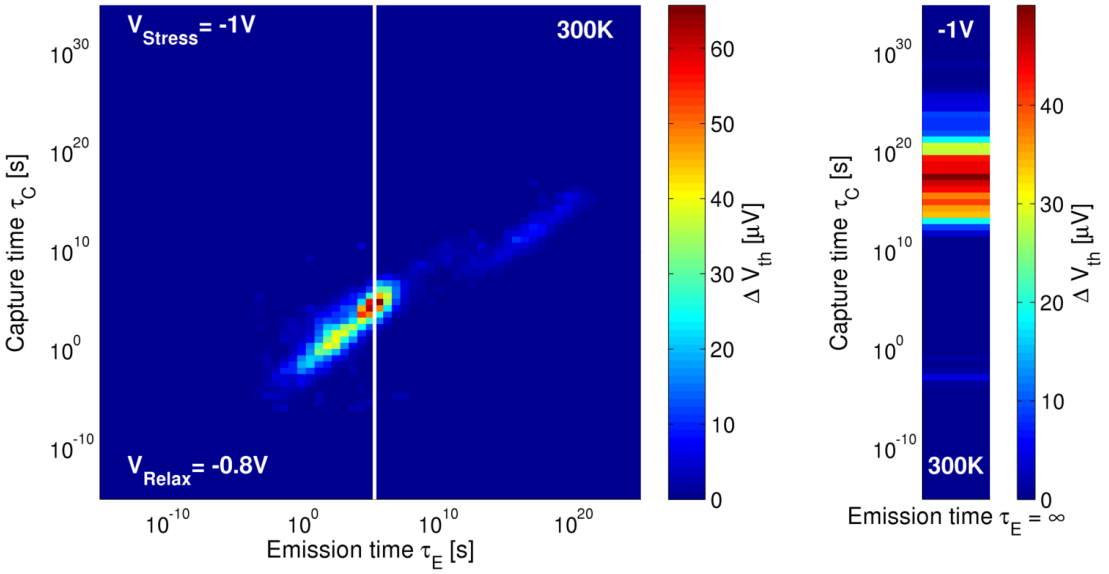


Figure 5.45: CET maps of traps that are only activated in the voltage range [-0.8 -1] V (see Panel c and d of Figure 5.42) for different relaxation voltages. The vertical line characterizes the emission time of maximal threshold voltage shift within the different CET maps.

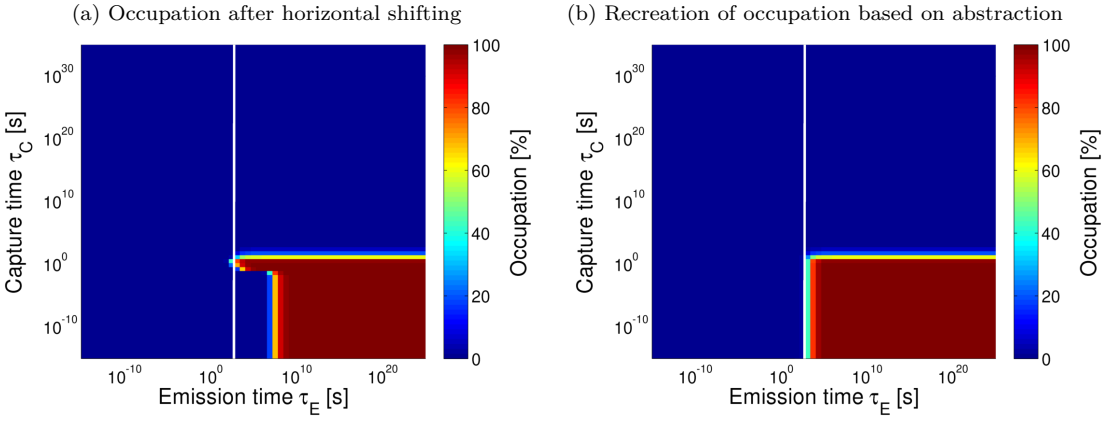


Figure 5.46: Possible occupation probability as a results of relaxation voltage conversion. The shift of the low emission time border due to the abstraction process is indicated with a constant vertical line.

described in Section 5.2.3. Instead of a model modification, this unlikely occupation state is handled in the way that such a state is identified and a fixed minimal slope width is used for the abstraction. This approach slightly shifts the low emission time border to a higher τ_E value within the recreated occupation state as shown in Figure 5.46b. Since this deviation is located in the short time region, it is expected that this simplification has negligible impact on the phase space model's accuracy within long time scenarios.

5.2.7.1 Evaluation

Time traces of simulated threshold voltage degradations for 22 nm technology and short time scenarios with stress voltage transitions are presented in Figures 5.47 and 5.48. Simulation results of the switching trap model (black) always serve as a reference within these figures, since the model directly supports these scenarios. The conversion method for the transient component that relies on the permanent shift in threshold voltage and the three abstraction parameters is used for the CET approach in Figure 5.47. Within the time intervals of low stress voltage, the voltage aware CET model simulates additional threshold voltage degradation that quickly regenerates within the first seconds of the high voltage intervals. This characteristic does clearly not correspond to the reference results of the switching trap model, which specifies a relaxation process within the low voltage intervals and a rapid degradation at the beginning of the high voltage intervals.

In Figure 5.48 the voltage transformation technique of Section 5.2.7 is utilized for the voltage aware CET approach and the phase space model. The simulation results of these two models correspond to the degradation and relaxation sequence of the reference model and the phase space model only introduces minor additional inaccuracies. In conclusion, the transformation approaches are able to support the increasing and decreasing threshold voltage shifts due to voltage variations. Results of the mission scenario aware analytical approach (MSA) demonstrate the time behavior that is already described in Section 3.1.5. Within the low voltage intervals, the MSA simulation results change from a relaxation characteristic to the degradation behavior of the scenario with constant stress voltage of -0.8 V. In addition, the MSA threshold voltage degradation is at the end of the second and third high stress voltage interval always hardly larger than the corresponding value at the end of the previous high voltage interval. Consequently, the simulation results of the voltage transformation enhanced CET model are always in a better agreement with the switching trap model. The phase space approach, which simulates the high side envelope of the CET degradation curve only at the end of each voltage interval, seems to also have the tendency to underestimate the degradation with increasing number of voltage intervals as stated for the MSA approach.

Evaluation results for scenarios with varying stress voltage and longer degradation times are shown in Figures 5.49 and 5.50. Again, the scenarios are defined using a probability of high

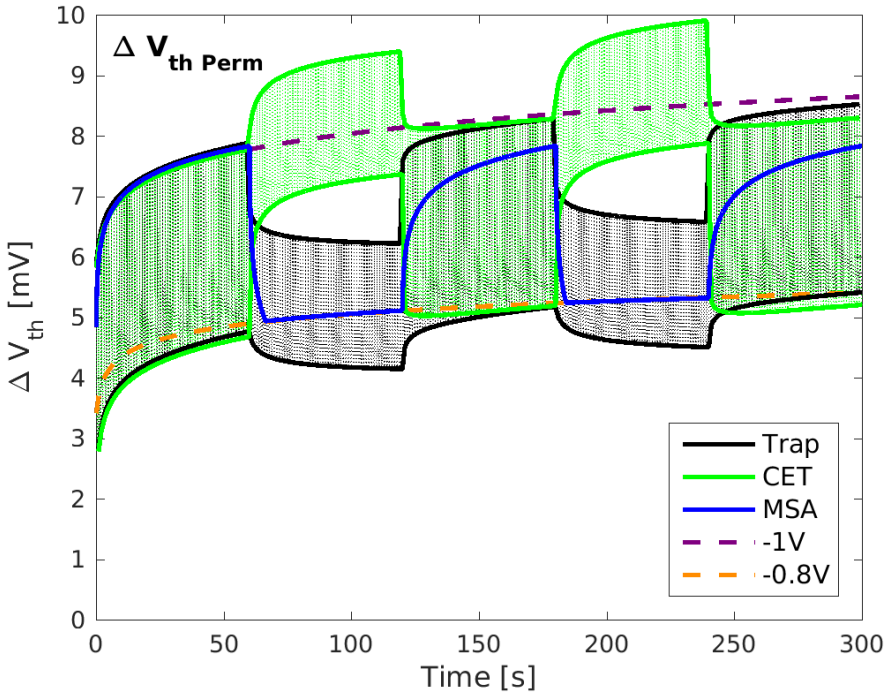


Figure 5.47: Time traces of simulated threshold voltage degradations for 22 nm technology and short time scenarios with 300 K, 1 Hz, 10% duty cycle and transitions between -1 and -0.8 V based on 1 minute intervals. The CET model uses the transformation based on permanent shift in threshold voltage and phase space results are omitted. Representation corresponds to Figure 5.22 on page 61.

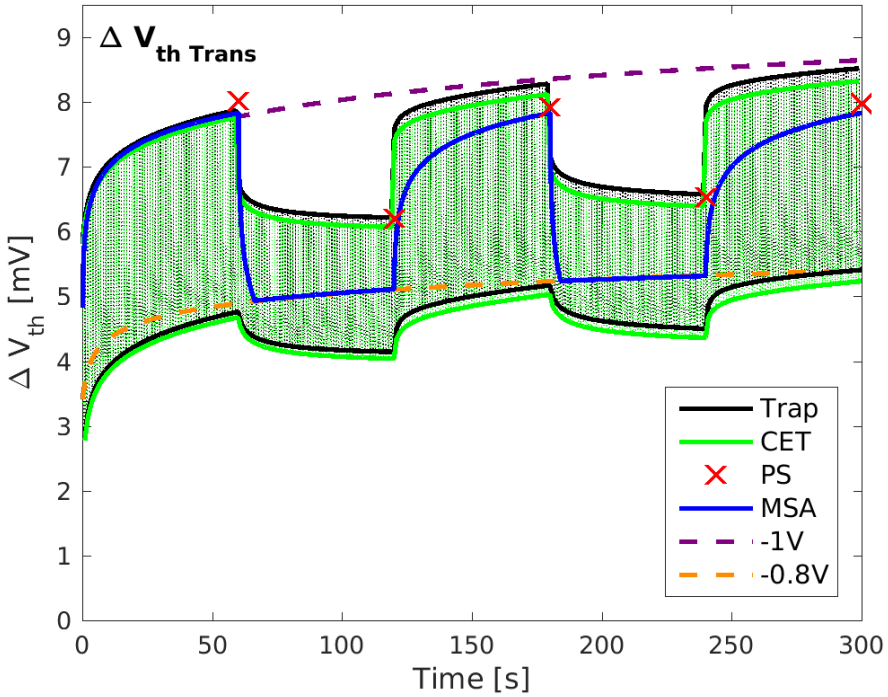


Figure 5.48: Time traces of simulated threshold voltage degradations for 22 nm technology and short time scenarios with 300 K, 1 Hz, 10% duty cycle and transitions between -1 and -0.8 V based on 1 minute intervals. CET and phase space models use the transformation based on transient shift in threshold voltage of individual columns. Representation corresponds to Figure 5.22 on page 61.

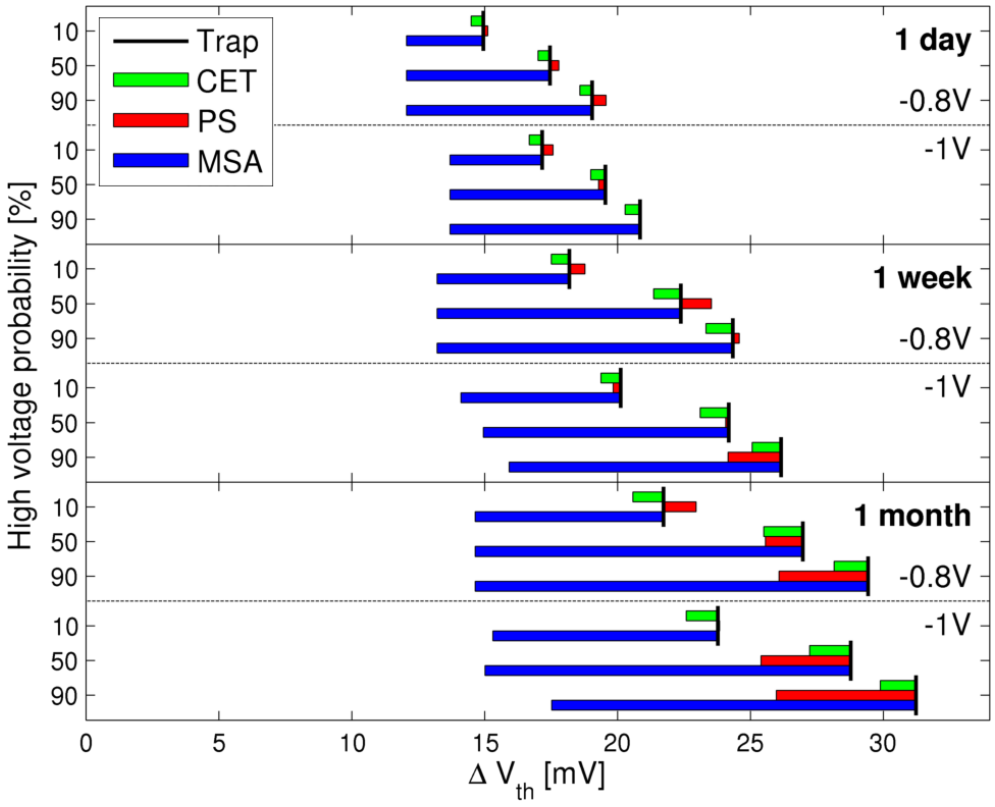


Figure 5.49: Simulated threshold voltage degradations for 22 nm technology and scenarios with 300 K, 1 Hz, 90% duty cycle and voltage variations between -0.8 and -1 V based on 10 minute intervals. Probabilities of high voltage state (-1 V), stress durations and voltage state within the final interval are represented by y-axis, panels and sub-panels, respectively. Short black vertical lines depict simulation results of the switching trap model (Trap) based on provided trap lists (see Section 5.1.1) as reference. Deviations between switching trap simulation results and results of CET model, phase space model (PS) and mission scenario aware analytical approach (MSA) are presented using green, red and blue bars, respectively.

		Duration			Voltage		High volt. prob.		
		1d	1w	1m	-0.8V	-1V	10%	50%	90%
CET	RMS of rel. error [%]	2.8	4.3	5.26	4.27	4.22	4.22	4.58	3.91
	Underestimation [%]	100	100	100	100	100	100	100	100
PS	RMS of rel. error [%]	1.81	4.19	11.62	5.5	8.58	2.81	6.28	10.42
	Underestimation [%]	33	50	67	22	78	17	67	67
MSA	RMS of rel. error [%]	43.04	62	78.72	65.68	60.12	40.51	68.22	74.82
	Underestimation [%]	100	100	100	100	100	100	100	100

Table 5.15: Root mean square of relative errors and underestimation percentages of the simulation results presented in Figure 5.49. Data representation corresponds to Table 5.6.

voltage state based on 10 minute intervals and all four models are always served with the same voltage variation characteristic within a single scenario. Both CET and phase space models use the described voltage transformation method, however only the phase space approach comprises the parameter abstractions. Tables 5.15 and 5.16 specify the root mean square of the relative errors in relation to the simulation results of switching trap model for each model, final voltage state and high voltage probability.

The CET model with additional voltage conversion method accurately simulates the threshold

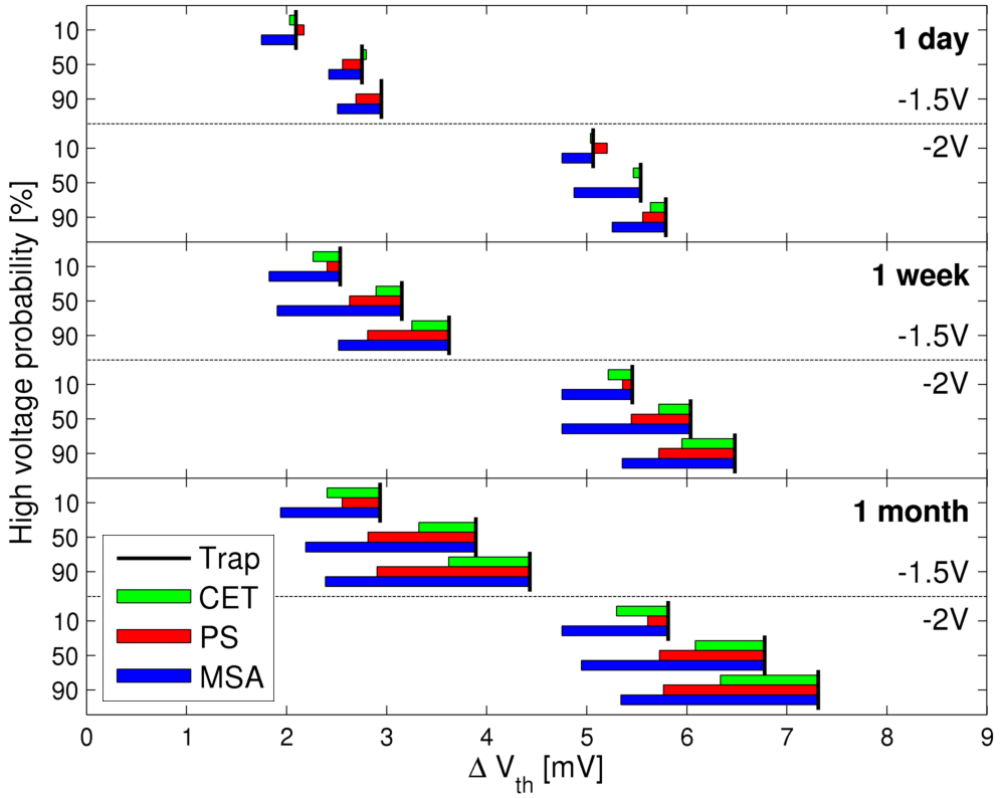


Figure 5.50: Simulated threshold voltage degradations for 130 nm technology and scenarios with 300 K, 1 Hz, 90% duty cycle and voltage variations between -1.5 and -2 V based on 10 minute intervals. Representation corresponds to Figure 5.49.

		Duration			Voltage		High volt. prob.		
		1d	1w	1m	-1.5V	-2V	10%	50%	90%
CET	RMS of rel. error [%]	1.91	8.98	17	13.49	8.19	11.18	9.46	12.61
	Underestimation [%]	83	100	100	89	100	100	83	100
PS	RMS of rel. error [%]	5.6	16.13	30.32	25.55	12.44	6.89	19.89	27.71
	Underestimation [%]	67	100	100	89	89	67	100	100
MSA	RMS of rel. error [%]	14.25	38.98	56.56	52.25	23.47	29.66	46.17	43.69
	Underestimation [%]	100	100	100	100	100	100	100	100

Table 5.16: Root mean square of relative errors and underestimation percentages of the simulation results presented in Figure 5.50. Data representation corresponds to Table 5.6.

voltage degradation for the 22 and 130 nm technologies. Relative errors of the simulation results slightly increase with increasing degradation time, but the probability of high voltage state and thus number of state transitions, which is considerably higher in the 50% probability state, does not have a distinct effect on the model's accuracy. Furthermore, the accuracy seems to not be directly affected by the final voltage state. Likewise, simulation results of the phase space model, which comprises the same voltage conversion approach, feature the same characteristics with slightly higher relative errors and an increased tendency to overestimate the degradation effect. The MSA approach always clearly underestimates the threshold voltage degradation, although the discrepancies between MSA and phase space results are significantly smaller for the 130 nm technology.

Finally, scenarios with possible power gating states as well as varying temperatures and stress voltages are used in order to evaluate the interaction of different influencing variables. In

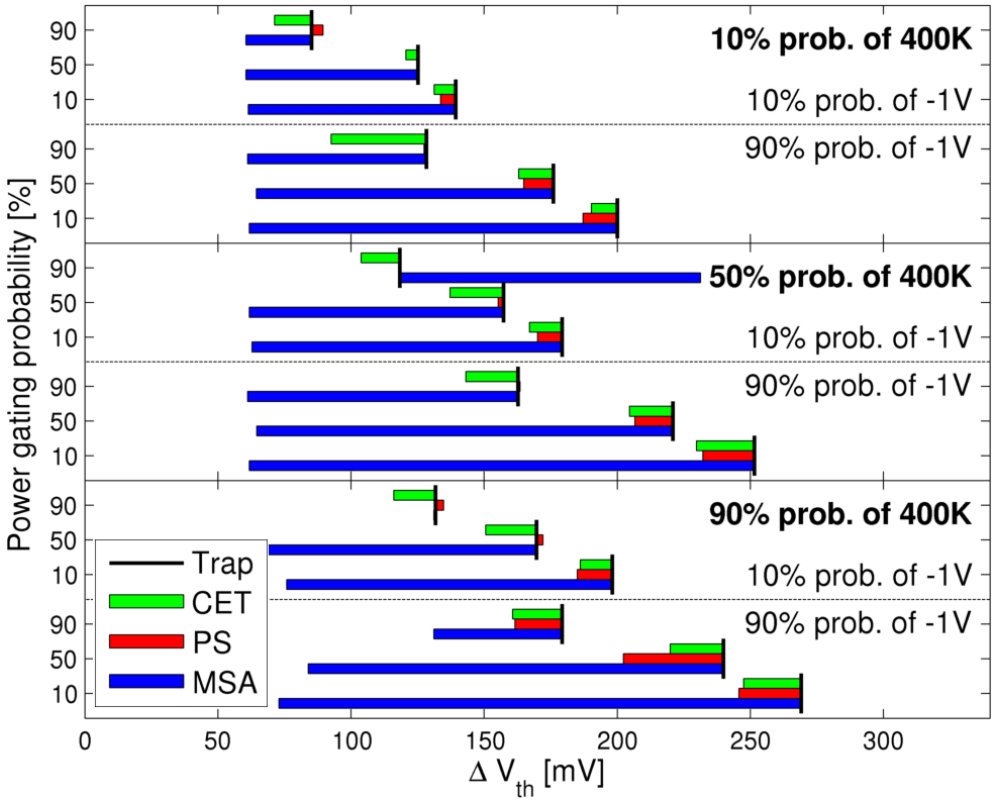


Figure 5.51: Simulated threshold voltage degradations for 22 nm technology and scenarios with 1 month stress duration, 1 Hz, 90% duty cycle, possible power gating states, temperature variations between 300 and 400 K and voltage variations between -0.8 and -1 V. All varying parameters rely on 10 minute intervals and probabilities of the different states of the parameters are independent of each other. Power gating probabilities of the intervals, probabilities of 400 K temperature state and probabilities of -1 V state are represented by y-axis, panels and sub-panels, respectively. Representation of simulation results corresponds to Figure 5.49.

		High temp. prob.			High volt. prob.		PG prob.		
		10%	50%	90%	10%	90%	10%	50%	90%
CET	RMS of rel. error [%]	18.35	11.58	10.6	11.93	15.7	7.36	9.98	20.72
	Underestimation [%]	100	100	100	100	100	100	100	100
PS	RMS of rel. error [%]	4.77	4.97	10.11	3.81	9.23	7.13	8.59	5
	Underestimation [%]	83	67	67	56	89	100	83	33
MSA	RMS of rel. error [%]	142.2	200.4	163.5	125.7	205.6	221	175.4	86.58
	Underestimation [%]	100	83	100	89	100	100	100	83

Table 5.17: Root mean square of relative errors and underestimation percentages of the simulation results presented in Figure 5.51. Data representation corresponds to Table 5.6.

particular, relative errors due to power gating, varying temperatures and varying stress voltages may add up. Evaluation results for corresponding scenarios with a degradation time of one month are presented in Figure 5.51 and Table 5.17. Here, a maximal degradation state with 400 K and -1 V is always used within the final interval.

Overall relative errors of the phase space approach in Figure 5.51 are always smaller than 10%. Furthermore, the accuracy of the phase space approach is even higher than the CET model's accuracy. A comparison with Tables 5.8, 5.13 and 5.15 shows that the relative error does not increase significantly due to simultaneous variation of several influencing variables.

The permanent component accounts for 11.4% of the overall threshold voltage degradation on average with minimal and maximal values of 0.6 and 27.1%, respectively. On the contrary, overall relative errors of the MSA approach are in the range of 86 to 221%. While the model tends to significantly underestimate the degradation effect in general, there are also examples with stress traces that cause clear deviations in relaxation phases resulting in an overestimation of the degradation effect.

5.2.7.2 Discussion

Scenarios with varying voltages due to DVFS can be supported by having separate phase space traces for different voltage ranges and using a conversion technique for the occupation probabilities. The conversion technique based on phase space abstraction parameters and transformations within the permanent component as utilized for temperature conversions cannot be used for stress voltage transformations, since simulated degradation traces may feature additional degradation within intervals that should actually contain relaxation. However, new one-dimensional conversion methods for the τ_C and τ_E dimension result in an accurate threshold voltage trace of a short time scenario and precise long time evaluation results. Support of varying voltages within the phase space approach has no impact on the phase space construction, as the same reference phase space is used for all interpolations, but causes a deterioration of the model's performance due to several phase space traces. Therefore, the presented voltage aware phase space approach is only applicable for scenarios with few different gate voltages and cannot be used for analogue conditions. The regeneration of defects of a higher voltage range at a lower stress voltage may be neglected for particular technologies or voltage ranges due to very high emission times. As a result, the model's performance is increased as the number of conversions and CET simulations is reduced and simulation results are slightly shifted to higher degradation voltages.

As relative errors due to different influencing variables do not add up in combined scenarios, it is not expected that the phase space model's accuracy decreases for realistic scenarios within the gate level degradation analysis. Furthermore, the model's accuracy should also increase with increasing share of the permanent component on overall threshold voltage degradation as stated before. In conclusion, the phase space model shows significantly improved accuracies within NBTI simulations in comparison to the MSA approach and should also offer a sufficient performance for the use case of gate level degradation analysis.

5.3 Phase Space Model of Single Transistor HCD Degradation

As stated in Section 3.2, HCD can also be characterized using CET maps. Major difference between HCD and NBTI degradation is the different stress condition. NBTI occurs when a transistor is in inversion and the gate terminal is negatively biased. On the other hand, HCD is generally believed to be caused by a considerable current flow in the channel [5]. By adjusting the stress conditions, the phase space approach can also directly be used for the HCD degradation mechanism. Since the emission times of HCD seem to be very high, a characterization with capture time (CT) map only may also be sufficient for realistic mission scenarios. Furthermore, recent measurements indicate that NBTI and HCD are independent of one another (see Section 3.2). In this way, HCD may be modeled similar to the permanent component of NBTI in a combined simulation flow of both effects. However, it was not possible to get characterization data for HCD degradation of a recent technology and it was thus unfortunately impossible to evaluate a phase space HCD model within this thesis.

5.4 Summary

A new NBTI modeling technique is proposed that is highly adapted to the use case of mission scenario aware gate level degradation analysis. Main idea is a lookup table approach to offer NBTI simulations with high performance and to support varying mission scenarios at the same

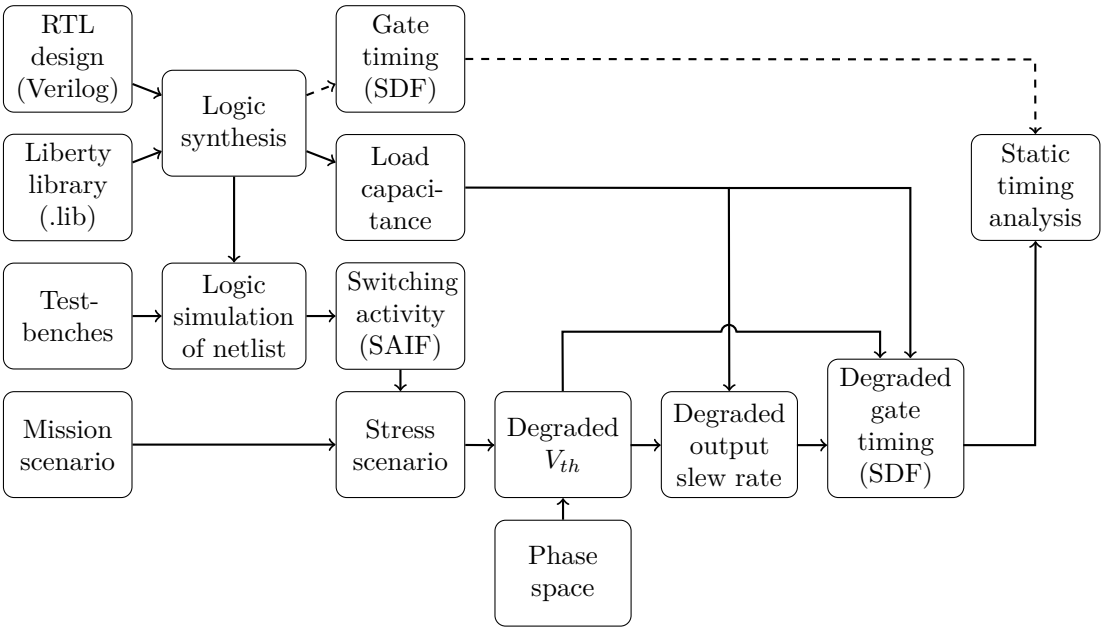


Figure 5.52: Proposed gate level degradation analysis flow extended by the phase space as an additional input argument for the estimation of threshold voltage degradation

time. The lookup table, which is called phase space, is based on an abstraction of the degradation state with 3 parameters and comprises pre-calculated transformations of the degradation state, which are calculated with a reference NBTI model for different scenarios. Since the CET model offers effective abstraction potentials it is used as reference model in the final version of the phase space model, although the CET model does not directly support varying temperatures nor gate voltages. Thus, additional modeling approaches are developed to facilitate CET as well as phase space simulations with varying temperatures and gate voltages as specified within the stress scenarios of the gate level degradation analysis. These additional modeling approaches cause a slight deterioration of the phase space model's performance, but have no impact on the generation of the phase space, which is constructed in a temperature and voltage independent manner. The extended gate level degradation analysis flow, which utilizes the phase space as an additional input argument for the estimation of threshold voltage degradation, is presented in Figure 5.52.

Within an extensive evaluation it is shown that the deviation between results of the phase space model and the reference switching trap model is almost always smaller than 15%. In many cases, the additional inaccuracy introduced by the phase space approach is in the same range as the inaccuracy that is caused by the CET approach. Furthermore, the model's accuracy should also increase with increasing share of the permanent component on overall threshold voltage degradation, which occurs for very long degradation times. Unfortunately, it was not possible to use these long degradation times within the extensive evaluation due to the performance of the reference model. However, evaluation examples for accelerated aging scenarios that also cause an increased share of the permanent component result in an increased accuracy of the phase space model. The mission scenario aware analytical (MSA) approach, which is the only other published model fulfilling the requirements of the gate level degradation analysis, shows significantly larger deviations to the reference results for all evaluation examples. In particular, mean deviations up to 400% arise for scenarios with varying temperatures.

The MSA model, as implemented within this thesis, uses model parameters that are directly specified for each scenario and thus does not comprise inaccuracies due to suboptimal parameters. On the other hand, the phase space model uses a direct recreation and CET simulation procedure for the emerging values of the abstraction parameters within the evaluation. In this way, the phase space results do not incorporate interpolation inaccuracies that are determined by the phase space's resolution. Since these additional inaccuracies of both models are not an integrated part of the respective modeling technique, it was decided to omit these inaccuracies and directly

evaluate the fundamental modeling techniques. In this way, the evaluation results of the phase space model cannot be used to assess the model's performance. Therefore, an assessment of the performance of the phase space model is stated in Chapter 7.

CHAPTER 6

Timing Degradation of CMOS Logic Gates

Realistic threshold voltage shifts of transistors within CMOS logic gates can already be calculated efficiently for varying mission scenarios using the transistors' unique stress scenarios as characterized by the corresponding XML files (see Section 4.1) and calculation techniques that are presented in Chapter 5. However, the main goal of the gate level degradation analysis is the estimation of the degraded timing of complete RT components. As stated in Section 4.2, this is done by generating SDF files with degraded gate delays for specific mission scenarios in order to use standard static timing analysis tools for the timing assessment of the complete RT component. Therefore, a method to calculate the degraded timings of CMOS logic gates based on the estimated threshold voltage shifts is required within the gate level degradation analysis. Additional requirements for this method are the support of worst case timing conditions as well as realistic conditions that are specified by the mission scenario and high performance demands. Such a timing estimation method, relying on well-known concepts that are presented in Section 2.2, is proposed and evaluated within this chapter.

6.1 Basic Concept of the Gate Timing Estimation

Based on calculated threshold voltage shifts of a gate's transistors, a signal's delay between input and output pin of a CMOS gate as well as slew rate of the output signal can directly be computed using SPICE. Here, the discrete factors type of the CMOS logic gate, respective input pin and type of output transition as well as the non discrete factors load capacitance, input slew rate, temperature, supply voltage and process variation node need to be also considered as stated in Section 2.2.1. These SPICE simulations can though not be executed during gate level degradation analysis due to performance constraints. Instead, the basic gate level liberty library approach that permits fast timing simulations by relying on pre-characterized lookup tables for the gate's delay and output slew rate is applied for the proposed method. Numerous SPICE simulations are thus performed for the gates of a particular technology in order to estimate the respective delay and output slew rate values. In particular, these simulations have to be performed for various combinations of the determining factors stated above. Similar to the liberty library approach, these values are stored within a database that is the foundation of the timing estimation within the gate level degradation analysis.

For non discrete determining factors reasonable sampling points have to be chosen. In contrast to the liberty library approach, the proposed method not only uses sampling points for load capacitance and input slew rate, but also incorporates several sampling points for the determining factors temperature and supply voltage in order to directly support timing estimations for different mission scenarios. In this regard, the liberty library approach is specifically designed for worst case assumptions and has to rely on different liberty library files for timing estimations at diverse temperatures as well as supply voltages. Furthermore, the impact of process variations can only be assessed based on different databases by both timing estimation methods. However, dependencies of degraded timings on process variation nodes are not studied within this thesis.

Main distinction between proposed method and liberty library approach is though the consideration of the additional determining factor threshold voltage shift of the gate's transistors. Since the threshold voltage shifts of a different number of transistors have to be considered for single and multiple-stage gates, the proposed method has to be also adapted for different kinds of gates. In order to reduce the size of the database, delay and output slew rate values of the sampling points of this non discrete determining factor are not directly stored within the database. Instead, parameters of regression curves of the respective delay and output slew rate values are used. This approach corresponds to linear delay regression techniques of some published gate level analysis methods [16, 17] that are described in Section 3.3.3. However, as the regression parameters of different gates are not specified within these publications, SPICE measurements for the characterization of regression parameters are needed nonetheless. In this way, an evaluation of the necessary simplifications and regression techniques, which is not incorporated within the stated publications, is also carried out.

During timing estimations of the gate level degradation analysis, delay and output slew rate values have to be estimated for a distinct combination of all determining factors. While all values of discrete factors are directly included within the database, non degraded delays and output

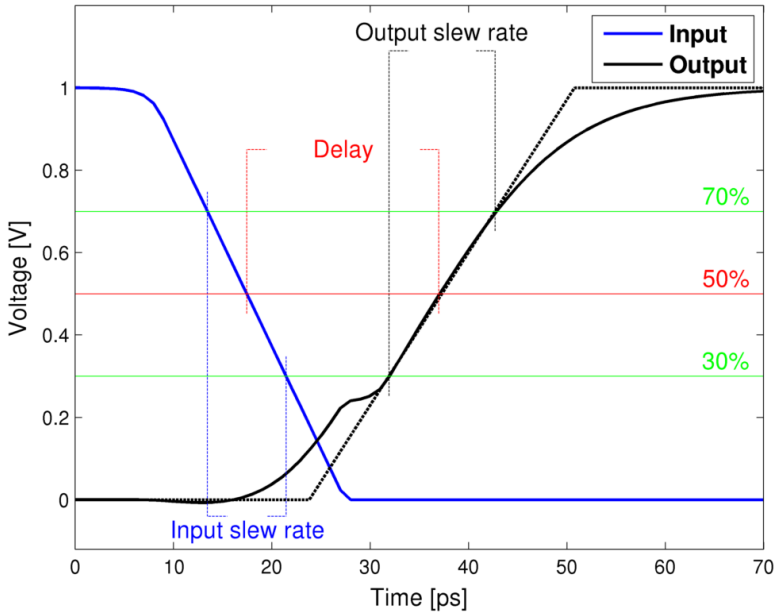


Figure 6.1: Primary SPICE measurement results of a rising output transition of an inverter for load capacitance 40 fF, input slew rate 8 ps and 298 K. Additional markings illustrate the gate's delay and slew rate definitions and the black dotted line denotes the linear regression of the output signal solely based on the output slew rate.

slew rates as well as parameters of the regression curve corresponding to required values of non discrete factors have to be interpolated on the grounds of utilized sampling points. Subsequently, actual degraded delays and output slew rates are calculated based on non degraded values and interpolated threshold voltage regression curves.

6.2 SPICE Measurements and Abstraction of Degradation Behavior

Within this section, SPICE measurements of timing degradations for different types of gates are presented and the threshold voltage regression techniques are evaluated. Measurements are based on the 45 nm NanGate „Free Process Design Kit“ and „Open Cell Library“ [25], since the 45 nm technology node is the latest technology with a complete and freely available set of required information and tools. The presented measurement data and regression techniques of different types of gates exemplify the process of building a complete database of the degraded timings of a recent technology node. Thus, measurements are performed for every state of discrete factors and reasonable sampling points are chosen for non discrete factors. In particular, load capacitances and input slew rates that are stated within the liberty library of the NanGate „Open Cell Library“ as well as different temperatures (273, 298, 323, 363 and 398 K) and supply voltages (0.95, 1.1 and 1.25 V) are used for the construction of the timing database.

Typical primary SPICE measurement results are shown in Figure 6.1. Within this example, a changing input signal at the input pin of an inverter causes a rising output transition. As illustrated in Figure 6.1, the gate's delay is characterized by the time lag between the 50% points of both signals and input and output slew rates are specified by the time lag between the 30% and 70% points of input and output signal, respectively. These secondary measurement results are solely used within the following figures. Input signal generation within SPICE is by default based on linear ramps that are specified by a slew rate value. In this way, separate SPICE simulations for two successive gates introduce minor inaccuracies due to the usage of the linear regression of the first gate's output signal as input signal of the subsequent gate. This is indicated by the black dotted line in Figure 6.1.

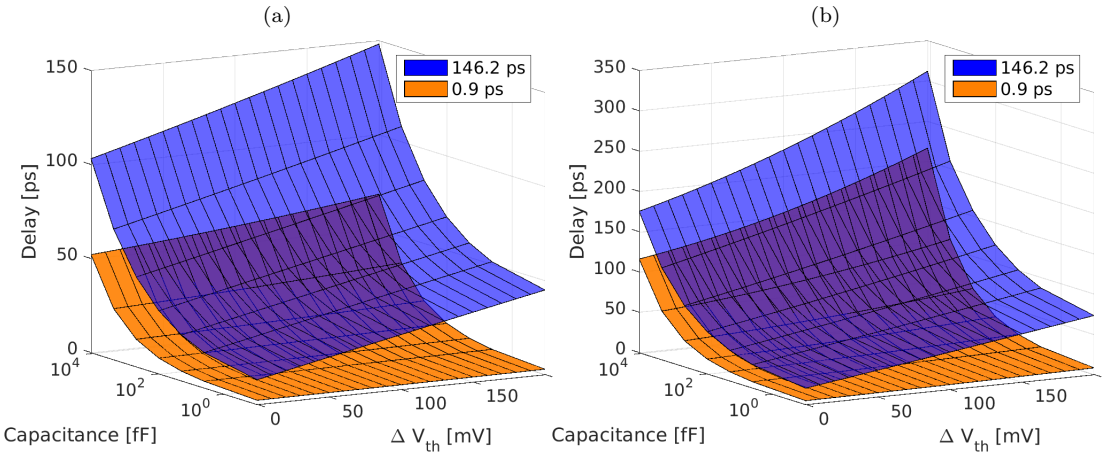


Figure 6.2: SPICE measurement results of the delay for an inverter's rising output transition as function of PMOS threshold voltage shift, load capacitance and input slew rate (color) for 273 K and 1.25 V (panel a) as well as 398 K and 0.95 V (panel b).

6.2.1 Inverter

At first, the timing degradation of an inverter is studied. Here, a rising or falling output transition has to be regarded as solely discrete factor. Sampling points that are stated above are used for non discrete factors. Since an inverter only contains a single PMOS transistor, a single threshold voltage shift has to be calculated in order to estimate the NBTI induced timing degradation. In particular, threshold voltage shifts up to 200 mV with 10 mV steps are used.

As stated in Section 5.3, the hot-carrier degradation effect may also be simulated based on the phase space approach. By respecting both NBTI and HCD induced threshold voltage shifts for the PMOS transistor as well as a threshold voltage degradation of the NMOS transistor that is caused by HCD, the timing degradation due to both degradation effects can be calculated. In this way, HCD can be integrated in the gate level degradation analysis without difficulty. However, an NBTI only analysis is used within this thesis, as the additional HCD component of the gate level degradation analysis cannot be evaluated due to missing HCD characterization data.

Different methods are used to estimate regression curves of the delay and output slew rate as a function PMOS threshold voltage shift. The main methods use tabular values of delay and output slew rate for all sampling points and inexistent threshold voltage degradation similar to the liberty library approach. In addition, the variations due to threshold voltage degradation are fitted with a linear as well as a quadratic function. These fits are performed with Matlab's fminsearch function that implements the simplex search method, since conformity with the tabular values of the non degraded case is used as an additional requirement. Size of the degraded timing database is thus twice and threefold as large as the timing information within the liberty library for linear and quadratic method, respectively. Furthermore, a second order Taylor series expansion with expansion point 7.42 fF, 30.2 ps, 323 K and 1.1 V is performed for the purpose of comparison.

6.2.1.1 Results

SPICE measurement results of the delay degradation for an inverter's rising output transition is shown in Figure 6.2. Increasing load capacitance, input slew rate and threshold voltage shift all cause an increase of the inverter's delay. Furthermore, surfaces of different input slew rates are not only displaced but demonstrate a completely different degradation behavior. The delay of the inverter also clearly depends on temperature and supply voltage, since maximal delay of the condition 398 K and 0.95 V is twice as large as within the condition 273 K and 1.25 V.

Dependence of the inverter's delay degradation on load capacitance, input slew rate, temperature and supply voltage is further analyzed in Figure 6.3. Each panel shows the effect of a single

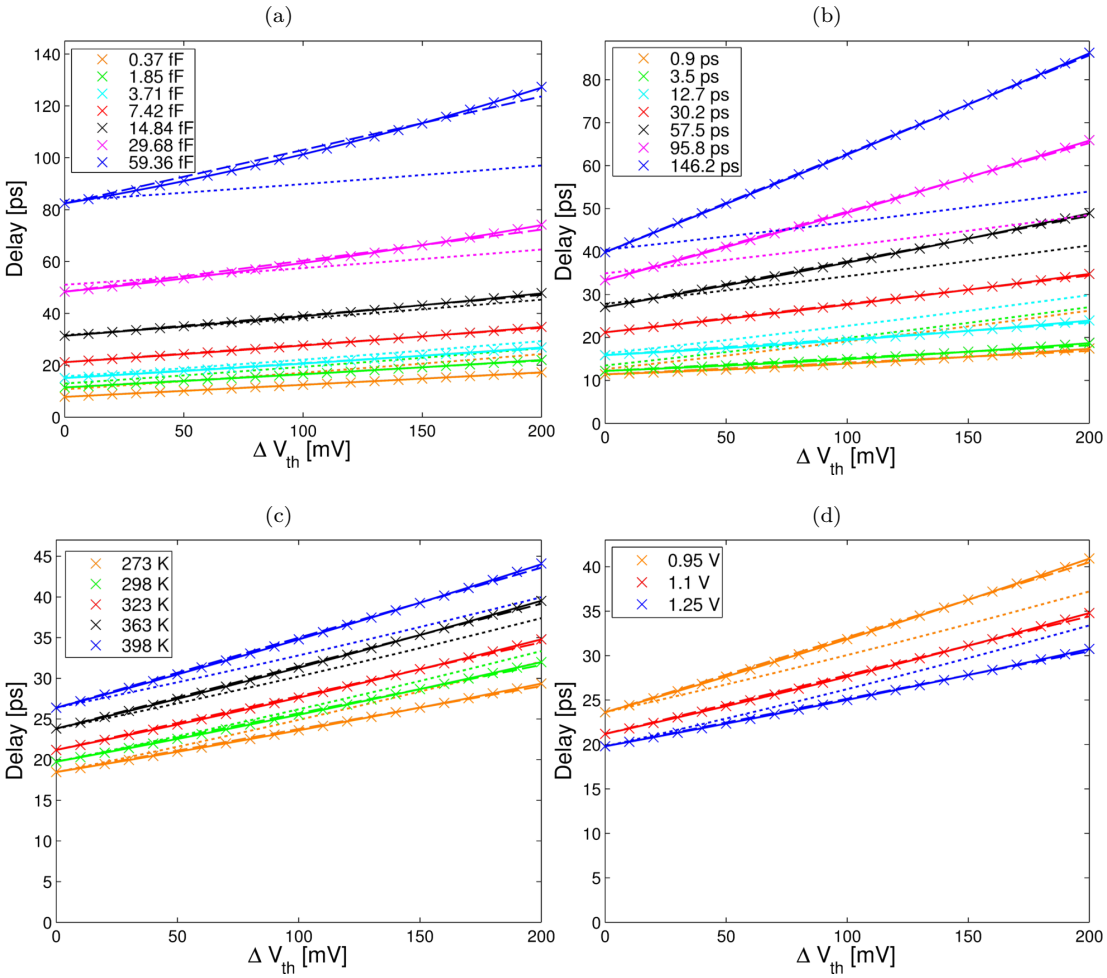


Figure 6.3: PMOS threshold voltage shift induced delay degradation for an inverter's rising output transition as function of load capacitance (panel a), input slew rate (panel b), temperature (panel c) and supply voltage (panel d). Red markings within each panel always illustrate the same data and determine the other parameter's fixed values within each panel in this way. Markers depict SPICE measurement results, while solid, dashed and dotted lines represent quadratic regressions, linear regressions and Taylor series expansions, respectively.

determining factor, while fixed values of 7.42 fF, 30.2 ps, 323 K and 1.1 V are used otherwise. In this way, the red markings within each panel always show the same degradation curve, whereas the other colors represent variations of a different determining factor in each case. As medium values are always used for three determining factors, maximal delays are considerably lower than within Figure 6.2b showing the combined effect of extreme values of load capacitance, input slew rate, temperature and supply voltage.

In addition to the impact of load capacitance, input slew rate and threshold voltage that is already stated above, an increase of temperature and supply voltage causes an increase and decrease of the inverter's delay, respectively. Regression curves within Figure 6.3 show that the dependence on PMOS threshold voltage shift is nearly linear. Only for 59 fF a slight deviation between linear regression curve and SPICE measurements is visible, which is perfectly characterized with the quadratic regression technique. Dotted lines, that represent the second order Taylor series expansions, are well suited for the non degraded case, as Figure 6.3 shows delay values of the Taylor series' expansion point. However, clear differences between measurement results and Taylor series expansions arise for increasing threshold voltages.

Figure 6.4 shows the degradation of the inverter's output slew rate for rising output transitions and various conditions of the non discrete factors. Although the slew rate degradation is slightly

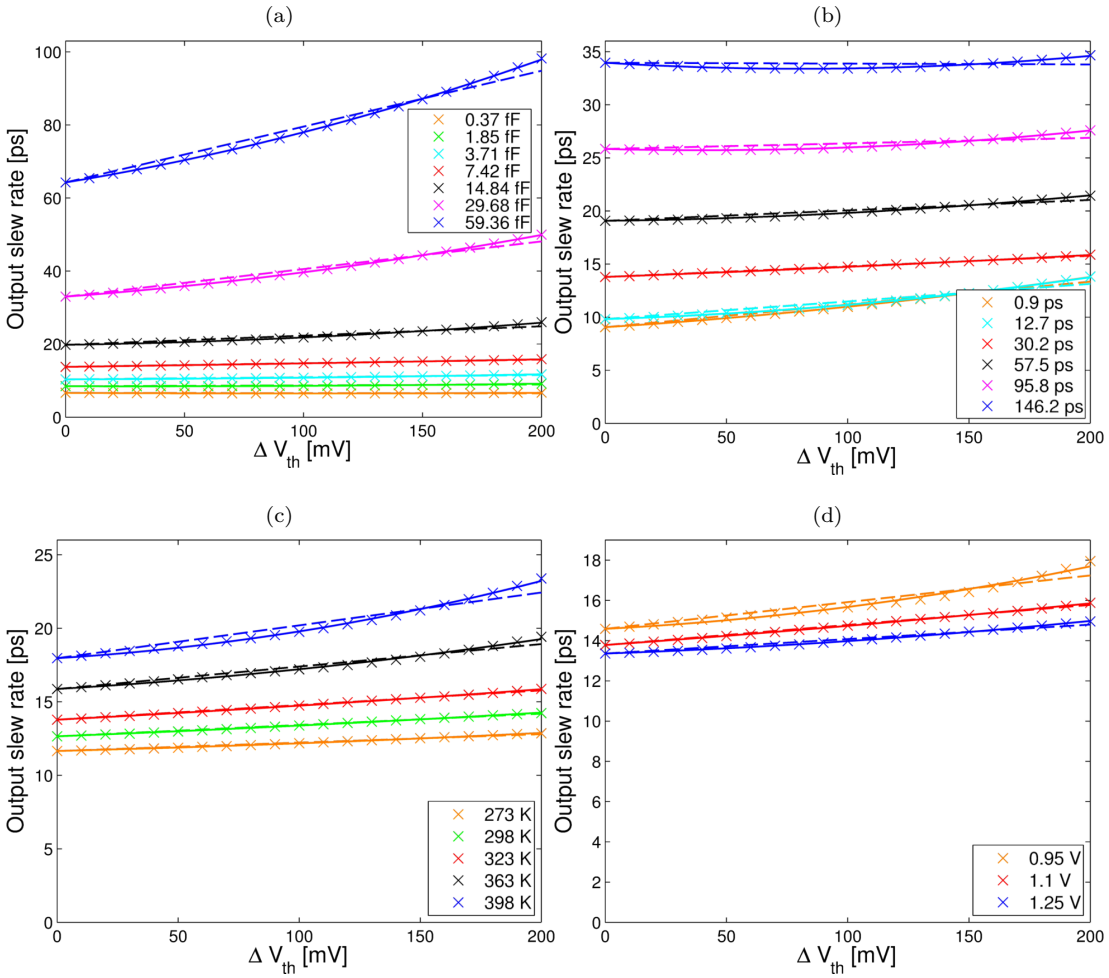


Figure 6.4: PMOS threshold voltage shift induced output slew rate degradation for an inverter's rising output transition as function of load capacitance (panel a), input slew rate (panel b), temperature (panel c) and supply voltage (panel d). Representation corresponds to Figure 6.3, whereas Taylor series expansions are omitted.

inferior to the delay degradation of Figure 6.3, dependence of output slew rates on PMOS threshold voltage shifts is clearly visible. Again, compliance with the SPICE measurement results is only slightly higher for the quadratic regression curves and the linear regressions may also be used as a precise abstraction of the slew rate degradation.

Impact of the NBTI induced PMOS threshold voltage shift of an inverter's falling output transition is presented in Figure 6.5. Whereas the output slew rate is clearly not affected by PMOS threshold voltage, the inverter's delay is slightly reduced for minor load capacitances. This is further analysis within Figure 6.6 by studying the delay degradation of an inverter's rising and falling output transition due to the interaction of PMOS and NMOS threshold voltage shift. Delay of the rising output transition is clearly increased and slightly diminished by PMOS and NMOS threshold voltage shift, respectively. The inverse dependence on PMOS and NMOS threshold voltage shift occurs for the falling output transition.

6.2.1.2 Discussion

Impacts of the non discrete factors PMOS threshold voltage, load capacitance, input slew rate, temperature and supply voltage are not independent from one another. In this way, different delay or output slew rate degradation characteristics occur for different combinations of load capacitance, input slew rate, temperature and supply voltage. Taylor series expansions are thus

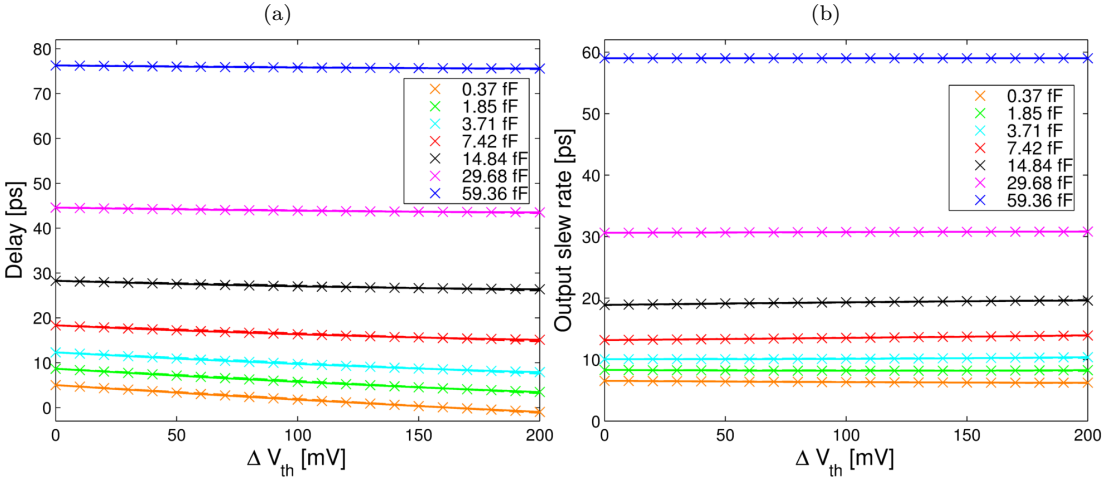


Figure 6.5: PMOS threshold voltage shift induced delay (panel a) and output slew rate degradation (panel b) for an inverter's falling output transition and different load capacitances, input slew rate 30.2 ps, 323 K and 1.1 V.

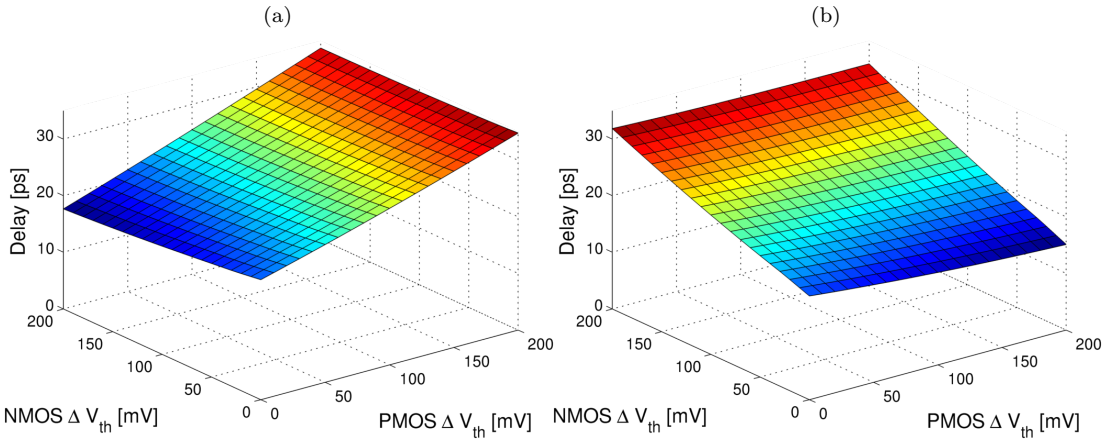


Figure 6.6: SPICE measurement results of the delay for an inverter's rising (panel a) and falling output transition (panel b) as function of PMOS and NMOS threshold voltage shift for 7.42 fF, 30.2 ps, 323 K and 1.1 V.

not suitable to characterize the measurement data as shown in Figure 6.3. However, liberty library like usage of a database for the non degraded case and characterization of the timing degradation within each condition based on quadratic or linear functions is in compliance with measurement results. This technique can thus be used to accurately simulate degraded timings due to NBTI induced PMOS threshold voltage shifts within the gate level degradation analysis.

While the NBTI degradation clearly causes a deterioration of the timing of rising output transitions, timings of falling transitions are slightly improved. However, these improvements are only marginal and do not compensate the NBTI timing degradations of rising transitions. Disregard of these improved timings will presumably not have an obvious implication on the results of the gate level degradation analysis. On the other hand, support for timing enhancements of falling transitions is though an improved simulation technique, as state-of-the-art gate level analysis methods (see Section 3.3.3) do not incorporate this effect. Thus, improved timings of falling transitions are included within the timing database that is generated as part of this thesis and impact of the support for improved timings on the timing simulation accuracy of complete circuits may be analyzed within subsequent studies.

NMOS threshold voltage shifts, which need to be considered when timing degradations due

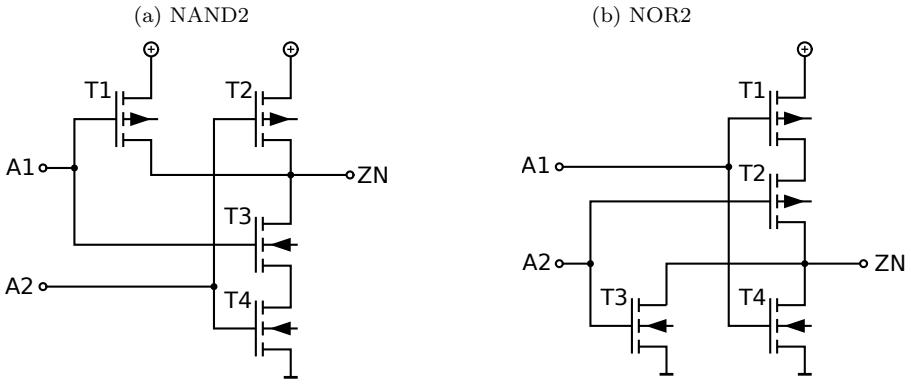


Figure 6.7: Examples of single-stage CMOS gates

to both NBTI and HCD effects are estimated, clearly cause a timing deterioration of the falling output transition. In this case, the same approach of a database with regression functions for the impact of NMOS threshold voltage on timing degradation can be used. The small interaction of PMOS and NMOS threshold voltage shifts (see Figure 6.6) may either be disregarded or NMOS and PMOS threshold voltage shift can be introduced as an additional non discrete factor within the database of rising and falling output transition, respectively.

6.2.2 Single-Stage Gates

The term single-stage CMOS gates denotes gates with more than one input pin and only one complementary PMOS NMOS pair between each input and output terminal. Two examples of this kind of gates are thus NAND and NOR gates that are shown in Figure 6.7. Timing characteristics of single paths between input and output terminals for this kind of gates are similar to the characteristics of the inverter with additional resistances and capacitances [67]. Hence, the same approach to estimate degraded timings that is already applied for the inverter can directly be used for single-stage gates. Each input pin of single-stage gates has to be analyzed separately as threshold voltage shifts of the respective PMOS and NMOS transistors can be different. Furthermore, capacitances and resistances of connection lines and the gate's other transistors influence the timing and require different timing database entries for each input pin.

Both quadratic as well as linear regression techniques described above are used for the threshold voltage induced timing degradations for each pin of different single-stage gates. The coefficient of determination R^2 is used to characterize the goodness of both regression techniques. In this way, necessity of the deployment of the more complex quadratic regression technique can be estimated.

6.2.2.1 Results

SPICE measurement results of delay and output slew rate degradations for each input pin of NAND2 and NOR2 gates are shown in Figure 6.8. Although, these measurements are performed for all combinations of non discrete factors' sampling points that are stated in Section 6.2, only datasets for rising output transitions and 7.42 fF, 30.2 ps, 323 K and 1.1 V are presented within this figure. For reference the respective degradation curves of the inverter are also replotted. Distinctive features of Figure 6.8 are varying degradation curves as well as timing characteristics of the non degraded case for different gates and input pins.

R^2 values of linear and quadratic fits of the timing degradation curves are presented in Table 6.1. For each gate's input pin, the quadratic regression method perfectly characterizes the SPICE measurement results and the linear regression technique only introduces minor inaccuracies. Since the R^2 values are based on measurement data for all combinations of non discrete factors' sampling points as well as rising and falling output transitions, it is evident that the used regression techniques are suitable for all these conditions.

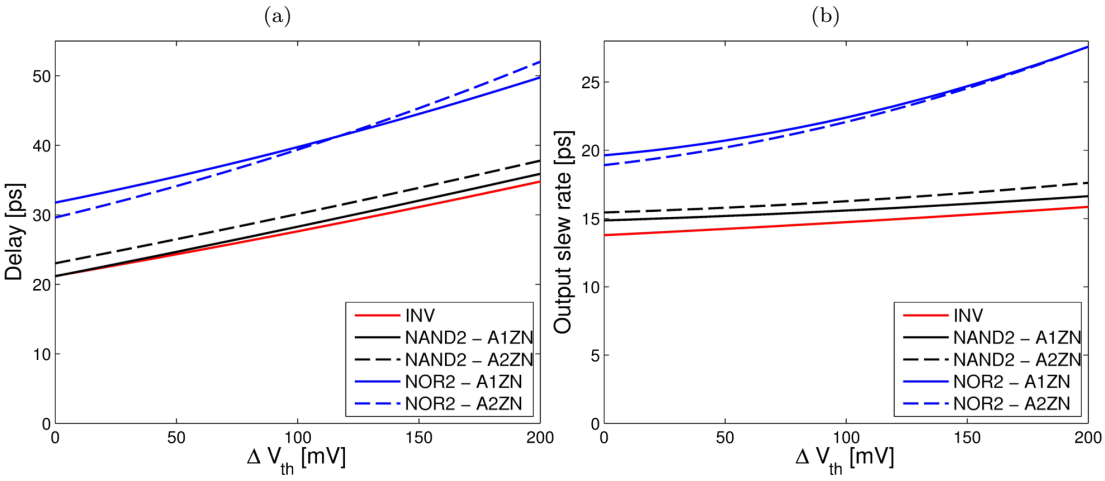


Figure 6.8: PMOS threshold voltage shift induced delay (panel a) and output slew rate degradation (panel b) for different single-stage gates and respective input pins. Measurement data corresponds to rising output transitions, load capacitance 7.42 fF, input slew rate 30.2 ps, 323 K and 1.1 V.

		INV	NAND2		NOR2	
			A1ZN	A2ZN	A1ZN	A2ZN
Delay	Linear fit	0.9998	0.9998	0.9998	0.9998	0.9998
	Quadratic fit	1	1	1	1	1
Output slew rate	Linear fit	0.9934	0.9976	0.9968	0.9995	0.9991
	Quadratic fit	1	0.9999	0.9999	1	1

Table 6.1: Coefficients of determination R^2 for linear and quadratic regression techniques of the timing degradation curves for different single-stage gates and respective input pins. R^2 values are based on measurement data for all combinations of non discrete factors' sampling points as well as rising and falling output transitions.

6.2.2.2 Discussion

Varying degradation curves and timing characteristics of the non degraded case illustrate the necessity of different timing database entries for each gate's input pin. Timing abstraction techniques of the proposed database are also well suited for single-stage gates, as abstracted timing degradation curves only marginally deviate from direct SPICE measurements. Since the linear abstraction method already provides highly accurate results, this method should be sufficient for the gate level degradation analysis. In this way, size of the degraded timing database is only twice as large as the timing information within the liberty library.

6.2.3 Multiple-Stage Gates

The term multiple-stage CMOS gates denotes gates with more than one complementary PMOS NMOS pair between each input and output terminal. Two examples of this kind of gates are thus AND and OR gates that are shown in Figure 6.9. Due to multiple PMOS NMOS pairs, timing characteristics of single paths between input and output terminals are affected by NBTI degradations of multiple PMOS transistors for this kind of gates. The approach to estimate degraded timings that is presented in the previous section can thus not directly be used for multiple-stage gates. In particular, several PMOS transistors with different degradation states have to be considered. This complexity can however be handled by a segmentation of multiple-stage gates into several single-stage gates within the timing analysis similar to [46]. In

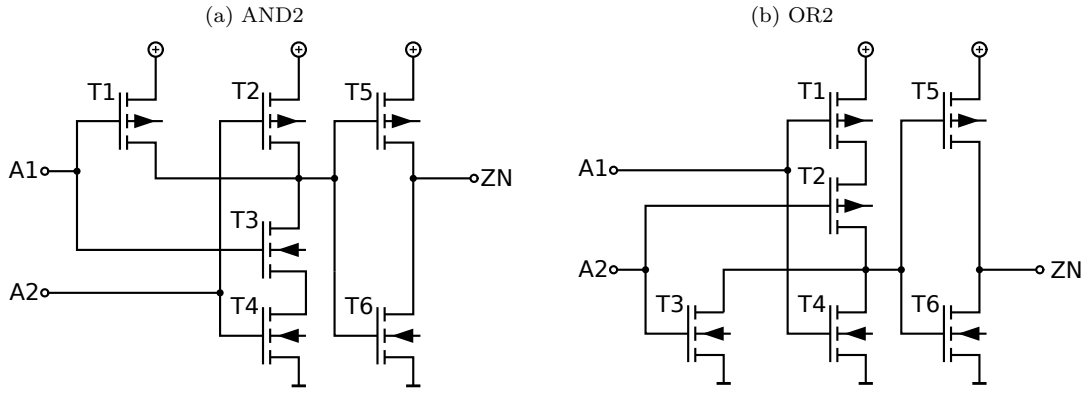


Figure 6.9: Examples of multiple-stage CMOS gates

this way, AND and OR gates of Figure 6.9 are subdivided into NAND and NOR gates and a subsequent inverter, respectively. The database approach for degraded timings can thus also be applied for multiple-stage gates.

Subdivision of multiple-stage gates into several single-stage gates requires some additional considerations. At first, signal probabilities of the subsequent single-stage gates have to be computed based on the signal probabilities of the input pins of multiple-stage gates, which are stated in the stress scenario. Secondly, load capacitances of the preceding single-stage gates due to internal wire loads and subsequent transistors have to be estimated using SPICE. By reason of different transistor sizes, the database entries of the respective single-stage gates may also not directly be used for single-stage parts within multiple-stage gates. The NAND2 part of the 45 nm NanGate AND2 gate for instance features smaller transistors as it has to only drive the low load capacitance of the INV part and the standard NAND2 database entry can thus not be used. Therefore, additional database entries are needed for the respective single-stage gates that feature the same characteristics as the single-stage parts within different multiple-stage gates.

6.2.3.1 Results

Delay of the falling output transition of an AND2 gate as function of PMOS threshold voltage shift of the transistor T1 (see Figure 6.9a) is shown in Figure 6.10 for three different load capacitances. In order to evaluate the subdivision of multiple-stage gates into several single-stage gates, delay of the respective NAND2 gate with the same PMOS threshold voltage degradation is also presented. As this gate only drives internal wires and the subsequent inverter, different load capacitances do not affect the NAND's delay. In particular, the NAND's load capacitance is specified within a previous SPICE measurement and is set to 2.13 fF. Furthermore, delay of an inverter that drives the three different load capacitances and receives a linear input signal that is defined by the NAND's output slew rate is also included in Figure 6.10. The visible marginal degradation of the inverter's delay is thus caused by the NAND's output slew rate degradation. As can be seen in Figure 6.10, the addition of the delays of NAND and inverter nearly perfectly match the delay degradation of the AND2 gate for all three load capacitances.

6.2.3.2 Discussion

Proposed method of subdivision of multiple-stage gates into several single-stage gates in combination with the database approach for timing estimation of single-stage gates accurately simulates degraded timings of multiple-stage gates. Only marginal inaccuracies may arise due to the abstraction of curved output signals using linear functions, that are characterized by slew rate values, and utilization of these linear functions for subsequent single-stage parts (see Figure 6.1).

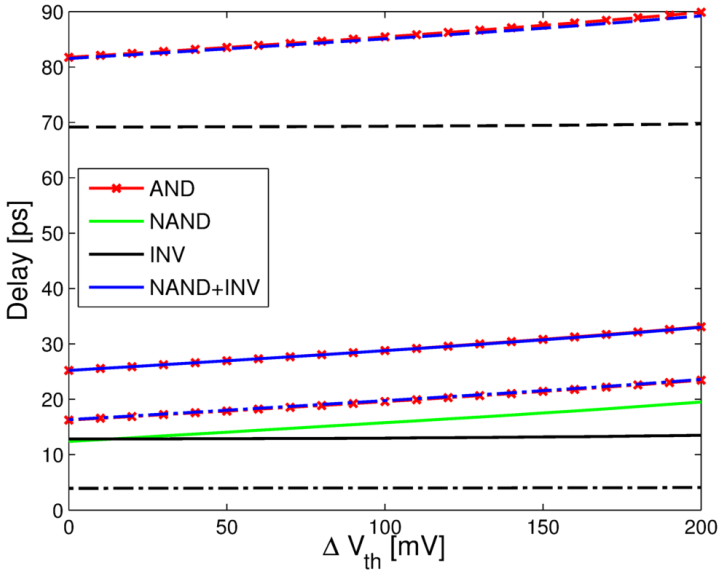


Figure 6.10: SPICE measurement results of the delay for an AND2's falling output transition as function of PMOS threshold voltage shift of the transistor T1 (see Figure 6.9a) for input slew rate 30.2 ps, 323 K, 1.1 V and load capacitance 0.37 fF (dash-dotted line), 7.42 fF (solid line) and 59.36 fF (dashed line). Individual delays of NAND and INV part of the AND gate as well as addition of both delays are presented using green, black and blue lines, respectively.

6.3 Summary

Main goal of the gate level degradation analysis is the estimation of the degraded timing of complete RT components due to NBTI by generating SDF files with degraded gate delays for specific mission scenarios. As realistic threshold voltage shifts of transistors within CMOS logic gates can already be calculated efficiently for varying mission scenarios using the phase space approach of Chapter 5, this chapter proposes a method to calculate the degraded timings of CMOS logic gates based on the estimated threshold voltage shifts. The method relies on SPICE measurements of single-stage gates for every input pin and type of output transition as well as reasonable sampling points of load capacitance, input slew rate, temperature, supply voltage and PMOS threshold voltage shift. While each measured delay and output slew rate value of the non degraded case is directly stored within a database similar to the liberty library approach, the impact of the threshold voltage shift is initially abstracted with linear or quadratic regression techniques in order to reduce the size of the database. Similar linear techniques are already applied within different gate level analysis approaches [16, 17], which however do not provide regression parameters for different gates and thus cause the necessity of SPICE characterization measurements within this thesis. Overall goodness of the linear regression technique with R^2 values always greater than 0.99 is already very high for both rising and falling output transitions that are subjected to considerable timing degradations and slight timing improvements, respectively.

Within the gate level degradation analysis non degraded output slew rates and delay values as well as linear degradation regression parameters are interpolated for respective load capacitance, input slew rate, temperature, supply voltage and PMOS threshold voltage shift using the timing databases as shown in Figure 6.11. Thus, the proposed timing estimation method directly supports temperature and supply voltage values of worst case timing conditions as well as realistic conditions that are specified by the mission scenario. Furthermore, high performance demands of the generation of SDF files are fulfilled due to the pre-characterized timing data. Due to a subdivision of multiple-stage gates into several single-stage gates during timing estimations, this type of gates can also be supported by the proposed method without introducing noteworthy inaccuracies. Finally, the timing estimation technique can also easily be extended in order to also support timing impacts of HCD induced NMOS threshold voltage shifts.

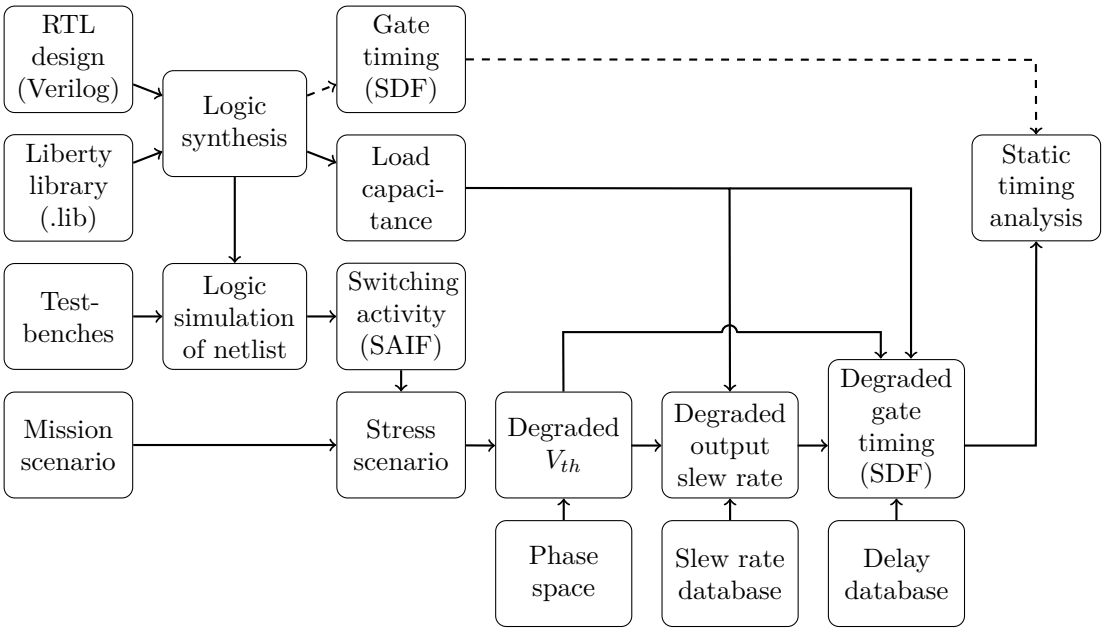


Figure 6.11: Proposed gate level degradation analysis flow extended by slew rate and delay degradation database as additional input arguments for the estimation of degraded timings

CHAPTER 7

Application Example: Degradation Analysis of a Single RT Component

Main objective of this chapter is to present a proof of concept for the proposed mission scenario aware gate level degradation analysis and to assess the impact of static worst case assumptions for NBTI induced delay degradations of single RT components. Thus, the proposed gate level degradation analysis flow of Chapter 4 is applied in combination with the phase space model and the database for degraded timings, which are already evaluated in Chapter 5 and 6, respectively. For this reason, a complete phase space has to be first generated, which also facilitates a performance evaluation of the phase space model. Finally, multiple degradation analyses are performed for a single RT component to assess the impact of different mission scenarios on the component's degraded delay.

7.1 Generation of a Pre-Characterized Phase Space

Key component of the proposed gate level degradation analysis is the phase space model, which should facilitate mission scenario aware yet efficient threshold voltage simulations. Although the simulation performance is a major design objective of the phase space approach, Chapter 5 only focuses on the evaluation of the fundamental modeling techniques and omits the generation of a phase space. As degradation analyses should be performed for a complete RT component and long time mission scenarios within this chapter, it is thus necessary to generate a pre-characterized phase space, which should clearly improve the performance of the phase space model.

The phase space is generated as illustrated in Figure 5.4 on page 41 by utilizing the CET model for the 22 nm technology as base model. The time step of the phase space generation, which can be set to arbitrary values, affects the phase space approach in several ways. At first, the phase space time step determines the minimal duration of constant stress phases for the utilized stress scenarios, as constant stress is presupposed for each phase space time step. Furthermore, longer time steps facilitate more efficient and more precise long time simulations, but cause higher computational efforts of the phase space generation. Due to these constraints, a phase space time step of 10 minutes is chosen within this chapter. Initial values of the abstraction parameters have to be chosen with respect to possible degradation states within the reference map (see Section 5.2.6) for the 22 nm technology and reasonable maximal abstraction parameter values have to be incorporated within the phase space generation, as a subsequent phase space simulation must not leave the area of precomputed values. In particular, 97 different values between 1 mV and 1.4 V, that approximately follow a logarithmic distribution, are used for the abstraction parameter transient shift in threshold voltage $\Delta V_{th\ Trans}$. The abstraction parameter "maximum capture time" is sampled by two values within each decade between 10^{-2} s and 10^{16} s with a total number of 37 different values. These very high maximum capture time values, which cannot directly occur within a stress scenario, are required as the conversion technique for phase space simulations with varying temperatures (see Section 5.2.5) may induce a major amplification of this abstraction parameter. Furthermore, initial values of 0, 2, 2.5, 3, 3.5, 4, 5, 6, 8, 12, 18 and 30 are used for the abstraction parameter "slope width".

As described in Section 5.2.1, parameters characterizing the constant stress scenario of a 10 minute interval are also required for the phase space generation. Due to the usage of the reference map approach within the phase space model, the parameters temperature and supply voltage don't need to be incorporated within the phase space. However, the parameters duty frequency and duty cycle of the NBTI stress signal as defined by Figure 3.12 have to be considered. While duty frequencies of 1 and 10 Hz are directly chosen to minimize the computational effort of the phase space generation, a small analysis is carried out to identify the required values of the parameter duty cycle. CET simulations are hence performed for the 22 nm technology and stress scenarios that consist of two 10 minute stress intervals each with one of 35 different duty cycle values in the range between 1 and 99%. In detail, all 35 different duty cycles of the second stress interval are examined for each duty cycle value of the first interval, resulting in 35^2 different stress scenarios. The simulated threshold voltage shifts of the scenarios having duty cycles of 1, 20, 50, 80 and 99% within the first interval are shown in Figure 7.1. As can be seen, all curves have a typical S-like shape with the main difference of diverse gradients in central part of the curves.

Since the main S-like dependence on duty cycle is preserved even if the duty cycle is changed

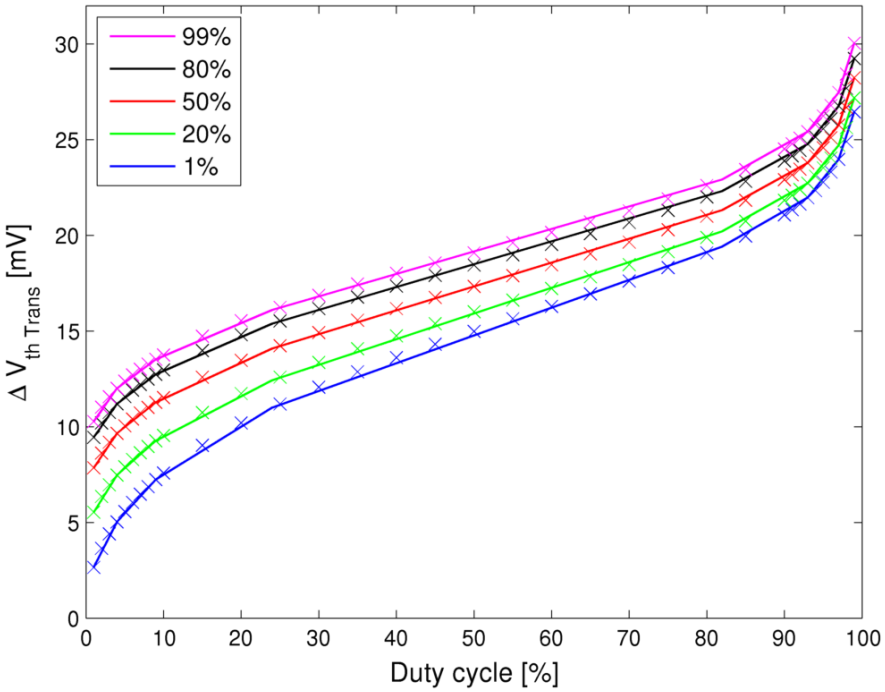


Figure 7.1: Transient threshold voltage shifts for 22 nm technology as function of duty cycle of a 10 min NBTI stress scenario with 10 Hz, -0.8 V and 400 K. Duty cycle within a previous 10 min stress interval, resulting in an overall stress duration of 20 min, is color coded. Lines represent piecewise linear fits between identified distinctive duty cycles of 1, 4, 9, 24, 82, 93, 97 and 99%.

within a stress scenario, a reasonable choice of duty cycle values for the phase space generation may reduce the computational effort of the generation without affecting the phase space model's accuracy. In order to identify number and values of the required duty cycles, the S-like shaped curves are further analyzed by performing piecewise linear fits between distinctive duty cycle values, as the phase space model uses linear interpolations between basic phase space entries. Thus, Matlab's "fminsearch" function, that is based on the simplex search method, is used to estimate optimal duty cycle values for different numbers of distinctive duty cycles. Main result of this analysis is that 8 different distinctive duty cycles with the values 1, 4, 9, 24, 82, 93, 97 and 99% already characterize the threshold voltage degradations within the 35 different S-like shaped curves with marginal linear interpolation error (see Figure 7.1). Furthermore, the increase in interpolation accuracy due to a higher number of distinctive duty cycles is negligible. Hence, these duty cycle values are used for the generation of the phase space. In this way, the computational effort of the generation process is reduced considerably, while the generated phase space is nearly as precise as a phase space based on all 35 duty cycle values.

Finally, CET simulations have to be performed for all 689k combinations of parameter values. As two parallel processes were utilized to calculate the phase space, the overall phase space generation took approximately two and a half weeks using a compute server with 4 AMD Opteron octa-core processors and 500 GB RAM. By using more parallel processes or even different machines, this computation duration can be reduced considerably. A three-dimensional quiver plot representation of the resulting phase space for the condition 82% duty cycle and 10 Hz duty frequency is shown in Figure 7.2. In particular, the arrows for extremal initial values of transient shift in threshold voltage and slope width are always directed inwards. Due to this characteristic of the phase space and the usage of very large initial values of the parameter "maximum capture time" that exceed reasonable stress durations, it is ensured that subsequent phase space simulation cannot leave the area of precomputed values.

Figure 7.3 shows a more meaningful representation by omitting several phase space entries. In detail, the representation uses a reduced phase space resolution and is restricted to the abstraction parameters $\Delta V_{th\ Trans}$ and maximum capture time. Thus, the decrease of $\Delta V_{th\ Trans}$ for very

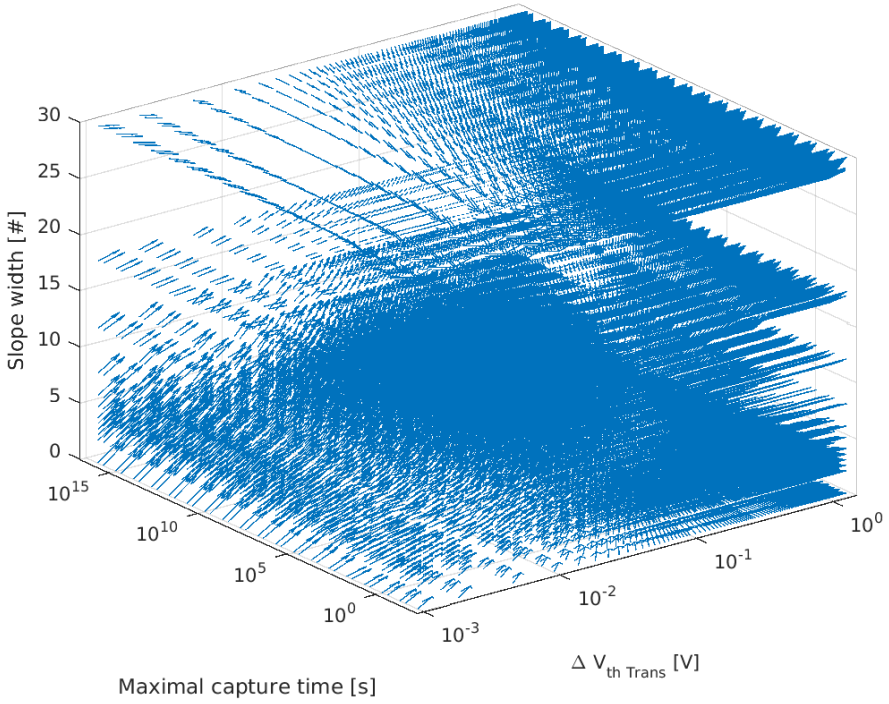


Figure 7.2: Quiver plot of the calculated phase space for 22 nm technology, 10 Hz and 82% duty cycle. Arrows are scaled uniformly to improve the visualization by reducing the overlap of different arrows.

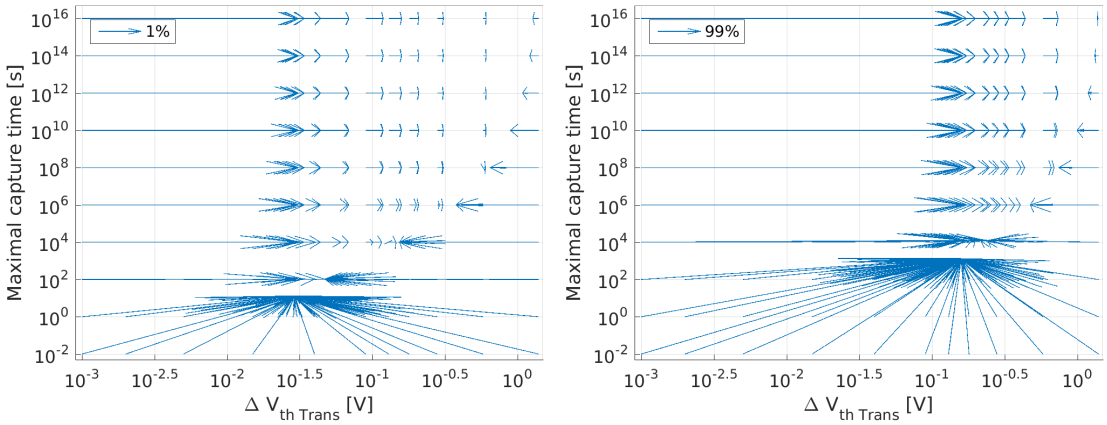


Figure 7.3: Quiver plot of the calculated phase space for 22 nm technology, 10 Hz and 1% (left panel) as well as 99% duty cycle (right panel). Representation is simplified by omission of the slope width dimension and several intermediate phase space entries. In contrast to Figure 7.2, correct arrow length are displayed.

high initial values of this abstraction parameter is directly visible. This decrease is caused by an initial occupation of the very short emission time region that is discharged during a 10 minute stress interval. Furthermore, the effect that initial values of maximum capture time do not decrease within a stress interval, which is already stated in Section 5.2.4 can also be seen. In addition, the different panels of Figure 7.3 illustrate the transformation of phase space entries due to different duty cycles. Main effect of higher duty cycles is hence a larger increase of $\Delta V_{th\ Trans}$ and maximum capture time within the phase space. In this way, arrow lengths within the quiver plot representation are increased due to higher duty cycles.

The evaluation of the phase space approach that is carried out in Chapter 5 does not

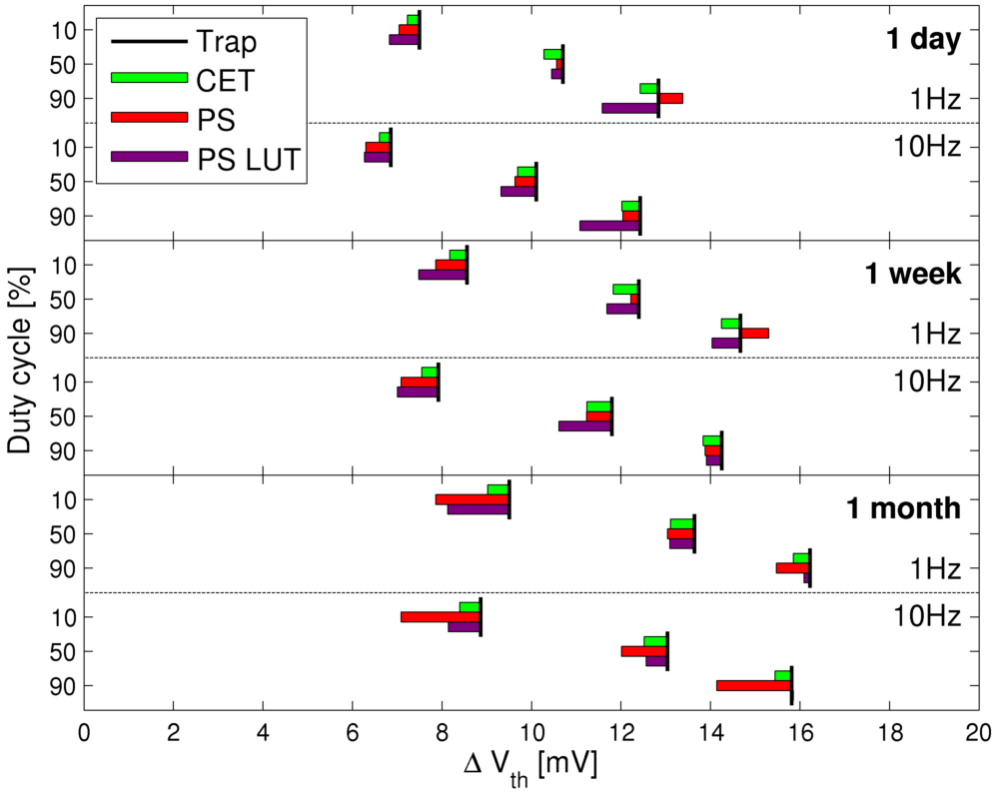


Figure 7.4: Simulated threshold voltage degradations for 22 nm technology and scenarios with 300 K, -0.8 V and different duty cycles (y-axis), stress durations (panels) and frequencies (sub-panels). Short black vertical lines depict simulation results of the switching trap model (Trap) based on provided trap lists (see Section 5.1.1) as reference. Deviations between switching trap simulation results and results of CET model, phase space model (PS) and phase space model based on the generated phase space lookup table (PS LUT) are presented using green, red and purple bars, respectively. Apart from the simulation results of the phase space model based on the generated lookup table (PS LUT), results are replotted from Figure 5.33.

incorporate interpolation inaccuracies that are determined by the phase space's resolution. Thus, an analysis of the accuracy is required for the complete phase space model that relies on the generated phase space. Figure 7.4 hence contains evaluation results for the generated phase space and several constant scenarios with different duty cycles, stress durations and frequencies. As these scenarios correspond to Figure 5.33 on page 71, the respective phase space simulation results without interpolation inaccuracies are replotted for comparison. Furthermore, the corresponding relative errors that are averaged for each duration, frequency and duty cycle are presented in Table 7.1.

As mean relative errors of both phase space models are in the range of 10%, the usage of the generated phase space does not cause a clear decline of the phase model's accuracy. In addition, a comparison of both phase space models' simulation results reveals that the interpolation inaccuracies within the generated phase space can lead to over- as well as underestimations of the threshold voltage degradation. Interestingly, the relative error of the lookup table based phase space model is even clearly reduced for the simulation duration of one month as can be seen in Table 7.1. Since the conversion techniques that are utilized within the phase space model to support scenarios with varying temperatures or supply voltages do not depend on the model's pre-generated phase space, a reassessment of the complete phase space approach is not required for these scenarios. In summary, the phase space resolution should thus be sufficient to facilitate an accurate simulation of the NBTI induced threshold voltage degradation.

		Duration			Frequency		Duty cycle		
		1d	1w	1m	1Hz	10Hz	10%	50%	90%
CET	RMS of rel. error [%]	3.82	4.31	4.19	4.04	4.18	4.72	4.44	2.95
	Underestimation [%]	100	100	100	100	100	100	100	100
PS	RMS of rel. error [%]	5.39	6.71	14.9	8.45	11.23	15.34	4.98	6
	Underestimation [%]	83	83	100	78	100	100	100	67
PS LUT	RMS of rel. error [%]	9.39	9.7	8.17	9.39	8.82	12.45	6.72	6.99
	Underestimation [%]	100	100	83	100	89	100	100	83

Table 7.1: Root mean square of relative errors and underestimation percentages of the simulation results presented in Figure 7.4. Stated data of CET model, phase space model (PS) and lookup table based phase space model (PS LUT) refer to the corresponding simulation results of the switching trap model (Trap) and are divided by the different stress scenarios of the associated figure.

7.2 Performance of the Phase Space Model

In order to assess the average performance of the phase space model and to compare the performances of different NBTI models, simulation durations of several NBTI simulations that are performed on the same computer are examined. In particular, the performance evaluation comprises the phase space PS approach in combination with the pre-characterized phase space of Section 7.1, the capture-emission time CET model as described in Section 5.1.2 and the switching trap model of Section 5.1.1. Stress scenarios with 1 week stress duration, 350 K, -0.8 V, 85% duty cycle and duty frequencies of 1 Hz and 10 Hz are utilized. Mean values SD and corresponding standard deviations $\overline{SD}$ of the simulation durations are estimated on the basis of 100, 10 and 5 repeated measurements for phase space, CET and switching trap model, respectively.

Resulting simulation durations of the three different NBTI models are shown in Table 7.2. As can be seen, the phase space model's performance with simulation durations of approximately 3 s is considerably better than the performances of CET and switching trap models. Furthermore, simulation durations of the phase space model do not depend on the duty frequency of the NBTI stress condition. In contrast, both CET and switching trap models exhibit a distinct increase in simulation duration for the duty frequency 10 Hz. Possible improvements of the simulation duration that may be caused by substitution of CET or switching trap model with the phase space approach are also stated in Table 7.2. In detail, the acceleration factor AF, that is defined by Equation 7.1, is used to characterize the potential performance improvements. Uncertainty of the acceleration factor $\overline{AF}$ is precisely specified by the standard deviations of the corresponding simulation durations and the Gaussian error propagation rule as stated in Equation 7.2.

$$AF = \frac{SD\,(\text{Physical model})}{SD\,(\text{PS model})} \tag{7.1}$$

$$\overline{AF} = \sqrt{\left(\frac{\overline{SD}\,(\text{Physical model})}{SD\,(\text{PS model})}\right)^2 + \left(\frac{SD\,(\text{Physical model}) \cdot \overline{SD}\,(\text{PS model})}{SD\,(\text{PS model})^2}\right)^2} \tag{7.2}$$

The performances of CET and switching trap models should further decline for higher duty frequencies, as the number of computational operations increases with increasing duty frequency. In contrast, the performance of the phase space model is independent of duty frequency, since the frequency only affects the computational effort of the phase space generation. Substitution of CET or switching trap model with the phase space approach should thus cause even higher performance improvements for more realistic stress scenarios with higher duty frequencies.

Stress scenarios with changing temperatures or supply voltages require additional conversion techniques as described in Section 5.2 that slightly diminish the phase space model's performance.

	Duty frequency [Hz]	Simulation duration [s]	Simulation acceleration of phase space model [factor]
Phase space model	1	3.29 ± 0.16	
	10	3.22 ± 0.12	
CET model	1	363.7 ± 12.1	110.5 ± 6.5
	10	3664.5 ± 40.1	1138 ± 44
Switching trap model	1	47601 ± 317	14468 ± 710
	10	74725 ± 5872	23207 ± 2018

Table 7.2: Averaged simulation durations of different NBTI models for stress scenarios with 1 week stress duration and duty frequencies of 1 Hz and 10 Hz

As the published CET model [38] does not support these stress scenarios, the same conversion techniques have to be also integrated within the CET model. Thus, the CET model's performance is reduced for these varying scenarios in almost the same manner. In contrast, the switching trap model should directly support these varying scenarios without a noticeable performance impact. However, the switching trap model's performance is clearly inferior to the performance of the CET model and does not fulfill the performance demands of the gate level degradation analysis.

7.3 Degradation Analysis of a 4-Bit Adder

As a proof of concept, the proposed gate level degradation analysis is utilized to assess the delay degradation of a complete RT component for different mission scenarios. Without the loss of generality, a 4-bit adder that only consists of INV, NAND2 and NOR2 gates is analyzed. Overall lifetimes of 1 and 10 years are utilized and delay degradations caused by a worst case and a more realistic mission scenario are examined. While a constant temperature of 400 K as well as a permanent operation state of the RT component are assumed within the worst case scenario, the realistic scenario only expects a 6 h operation state per day and utilizes temperatures of 350 K and 300 K within operation and off state, respectively. However, the last operation state of the realistic scenario, which directly precedes the delay degradation estimation, also uses the worst case temperature of 400 K. The realistic scenario hence assumes that the worst case state occurs only once after a long time of typical usage. As stated in Section 4.1, these different mission scenarios are characterized within stress scenario XML files.

As individual stress states are needed for each device within the RT component to precisely simulate the timing degradation, signal and transition probabilities of each gate's input pins are estimated by industrial tools and are also provided within the XML files. Based on the definition of NBTI stress conditions presented in Section 3.3.1, the duty cycle of the NBTI stress signal is directly determined by signal probability and gate type. However, the conversion between transition probabilities and duty frequencies of the corresponding NBTI stress conditions relies on the clock frequency of the RT component. In this way, a very low clock frequency is chosen to ensure that the duty frequencies of the NBTI stress conditions are in the range between 1 Hz and 10 Hz. Although the corresponding clock frequency is unrealistic, the principle impact of different transition probabilities on the threshold voltage degradation is considered. As the main objectives of this chapter are the proof of concept for the proposed degradation analysis and the assessment of the impact of static worst case assumptions, the usage of a realistic clock frequency would only provide a small benefit but would cause a clear increase of the computational effort of the phase space generation due to the performance of the CET model being dependent on duty frequency (see Table 7.2). Besides, there is only a small impact on the threshold voltage degradation and NBTI was even thought to be independent of a stress signal's frequency a few years ago as stated in Section 2.1.1.

Phase space simulations are performed to estimate threshold voltage degradations for each gate's input pins based on the stated stress conditions. In order to calculate each gate's timings, load capacitances are estimated by industrial tools as illustrated in Figure 6.11. In addition to

		Gate delay [ps]				
		Fresh	Worst case		Realistic mission scenario	
			1 year	10 years	1 year	10 years
U83 (NAND2_X1)	Rise	25.99	39.55	44.03	30.63	31.16
U74 (NAND2_X1)	Fall	27.32	25.76	25.27	26.87	26.78
U65 (NAND2_X1)	Rise	15.22	23.69	26.59	17.83	18.26
U64 (NAND2_X1)	Fall	30.01	29.82	29.74	29.96	29.95
U63 (INV_X1)	Rise	18.21	29.09	32.78	21.95	22.31
U55 (NAND2_X1)	Fall	17.29	17.06	16.77	17.35	17.33
U54 (NAND2_X1)	Rise	13.59	20.28	22.92	15.47	15.80
U53 (NAND2_X1)	Fall	23.43	23.17	23.05	23.36	23.35
U52 (INV_X1)	Rise	11.88	18.90	21.97	14.34	14.57
U46 (NAND2_X1)	Fall	15.84	15.85	15.87	15.87	15.86
U41 (NAND2_X1)	Rise	17.20	24.61	27.32	19.32	19.71
Data arrival time [ps]		215.97	267.78	286.31	232.95	235.07
Path delay degradation [%]			23.99	32.57	7.86	8.84
Maximal ΔV_{th} [mV]			175.99	233	61.64	70.46

Table 7.3: Delay degradation of a 4-bit adder due to worst case conditions as well as a realistic scenario

the analyses within the RT component, input slew rates and load capacitances are required for the gates that are connected to input and output terminals of the RT component, respectively. In accordance with the fan-out of 4 procedure [68], the fourfold capacitances of the driving gates are used as capacitances of the output terminals. A similar approach is used to identify a reasonable input slew rate. Therefore, a rectangular signal is passed through two inverters and the slew rate of the subsequent signal is estimated with the fourfold inverter capacitance as load capacitance using SPICE. Degraded slew rates can consequently be estimated for each internal net based on the timing database of Chapter 6. Afterwards, the final calculation of the gates' delays is realized. Based on these results, SDF files are generated for the stated scenarios as well as a "fresh" scenario without degradation to assess the delay transformation that is caused by NBTI. Finally, the overall delay of the RT component's critical path is estimated using the static timing analysis procedure within the industrial tool "Synopsys Design Compiler".

The resulting delays of every gate within the critical path as well as the path's overall delay are presented in Table 7.3 for each scenario. In accordance with Chapter 6, the NBTI effect causes a clear increase and slight decrease of a gate's delay for rising and falling output transitions, respectively. As the utilized timing database is based on the 45 nm technology node (see Section 6.2) and the NBTI model uses available characterization data of the 22 nm technology, the stated delay degradations within Table 7.3 do not correspond to a real technology. Thus, the essential result of the degradation analysis is the percentaged delay degradation for each mission scenario. In particular, the impact of different mission scenarios on the delay degradation is specified for the 22 nm technology by the differences between the percentaged delay degradations for these scenarios. The actual increase of the critical path's delay should though be considerably higher for the 22 nm technology, since the estimated threshold voltage degradations should result in a larger deceleration in comparison to the 45 nm technology node.

As can be seen in Table 7.3, the delay degradations that are caused by the realistic scenarios with final worst case state are considerable smaller than the worst case delay degradations. In detail, the realistic delay degradation after 10 years is 2.7 times smaller than the worst case degradation after a single year. Furthermore, the maximal threshold voltage degradations that occur within the RT component already demonstrate the principle impact of the different mission scenarios as presented in Table 7.3. However, the modeling of a RT component's failure rate based on the exceedance of a pre-defined threshold voltage boundary is very imprecise.

Instead, the noncompliance between predetermined delay specifications and the RT component's degraded delays for different mission scenarios can finally be used to estimate the overall failure rate of the RT component.

7.4 Summary

The generation of the required phase space is performed as first step of the application of the proposed mission scenario aware gate level degradation analysis. After a preliminary inspection to assess the required sampling values of the duty cycle parameter, 689k different phase space entries are calculated within approximately two and a half weeks based on the CET model for the 22 nm technology. This phase space generation has to be done only once for each technology and parameters like the phase space time step, which affects the approach's efficieny and support for rapidly changing stress scenarios, can be adapted to other needs if necessary. Trajectories of subsequent phase space simulations cannot leave the area of precomputed values as shown by an analysis of the generated phase space. Furthermore, a subsequent evaluation of the complete phase space model reveals that the interpolation inaccuracies within the phase space do not clearly affect the overall accuracy of the phase space approach. The utilized phase space resolution should thus be sufficient to facilitate an accurate simulation of the NBTI induced threshold voltage degradation.

Afterwards, measurements are carried out to assess the performance improvements of the complete phase space approach in comparison to the CET model. In this way, a decrease of simulation duration by factors of approximately 100 and 1000 are identified for duty frequencies of 1 and 10 Hz, respectively. This reduction of simulation duration is most likely even higher for more realistic frequencies as higher duty frequencies only induce an increase of the phase space generation effort within the phase space methodology. Besides, simulation durations of the phase space model are already more than 10^4 times faster than the switching trap model's simulation durations for these low duty frequencies.

A proof of concept for the proposed mission scenario aware gate level degradation analysis of Chapter 4 is finally provided by the simulation of the degraded delay of a 4-bit adder for different mission scenarios. By comparing the delay degradations of worst case condition and a more realistic scenario with final worst case state it is shown that a realistic delay degradation after 10 years is even 2.7 times smaller than the worst case degradation after a single year. Thus, static worst case assumptions for NBTI induced delay degradations of single RT components cause a considerable overestimation of the degradation effect.

CHAPTER 8

Conclusion

Within this thesis, a methodology is proposed that facilitates an accurate, efficient and reliable computation of the degraded timings of RT level components subjected to varying stress. The methodology exceeds the state-of-the-art in modeling of delay degradation by an integration of the proposed modeling techniques within an industrial tool flow and by a performance oriented yet varying mission scenario aware abstraction of the physical timing degradation models. Main advantages of the proposed methodology are thus an increased industrial relevance as degraded and non degraded timings are calculated similarly and a more realistic simulation of the degradation effect as authentic stress scenarios can be applied without a clear decline of the simulation's accuracy or performance. In this way, an accurate and efficient simulation of the expected timing degradation can be performed to examine the impact of static worst case assumptions on predicted failure rates of integrated circuits.

Specifically, a tool flow for a gate level degradation analysis is proposed that estimates the degraded timings of RT level components using mission scenarios that contain degradation relevant information about the application of an integrated circuit like ambient temperature variations in the course of several years and predicted off times. The deduction of the devices' varying stress states that is carried out using industrial tools facilitates individual NBTI simulations for each device's degraded threshold voltage under realistic conditions. Thus, a modeling technique is proposed for the estimation of NBTI induced threshold voltage shifts that supports varying stress states and permits numerous computations within reasonable time by utilizing a lookup table approach. The lookup table, which is called phase space, is based on an abstraction of the NBTI degradation state with 3 parameters and comprises pre-calculated transformations of the degradation state, which are calculated with a reference NBTI model for different scenarios. Since the CET model offers effective abstraction potentials it is used as reference model in the final version of the phase space model, although the CET model does not directly support varying temperatures nor gate voltages. Thus, additional modeling approaches are developed to facilitate CET as well as phase space simulations with varying temperatures and gate voltages as specified within the stress scenarios. These additional modeling approaches cause a slight deterioration of the phase space model's performance, but have no impact on the generation of the phase space, which is constructed in a temperature and voltage independent manner. Depending on each device's estimated threshold voltage shift, a timing database is subsequently utilized to efficiently calculate the degraded timings of CMOS logic gates. This database relies on SPICE simulations of single-stage gates that are carried out for every input pin and type of output transition as well as reasonable sampling points of load capacitance, input slew rate, temperature, supply voltage and PMOS threshold voltage shift. While each simulated delay and output slew rate value of the non degraded case is directly stored within the database similar to the liberty library approach, the impact of the threshold voltage shift is initially abstracted with a linear regression in order to reduce the database's size. Finally, the estimated degraded timings of CMOS logic gates are inserted within SDF files that are used as the main interface to perform static timing analyses of the complete circuit using verified industrial tools.

Individual evaluations of each stage of the proposed methodology are an essential part of this thesis to assess the methodology's overall accuracy. At first, an evaluation is performed for the simulation of degraded threshold voltages based on available NBTI measurement data of a 130 nm technology and a 22 nm FinFET technology. After several refinements of the initial phase space approach, it is shown that the deviation between results of the phase space model without interpolation inaccuracies and the reference switching trap model is almost always smaller than 15%. In many cases, the additional inaccuracy introduced by the phase space approach is in the same range as the inaccuracy that is caused by the CET approach. Furthermore, there is reason to assume that the model's accuracy should even increase for very long degradation times. Afterwards, a phase space with 689k different phase space entries is calculated within approximately two and a half weeks using two parallel processes on a compute server. Besides, this generation duration can be reduced considerably by further parallelization of the computation. A subsequent evaluation of the complete phase space model reveals that the interpolation inaccuracies within the phase space do not clearly affect the overall accuracy of the phase space approach. The utilized phase space resolution should thus be sufficient to facilitate an accurate simulation of the NBTI induced threshold voltage degradation. In contrast, the approach of [41], which is the only published model fulfilling the performance requirements of

the gate level degradation analysis and may thus be used as a replacement of the phase space model, shows significantly larger deviations to the reference results for all evaluation examples. In particular, mean deviations up to 400% arise for scenarios with varying temperatures. This modeling approach can thus not be applied within the gate level degradation analysis as the overall accuracy would greatly decline. After the assessment of the threshold voltage estimation, SPICE simulations based on the 45 nm NanGate „Free Process Design Kit“ and „Open Cell Library“ [25] are used to evaluate the linear regression between PMOS threshold voltage shifts and degraded timings of CMOS logic gates. It is thus shown that the overall goodness of the linear regression technique is very high, as the R^2 values are always greater than 0.99 for both rising and falling output transitions that are subjected to considerable timing degradations and slight timing improvements, respectively. In this way, it is shown that the overall accuracy of the proposed gate level degradation analysis is ensured as the additional computations within the analysis are performed by extensively verified industrial tools.

Since performance improvements of the NBTI simulation technique within the gate level degradation analysis are a major objective of this thesis, measurements are carried out to assess the performance of the complete phase space approach in comparison to the CET model. Thus, a decrease of simulation duration by factors of approximately 100 and 1000 are identified for duty frequencies of 1 and 10 Hz, respectively. This reduction of simulation duration is most likely even higher for more realistic frequencies as higher duty frequencies only induce an increase of the phase space generation effort within the phase space methodology. Besides, simulation durations of the phase space model are already more than 10^4 times faster than the switching trap model's simulation durations for these low duty frequencies. In consequence, the overall performance of the gate level degradation analysis should be sufficient to be integrated in the product development cycle without introducing considerable development delays.

A proof of concept for the applicability of the proposed gate level degradation analysis is finally provided by the simulation of the degraded delay of a 4-bit adder for different mission scenarios. Based on these results, the overall failure rate of an integrated circuit may afterwards be easily estimated by the noncompliance of each RT component's degraded timing with predetermined specifications. By comparing the delay degradations of worst case condition and a more realistic scenario with final worst case state it is shown that a realistic delay degradation after 10 years is even 2.7 times smaller than the worst case degradation after a single year.

In this way, the main research questions that are stated within the introduction are finally addressed. It is shown that static worst case assumptions concerning idle phases, temperatures and supply voltages cause considerable overestimation of the timing degradations of RT level components and thus failure rates of integrated circuits. Hence, mission scenario aware simulations are required to obtain more realistic predictions of the failure rates. In order to develop an appropriate simulation procedure, the thesis demonstrates that a methodology based on a 3 parameter characterization of the NBTI degradation state and a corresponding pre-generated lookup table facilitate a performance oriented yet varying mission scenario aware abstraction of physical timing degradation models. Besides, SDF files are identified as main interface to utilize industrial timing estimation procedures for the estimation of degraded timings of signal paths between registers.

Based on the main results of this thesis, follow-up research could initially focus on a reduction of the number of executed phase space simulations within the gate level degradation analysis. As the stress scenarios of different transistors within the same RT component usually only differ in stress probabilities but share the same temperature, supply voltage and off state profiles, it is not necessary to calculate each threshold voltage degradation separately. Furthermore, phase space simulations may also only be performed for particular stress probabilities and the correlation between stress probability and threshold voltage degradation that is already implied in Figure 7.1, may be utilized to interpolate the degradation caused by arbitrary stress probabilities. The performance of the gate level degradation analysis may thus be easily improved as a fixed number of phase space simulations are performed that is independent on the actual number of CMOS gates within the RT component. In addition, the phase space NBTI model may also be utilized as a major component within a completely different degradation analysis procedure. This is already shown by [56], which uses the phase space approach within a RT level model for the prediction of degraded timings.

Acknowledgment

Firstly, I would like to express my sincere gratitude to my advisor Prof. Dr.-Ing. Wolfgang Nebel for the continuous support of my Ph.D study and related research, for his patience, motivation, and immense knowledge. Besides my advisor, I would like to thank the rest of my thesis committee: Prof. Dr.-Ing. Steffen Paul, Prof. Dr.-Ing. Sergej Fatikow and Dr. Albert Sill for their insightful comments and encouragement.

A very special thanks goes out to Prof. Dr. Tibor Grasser and Gerhard Rzepa from TU Vienna, who provided me with detailed insights into physical NBTI modeling.

My sincere thanks also goes to Dr. Domenik Helms, who provided me an opportunity to join OFFIS. His guidance helped me in all the time of research and writing of this thesis. I am also thankful to Malte Metzdorf and Nils Koppaetzky for discussions that induced me to widen my research from various perspectives. I would also like to thank my other colleagues of the ANI and HDM groups for technical assistance, exchanges of knowledge, constructive feedback and venting of frustration.

I owe my sincere gratitude and a very special thanks to my parents, sister and parents-in-law for bearing me through the thick and thin of my Ph.D tenure. If it wasn't for their support, co-operation and encouragement in times of sickness this endeavor would not have been possible.

I bow in ovation to my wife, Anne, and my daughters, Jette and Jule, for their care and kindness. They have not only supported me with patience and tolerance but have also shown the course of directions by righteously correcting me in many ways from time to time. My words of thanks cannot compensate their contribution, yet with all humility I thank them for their noble gestures and splendid support.

Reef Eilers
Oldenburg
February 2017

List of Tables

List of Figures

Bibliography

[1] Kaushik Roy, Saibal Mukhopadhyay, and Hamid Mahmoodi-Meimand. “Leakage Current Mechanisms and Leakage Reduction Techniques in Deep-Submicrometer CMOS Circuits”. In: *Proceedings of the IEEE* 91.2 (2003).

[2] Jens Lienig. “Introduction toElectromigration-AwarePhysical Design”. In: *Proceedings of International Symposium on Physical Design* (2006), pp. 39–46.

[3] Salvatore Lombardo, James H. Stathis, Barry P. Linder, Kin Leong Pey, Felix Palumbo, and Chih Hang Tung. “Dielectric breakdown mechanisms in gate oxides”. In: *Journal of Applied Physics* 98 (2005).

[4] Tibor Grasser, B. Kaczer, W. Goes, Th. Aichinger, Ph. Hehenberger, and M. Nelhiebel. “A Two-Stage Model for Negative Bias Temperature Instability”. In: *Proceedings of IEEE International Reliability Physics Symposium* (2009).

[5] S.E. Tyaginov, I.A. Starkov, H. Enichlmair, J.M. Park, Ch. Jungemann, and T. Grasser. “Physics-Based Hot-Carrier Degradation Models”. In: *Electrochemical Society Transactions* 35.4 (2011).

[6] Mihir Choudhury, Vikas Chandra, Kartik Mohanram, and Robert Aitken. “Analytical model for TDDB-based performance degradation in combinational logic”. In: *Proceedings of Design, Automation & Test in Europe Conference* (2010).

[7] Malte Metzdorf, Reef Eilers, Domenik Helms, and Wolfgang Nebel. “Design methodology for a self–healing signal processing unit on the example of an FIR filter”. In: *Proceedings of edaWorkshop* (2012).

[8] Chris Angelini. *Intel Identifies Cougar Point Chipset Error, Halts Shipments*. 2011. URL: `http://www.tomshardware.com/news/cougar-point-sandy-bridge-sata-error,12108.html` (visited on 08/31/2016).

[9] Emmanuel Vincent. “Long Term Reliability of CMOS 65nm and 45nm”. In: *Workshop on "The issues of Long Term Reliability for industrial applications: Deep Sub Micron Challenges and perspectives" at European Symposium on Reliability of Electron Devices, Failure Physics and Analysis* (2008). STMicroelectronics Technology R&D, Crolles, France.

[10] Yun Xiang, Thidapat Chantem, Robert P. Dick, X. Sharon Hu, and Li Shang. “System-level reliability modeling for MPSoCs”. In: *Proceedings of IEEE/ACM/IFIP International Conference on Hardware/Software Codesign and System Synthesis* (2010), pp. 297–306.

[11] Cheng Zhuo, Kaviraj Chopra, Dennis Sylvester, and David Blaauw. “Process Variation and Temperature-Aware Full Chip Oxide Breakdown Reliability Analysis”. In: *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems* 30.9 (2011), pp. 1321–1334.

[12] Domenik Helms, Kim Grüttner, Reef Eilers, Malte Metzdorf, Kai Hylla, Frank Poppen, and Wolfgang Nebel. “Considering Variation and Aging in a Full Chip Design Methodology at System Level”. In: *Proceedings of Electronic System Level Synthesis Conference* (2014).

[13] Sanjay V. Kumar, Chris H. Kim, and Sachin S. Sapatnekar. "NBTI-Aware Synthesis of Digital Circuits". In: *Proceedings of ACM/IEEE Design Automation Conference* (2007), pp. 370–375.

[14] Laurence W. Nagel and D.O. Pederson. "SPICE (Simulation Program with Integrated Circuit Emphasis)". In: UCB/ERL M382 (1973). URL: `http://www2.eecs.berkeley.edu/Pubs/TechRpts/1973/22871.html`.

[15] Dominik Lorenz, Martin Barke, Daniel Mueller-Gritschneder, Georg Georgakos, and Ulf Schlichtmann. "Aging model for timing analysis at register-transfer-level". In: *Proceedings of ACM/IEEE International Workshop on Timing Issues in the Specification and Synthesis of Digital Systems* (2010).

[16] Vincent Huard, E. Pion, F. Cacho, D. Croain, V. Robert, R. Delater, P. Mergault, S. Engels, P. Flatresse, N. Ruiz Amador, and L. Anghel. "A Predictive Bottom-up Hierarchical Approach to Digital System Reliability". In: *Proceedings of IEEE International Reliability Physics Symposium* (2012).

[17] Yu Cao, Jyothi Velamala, Ketul Sutaria, Mike Shuo-Wei Chen, Jonathan Ahlbin, Ivan Sanchez Esqueda, Michael Bajura, and Michael Fritze. "Cross-Layer Modeling and Simulation of Circuit Reliability". In: *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems* 33.1 (2014), pp. 8–23.

[18] 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).

[19] Hans Reisinger, Oliver Blank, Wolfgang Heinrigs, Wolfgang Gustin, and Christian Schlünder. "A Comparison of Very Fast to Very Slow Components in Degradation and Recovery Due to NBTI and Bulk Hole Trapping to Existing Physical Models". In: *IEEE Transactions on Device and Materials Reliability* 7.1 (2007), pp. 119–129.

[20] C. Shen, M.-F. Li, C.E. Foo, T. Yang, D.M. Huang, A. Yap, G.S. Samudra, and Y.-C. Yeo. "Characterization and Physical Origin of Fast V_{th} Transient in NBTI of pMOSFETs with SiON Dielectric". In: *Proceedings of International Electron Devices Meeting* (2006).

[21] Tibor Grasser, B. Kaczer, H. Reisinger, P.-J. Wagner, and M. Toledano-Luque. "On the Frequency Dependence of the Bias Temperature Instability". In: *Proceedings of IEEE International Reliability Physics Symposium* (2012).

[22] Tibor Grasser, Ben Kaczer, Wolfgang Goes, Hans Reisinger, Thomas Aichinger, Philipp Hehenberger, Paul-Jürgen Wagner, Franz Schanovsky, Jacopo Franco, María Toledano Luque, and Michael Nelhiebel. "The Paradigm Shift in Understanding the Bias Temperature Instability: From Reaction–Diffusion to Switching Oxide Traps". In: *IEEE Transactions on Electron Devices* (2011).

[23] Moonju Cho, Philippe Roussel, Ben Kaczer, Robin Degraeve, Jacopo Franco, Marc Aoulaiche, Thomas Chiarella, Thomas Kauerauf, Naoto Horiguchi, and Guido Groeseneken. "Channel Hot Carrier Degradation Mechanism in Long/Short Channel n-FinFETs". In: *IEEE Transactions on Electron Devices* 60.12 (2013).

[24] Bipul C. Paul, Kunhyuk Kang, Haldun Kufluoglu, Muhammad A. Alam, and Kaushik Roy. "Impact of NBTI on the Temporal Performance Degradation of Digital Circuits". In: *IEEE Electron Device Letters* 26.8 (2005), pp. 560–562.

[25] NanGate Inc. *NanGate FreePDK45 Open Cell Library*. 2011. URL: `http://www.nangate.com/?page_id=2325` (visited on 08/28/2016).

[26] IEC/IEEE Standard 61523-3/1497. *Delay and power calculation standards - Part 3: Standard Delay Format (SDF) for the electronic design process*. 2004.

[27] Robert B. Hitchcock, Gordon L. Smith, and David D. Cheng. "Timing Analysis of Computer Hardware". In: *IBM Journal of Research and Development* 26.1 (1982), pp. 100–105.

[28] Synopsys Inc. *Switching Activity Interchange Format*. 2016. URL: `http://www.synopsys.com/community/interoperability/pages/tapinsaif.aspx` (visited on 08/30/2016).

[29] Shigeo Ogawa and Noboru Shiono. “Generalized diffusion-reaction model for the low-field cahrge-buildup instability at the Si-SiO_2 interface”. In: *Physical Review B* 51.7 (1995).

[30] S. Chakravarthy, A.T. Krishnan, V. Reddy, C.F. Machala, and S. Krishnan. “A Comprehensive Framework For Predictive Modeling of Negative Bias Temperature Instability”. In: *Proceedings of IEEE International Reliability Physics Symposium* (2004), pp. 273–282.

[31] M.A. Alam and S. Mahapatra. “A comprehensive model of PMOS NBTI degradation”. In: *Microelectronics Reliability* 45 (2005), pp. 71–81.

[32] Felix Salfelder and Lars Hedrich. “An NBTI model for efficient transient simulation of analogue circuits”. In: *Proceedings of edaWorkshop* (2011).

[33] Jiro Ushio, Yutaka Okuyama, and Takuya Maruizumi. “Electric-field dependence of negative-bias temperature instability”. In: *Journal of Applied Physics* 97 (2005).

[34] Souvik Mahapatra, Vrajesh D. Maheta, Ahmad Ehteshamul Islam, and Muhammad Ashraful Alam. “Isolation of NBTI Stress Generated Interface Trap and Hole-Trapping Components in PNO p-MOSFETs”. In: *IEEE Transactions on Electron Devices* (2008).

[35] Gerhard Rzepa. “Microscopic Modeling of NBTI in MOS Transistors”. MA thesis. Vienna University of Technology, 2013.

[36] Yannick Wimmer, Al-Moatasem El-Sayed, Wolfgang Gös, Tibor Grasser, and Alexander L. Shlunger. “Role of hydrogen in volatile behaviour of defects in SiO_2-based electronic devices”. In: *Proceedings of the Royal Society A - Mathematical, Physical and Engineering Sciences* 472 (2016), pp. 1–23.

[37] Kay-Uwe Giering, Christoph Sohrmann, Gerhard Rzepa, Leonhard Heiß, Tibor Grasser, and Roland Jancke. “NBTI modeling in analog circuits and its application to long-term aging simulations”. In: *IEEE International Integrated Reliability Workshop Final Report* (2014), pp. 29–34.

[38] Hans Reisinger, Tibor Grasser, Wolfgang Gustin, and Christian Schlünder. “The statistical analysis of individual defects constituting NBTI and its implications for modeling DC- and AC-stress”. In: *Proceedings of IEEE International Reliability Physics Symposium* (2010).

[39] Wenping Wang, Shengqi Yang, Sarvesh Bhardwaj, Rakesh Vattikonda, Sarma Vrudhula, Frank Liu, and Yu Cao. “The Impact of NBTI on the Performance of Combinational and Sequential Circuits”. In: *Proceedings of ACM/IEEE Design Automation Conference* (2007), pp. 364–369.

[40] Boby Bindu, Wolfgang Goes, Ben Kaczer, and Tibor Grasser. “Analytical Solution of the Switching Trap Model for Negative Bias Temperature Stress”. In: *IEEE International Integrated Reliability Workshop Final Report* (2009), pp. 93–96.

[41] Martin Barke, Veit Kleeberger, Christoph Werner, Doris Schmitt-Landsiedel, and Ulf Schlichtmann. “Analysis of Aging Mitigation Techniques for Digital Circuits Considering Recovery Effects”. In: *Proceedings of edaWorkshop* (2013), pp. 19–24.

[42] Martin Barke. “Aging Aware Robustness Validation of Digital Integrated Circuits”. PhD thesis. 2014.

[43] Simon Tam, Ping-Keung Ko, and Chenming Hu. “Lucky-Electron Model of Channel Hot-Electron Injection in MOSFET’s”. In: *IEEE Transactions on Electron Devices* 31.9 (1984).

[44] Vincent Huard, Florian Cacho, Xavier Federspiel, and Pascal Mora. “Hot-Carrier Injection Degradation in Advanced CMOS Nodes: A Bottom-Up Approach to Circuit and System Reliability”. In: *Hot Carrier Degradation in Semiconductor Devices*. Ed. by Tibor Grasser. 2014, pp. 401–444.

[45] Yury Illarionov, Anderson Smith, Sam Vaziri, Mikael Ostling, Thomas Mueller, Max Lemme, and Tibor Grasser. “Hot-Carrier Degradation and Bias-Temperature Instability in Single-Layer Graphene Field-Effect Transistors: Similarities and Differences”. In: *IEEE Transactions on Electron Devices* 62.11 (2015).

[46] Dominik Lorenz, Georg Georgakos, and Ulf Schlichtmann. "Aging Analysis of Circuit Timing Considering NBTI and HCI". In: *Proceedings of IEEE International On-Line Testing Symposium* (2009), pp. 3–8.

[47] Cadence. *Reliability Simulation in Integrated Circuit Design*. White paper. 2003.

[48] Amir H. Ajami, Kaustav Banerjee, Amit Mehrotra, and Massoud Pedram. "Analysis of IR-Drop Scaling with Implications for Deep Submicron P/G Network Designs". In: *Proceedings of IEEE International Symposium on Quality Electronic Design* (2003).

[49] Michael DeBole, K. Ramakrishnan, Varsha Balakrishnan, Wenping Wang, Hong Luo, Yu Wang, Yuan Xie, Yu Cao, and N. Vijaykrishnan. "A Framework for Estimating NBTI Degradation of Microarchitectural Components". In: *Proceedings of Asia and South Pacific Design Automation Conference* (2009), pp. 455–460.

[50] Dominik Lorenz. "Aging Analysis of Digital Integrated Circuits". PhD thesis. 2012.

[51] Evelyn Mintarno, Vikas Chandra, David Pietromonaco, Robert Aitken, and Robert W. Dutton. "Workload dependent NBTI and PBTI analysis for a sub-45nm commercial microprocessor". In: *Proceedings of IEEE International Reliability Physics Symposium* (2013).

[52] Halil Kükner, Moustafa Khatib, Sébastien Morrison, Pieter Weckx, Praveen Raghavan, Ben Kaczer, Francky Catthoor, Liesbet Van der Perre, Rudy Lauwereins, and Guido Groeseneken. "Degradation Analysis of Datapath Logic Subblocks under NBTI Aging in FinFET Technology". In: *Proceedings of International Symposium on Quality Electronic Design* (2014), pp. 473–479.

[53] N. Ruiz Amador, V. Huard, E. Pion, F. Cacho, D. Croain, V. Robert, S. Engels, P. Flatresse, and L. Anghel. "Bottom-up digital system-level reliability modeling". In: *Proceedings of IEEE Custom Integrated Circuits Conference* (2011).

[54] Dominik Lorenz, Martin Barke, and Ulf Schlichtmann. "Aging analysis at gate and macro cell level". In: *Proceedings of IEEE/ACM International Conference on Computer-Aided Design* (2010), pp. 77–84.

[55] Malte Metzdorf, Reef Eilers, Domenik Helms, Kay-Uwe Giering, Roland Jancke, Gerhard Rzepa, Tibor Grasser, Markus Karner, Praveen Raghavan, Ben Kaczer, Dan Alexandrescu, Adrian Evans, Gunnar Rott, Peter Rotter, Hans Reisinger, Wolfgang Gustin, and Wolfgang Nebel. "MoRV: Modelling Reliability under Variability". In: *Proceedings of Silicon Errors in Logic - System Effects* (2016).

[56] Nils Koppaetzky, Malte Metzdorf, Reef Eilers, Domenik Helms, and Wolfgang Nebel. "RT Level Timing Model for Aging Prediction". In: *Proceedings of Design, Automation & Test in Europe Conference* (2016).

[57] Malte Metzdorf, Domenik Helms, Reef Eilers, and Wolfgang Nebel. "Abstracting TCAD aging models above the circuit level". In: *Proceedings of Design, Automation & Test in Europe Conference* (2015).

[58] G. Rzepa, M. Waltl, W. Goes, B. Kaczer, and T. Grasser. "Microscopic Oxide Defects Causing BTI, RTN, and SILC on High-K FinFETs". In: *Proceedings of International Conference on Simulation of Semiconductor Processes and Devices* (2015).

[59] Gunnar Andreas Rott, Karina Rott, Hans Reisinger, Wolfgang Gustin, and Tibor Grasser. "Mixture of negative bias temperature instability and hot-carrier driven threshold voltage degradation of 130 nm technology p-channel transistors". In: *Microelectronics Reliability* 54.9 (2014), pp. 2310–2314.

[60] Ahmet Unutulmaz, Domenik Helms, Reef Eilers, Malte Metzdorf, Ben Kaczer, and Wolfgang Nebel. "Analysis of NBTI Effects on High Frequency Digital Circuits". In: *Proceedings of Design, Automation & Test in Europe Conference* (2016).

[61] Reef Eilers, Malte Metzdorf, Sven Rosinger, Domenik Helms, and Wolfgang Nebel. "Phase space based NBTI model". In: *Proceedings of Integrated circuit and system design: Power and timing modeling, optimization and simulation* (2012).

[62] Domenik Helms, Reef Eilers, Frank Oppenheimer, and Wolfgang Nebel. “RELY – Design for Reliability of SoCs”. In: *Proceedings of IEEE European Test Symposium* (2013).

[63] M.A. Alam, H. Kufluoglu, D. Varghese, and S. Mahapatra. “A comprehensive model for PMOS NBTI degradation: Recent progress”. In: *Microelectronics Reliability* 47 (2007), pp. 853–862.

[64] Reef Eilers, Malte Metzdorf, Domenik Helms, and Wolfgang Nebel. “Efficient NBTI Modeling Technique Considering Recovery Effects”. In: *Proceedings of International Symposium on Low Power Electronics and Design* (2014).

[65] Reef Eilers, Carl Bergemann, Domenik Helms, and Wolfgang Nebel. “Abstraktion des Switching-Trap-Modells für eine leistungsorientierte NBTI-Simulation”. In: *Proceedings of edaWorkshop* (2013), pp. 59–64.

[66] Hans Reisinger, Tibor Grasser, Karsten Ermisch, Heiko Nielen, Wolfgang Gustin, and Christian Schlünder. “Understanding and Modeling AC BTI”. In: *Proceedings of IEEE International Reliability Physics Symposium* (2011).

[67] Marko Hoyer, Domenik Helms, and Wolfgang Nebel. “Modelling the Impact of High Level Leakage Optimization Techniques on the Delay of RT-Components”. In: *Proceedings of Integrated circuit and system design: Power and timing modeling, optimization and simulation* (2007), pp. 171–180.

[68] M.S. Hrishikesh, Doug Burger, Norman P. Jouppi, Stephen W. Keckler, Keith I. Farkas, and Premkishore Shivakumar. “The Optimal Logic Depth Per Pipeline Stage is 6 to 8 FO4 Inverter Delays”. In: *Proceedings of the 29th Annual International Symposium on Computer Architecture* (2002), pp. 14–24.

[69] Tibor Grasser, Hans Reisinger, Paul-Jürgen Wagner, and Ben Kaczer. “Time-dependent defect spectroscopy for characterization of border traps in metal-oxide-semiconductor transistors”. In: *Physical Review B* 82 (2010).